GENETIC SCREENING OF PATIENTS WITH BLADDER EXSTROPHY EPISPADIAS COMPLEX

By

Istiak Mahfuz

Thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

Hudson Institute of Medical Research

June 2015

Dedication

Dedicated
To BEEC Patients and their Families

Table of Contents

List of Tables

List of Appendix Table

List of Figures

Abstract

Bladder Exstrophy Epispadias Complex (BEEC) is a serious congenital anomaly. Babies affected are born with an exposed bladder, and leak urine continually. The high incidence among the offspring, siblings and identical twins of BEEC patients suggests a genetic aetiology. We previously showed that *TP63* is a candidate gene of this disorder. There are two main isoforms, *TAP63* and *ΔNP63*, and *ΔNp63* is expressed predominantly in the murine bladder epithelium. Notably, *p63$^{-/-}$* mice developed a condition similar to BEEC. In addition, the *p63$^{-/-}$* mutation led to an increase in apoptosis, decrease in cell proliferation and absence of smooth muscle development in the bladder. We found that *ΔNP63* expression is reduced in BEEC patients' bladders, whereas *TAP63* expression is increased. Although, no mutations were found in *P63* coding sequence, we did identify in/del polymorphisms in the *ΔNP63* promoter that were associated with significantly increased risk of BEEC. The analysis of the *TAP63* promoter region for possible mutations did not reveal any variants that might contribute to the risk of BEEC.

Study of the literature showed that *p63* acts upstream of *Perp,* and that *Perp* and desmosomes co-localize in the plasma membrane to form a desmosomal adhesive assembly complex to maintain tissue integrity. These findings about *p63*, *Perp* and Desmosome lead to a new hypothesis that the *TP63→PERP↔Desmosome* pathway may play a major role in BEEC pathogenesis. In order to determine the functional significance of the pathway, mutation screening was carried out in *PERP* and *DSP* exons and promoters. Although several common SNPs were found in *DSP* and *PERP* coding regions, no rare, potentially causative variants were identified. Interestingly, as

part of this study, bacterial DNA contamination was found in buccal DNA which can be a major component of buccal-derived DNA samples. This has significant implications for other studies using buccal DNA for genetic analysis.

To further explore the role of epigenetic regulation in BEEC pathogenesis, ChIP analysis with Histone antibody (H3k4me3 and H3K27ac) and a high throughput sequencing (ChIP-seq) was carried out in normal human urothelial cells. Our genome wide mapping of H3K4me3 and H3K27ac enrichment, and also the overlapping regions of these histone modifications with TP63 binding, showed that intron 1 of *PERP* and *DSP* are candidate loci for BEEC pathogenesis. *De novo* motif searches on H3K4me3 and H3K27ac enriched regions have identified significant enrichment of IRF2 and SCL transcription factor, respectively. IRF2 recruits RELA in the nucleus to activate NFƙB complex, which in turn expresses cell adhesion molecules and maintains epithelial integrity during development via IRF-RELA-NFƙB signalling pathway. SCL acts via the *SHH* pathway to induce mesenchyme development, which in turn is involved in epithelial-mesenchymal interaction in the development of bladder smooth muscle via SCL-SHH pathway. Based on our epigenetic findings, we have speculated the significant involvement of the IRF-RELA-NFƙB and SCL-SHH pathways in the aetiology of BEEC.

Overall, this research will lead to a better understanding of BEEC pathogenesis. Moreover, the epigenetic data will be a valuable resource for future BEEC studies, in order to better determine the role of genetic and epigenetic variation in BEEC pathogenesis.

PART A: General Declaration

Monash University

Declaration for thesis based or partially based on conjointly published or unpublished work

General Declaration

In accordance with Monash University Doctorate Regulation 17.2 Doctor of Philosophy and Research Master's regulations the following declarations are made:

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes 4 original papers published in peer reviewed journals and 1 unpublished publications. The core theme of the thesis is 'Genetic screening of patient with BEEC'. The ideas, development and writing up of all the chapters in the thesis were the principal responsibility of myself, the candidate, working within the MIMR-PHI Institute of Medical Research under the supervision of Professor Wei Cheng.
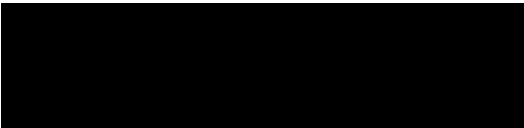
The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of chapter 1 to 6 my contribution to the work involved the following:

Thesis chapter	Publication title	Publication status*	Nature and extent of candidate's contribution
1	✓ New insight into the pathogenesis of Bladder Exstrophy Epispadias Complex.	Published	Drafted manuscript, Review of literature, Conceived the idea
2	N/A	N/A	Writing and preparing drafts
3	✓ Insertion/deletion polymorphisms in the DeltaNp63 promoter are a risk factor for bladder exstrophy epispadias complex.	Published	Performed experiment, Analyzed data, Contributed to the manuscript
3	✓ No TAP63 promoter mutation is detected in bladder exstrophy-epispadias complex patients.	Published	Designed experiment, Analyzed data, Contributed to the manuscript
4	✓ Mutation screening of *TP63*-related target genes *PERP* and *DSP* in Bladder Exstrophy Epispadias Complex	Submitted	Designed and performed experiments, Analyzed the data, Drafted manuscript
4	✓ Identification of *Streptococcus parasanguinis* DNA contamination in human buccal DNA samples.	Published	Designed and performed the experiment, Analyzed the data, Drafted manuscript
5	N/A	N/A	Writing and preparing drafts
6	N/A	N/A	Writing and preparing drafts

[* For example, 'published'/ 'in press'/ 'accepted'/ 'returned for revision']

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Signed: …███████████████████████…

Date: …………………………………………………
27th December 2014

Acknowledgements

A heart full of thanks is due to my fellow colleagues Dr. Simon Wilkins, Dr Rui Liang, and Suzan de Boer for help and support during the initial phase of my PhD studies. I would also like to convey a heart full of indebtedness to all other esteemed colleagues and researchers of the MIMR-PHI Institute of Medical Research, for their kind cooperation, valuable suggestions, encouragement and inspiration during the present research work, which will remain ever alive in the sanctuary of my heart.

Special thanks to all of my friends who inspired me and gave me incentives to strive towards my goal.

Finally, I would like to convey an ocean of gratefulness, devotion and indebtedness to my parents and parents in law, my younger brother and sister, and sister in law for their inspiration. Words cannot express how grateful I am to my family for the sacrifices they made during this period of study. Your prayer for me was what sustained me thus far.

At last but not the least, I would like express appreciation to my beloved wife, Sadia Sabrina Alam, who spent sleepless nights along with me and was always my support in the moments when there was no one to answer my queries. I am also grateful to her for her continuous encouragement, inspiration, selfless support and nurturing me during the ups and downs for the whole tenure of the study.

Istiak Mahfuz

Chapter 1

General Introduction

PART B: Suggested Declaration for Thesis Chapter

Monash University

Declaration for Thesis Chapter 1

Declaration by candidate

In the case of Chapter 1 (Based on the article: New insights into the pathogenesis of bladder exstrophy-epispadias complex), the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Drafted and prepared the manuscript, Review of literature, Conceived the idea	80%

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Darling, T.	Contributed, read and approved the final manuscript	5%
Wilkins, S.	Contributed, read and approved the final manuscript	
White, S.	Coordination and drafted the manuscript	
Cheng, W.*	Conceived the idea, Coordination and drafting of the manuscript	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work*.

Candidate's Signature		Date: 30/12/2014
Main Supervisor's Signature		Date: 30/12/2014

*Note: Where the responsible author is not the candidate's main supervisor, the main supervisor should consult with the responsible author to agree on the respective contributions of the authors.

Chapter 1: General Introduction

This chapter is based on my published review, outlining new molecular insights into the pathogenesis of bladder exstrophy epispadias complex (Mahfuz et al., 2013b).

1.1 Bladder Exstrophy Epispadias Complex- Clinical Background

The urinary bladder is the largest smooth muscle organ in human body. The bladder serves to collect, store and pass urine in a highly coordinated fashion, and is present only in placental mammals and teleost fish. In mammals, the bladder develops from the ventral aspect of the cloaca. Congenital urological abnormalities such as Bladder Exstrophy Epispadias Complex (BEEC) greatly impair bladder function and quality of life. BEEC is manifested as a cluster of ventral midline defects (Gearhart and Jeffs, 1989) (Figure 1. A-D) which include; i) separation of the external genitalia (epispadias), ii) separation of the pubic bones and the rectus abdominis muscles, iii) ventral bladder and abdominal wall defects (bladder exstrophy), iv) herniation of intestines via a ventral abdominal defect (exomphalos), v) ventrally displaced or imperforate anus, and vi) additional ventral defect in hind gut (cloacal exstrophy). The spectrum of severity ranges from Epispadias to Cloacal Exstrophy, the most severe form.

The overall incidence of BEEC is 1 in 30,000 live births (Ebert et al., 2009), with a greater proportion of males affected. Without treatment, affected babies continuously leak urine, resulting in skin excoriation, a progressive decline in renal function, a constant stench, and severe psychosocial strain for both patients and parents. BEEC patients also have a 700 fold increased chance of developing bladder cancer. Complications such as urinary and faecal incontinence (Ricketts et al., 1991), reduced bladder volume, genital disfigurement, sexual dysfunction, and psychological stress

(Reiner and Gearhart, 2004) are common and debilitating. Treatment of BEEC requires a series of major reconstructive surgeries. Multiple operations and frequent hospital visits result in a lifelong medical, social and financial burden for patients, parents and the community at large.
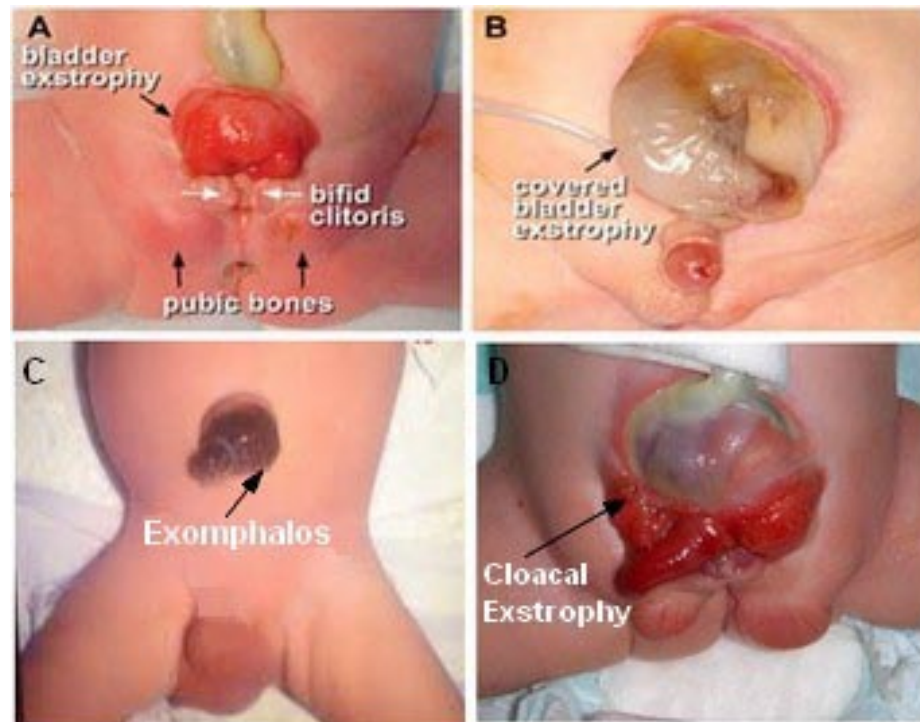
Figure 1: Different BEEC phenotypes. A. BEEC in a girl with exposed bladder and separated pubic symphysis. B. BEEC variant in a boy with bifid scrotum and covered bladder exstrophy C. Herniation of umbilical cord (exomphalos), D. Cloacal exstrophy in male (Image A,B courtesy of Dr. J. L. Salle, Hospital for Sick Children, Toronto, Canada)..

1.2 Bladder Structure and Development

1.2.1 Anatomy and Histology

The bladder is a hollow muscular organ forming the main urinary reservoir (Figure 2). The shape and size of the bladder varies according to the amount of urine the organ contains. Anatomically, the bladder consists of a dome or fundus and a relatively immobile base. The ureters, carrying urine from each kidney, enter the bladder dorso-

laterally at the ureteric orifices on each side. The area between the ureteric orifices and the urethra forms the trigone which is incorporated into the bladder from the ureters (Sinnatamby, 2011). Urine exits the bladder through the urethra which is controlled by a sphincter.
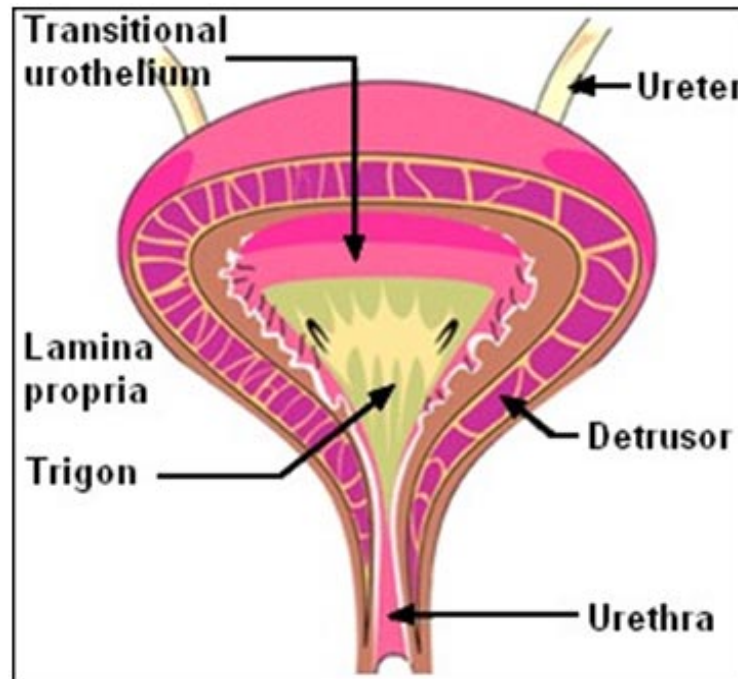
Figure 2: Structure of the bladder. Modified from Women's Health Matters (2015).

The waterproof urothelium is endodermal in origin except for the trigone area, which is mesodermal in origin. The urothelium consist of three different cell layers; the i) basal cells, ii) intermediate cells or transitional epithelium and iii) umbrella cells or superficial epithelium (Figure 3). The basal layer of this stratified epithelium expresses *p63* (Karni-Schmidt et al., 2011; Yang et al., 1999), a stem/progenitor cell marker in stratified epithelium (Figure 3, label 6). *p63* expression is also found in the intermediate cells of urothelium (Castillo-Martin et al., 2010). The most superficial layer contains water-proof umbrella cells which express uroplakin (Wu et al., 1994; Yu et al., 1990) (Figure 3, label 8) but not *p63* (Castillo-Martin et al., 2010).

3
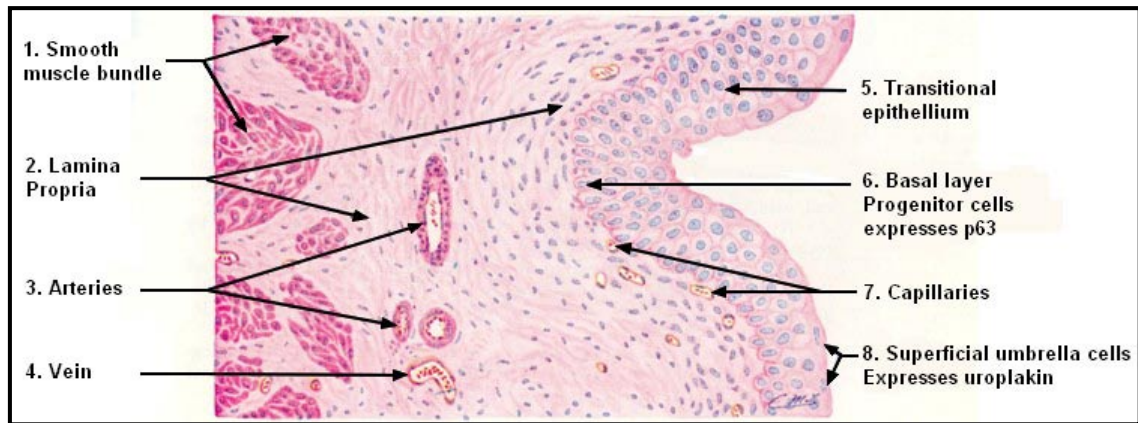
Figure 3: Histology of the Bladder. Modified from di Fiore and Schmidt (1981).

Beneath the urothelium lies a layer of connective tissue, the lamina propria. It consists of smooth-muscle called the detrusor muscle, organized into inner longitudinal, middle transverse and outer longitudinal layers, with fibres of each layer arranged in different directions. Unlike gut peristalsis, the detrusor muscle is tailored for mass contraction, resulting in the bladder emptying. The outermost layer of bladder is the serosa.

1.2.2 Embryological Development

The embryological development of the bladder is closely intertwined with that of the hindgut. During murine embryo gastrulation (E7.5), the embryo is stratified into three layers - the ectoderm, mesoderm and endoderm. The gut tube develops as a result of the folding of the endodermal layer, which begins at the anterior (anterior intestinal portal, AIP) and posterior (caudal intestinal portal, CIP) ends. This folding fuses at the ventral midline, at the level of the future umbilicus. Along the anterior-posterior axis, the gut develops into: i) the foregut (pharynx, oesophagus and stomach), ii) the midgut (small intestine) and iii) the hindgut (colon and bladder) (Moore and Persaud, 2003).

Each of the three germ layers contributes to the formation of the gut; the endoderm to the epithelium, the mesoderm to smooth muscle and ectoderm to the enteric nervous system as well as the most anterior and posterior parts of the gut (oral and anal lumens)

4

respectively. Along the radial axis, the gut assumes the general plan of epithelial lining, sub-mucosal (lamina propria), smooth muscle layers and outer lining of serosa (Dudek and Fix, 1998).
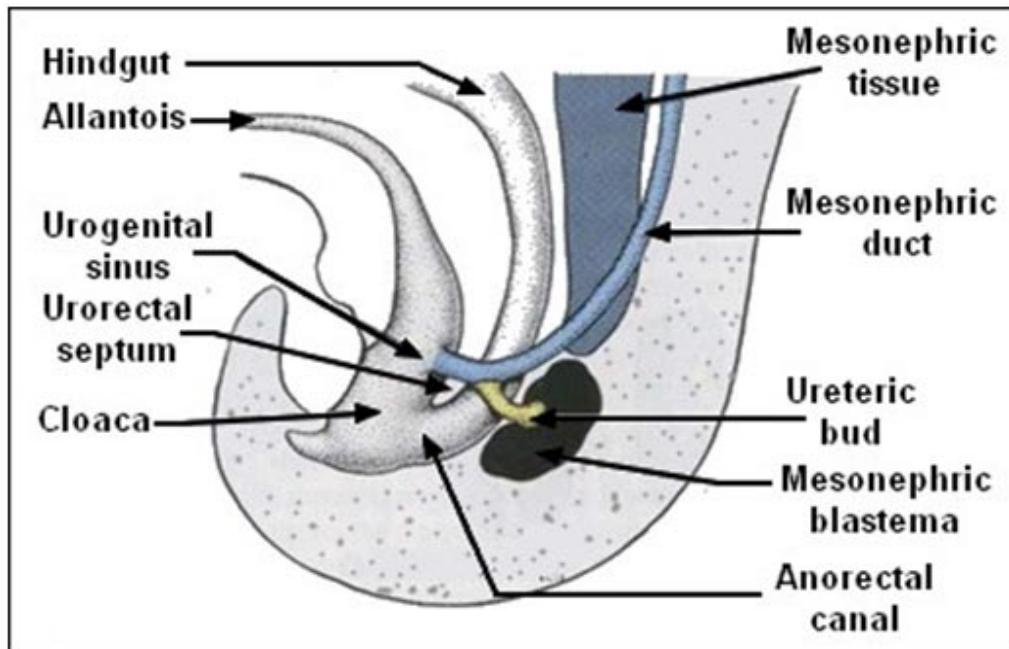
Figure 4: A sagittal representation of the cloaca being partitioned into urogenital sinus and anorectal canal. Modified from Langman and Sadler (1995).

During embryogenesis, allantois from the body stalk and hindgut converge into a cavity, the cloaca, at the posterior end of the embryo (Patten and Barry, 1952). The bladder then develops from the ventral part of cloaca. Between the 4th and 7th week of development (E12 in mice), the urorectal septum partitions the cloaca into the urogenital sinus ventrally and anorectal canal dorsally (Arey, 1965) (Figure 4). The urogenital sinus can be subdivided into four parts (Figure 5) (Skandalakis and Gray, 1994):
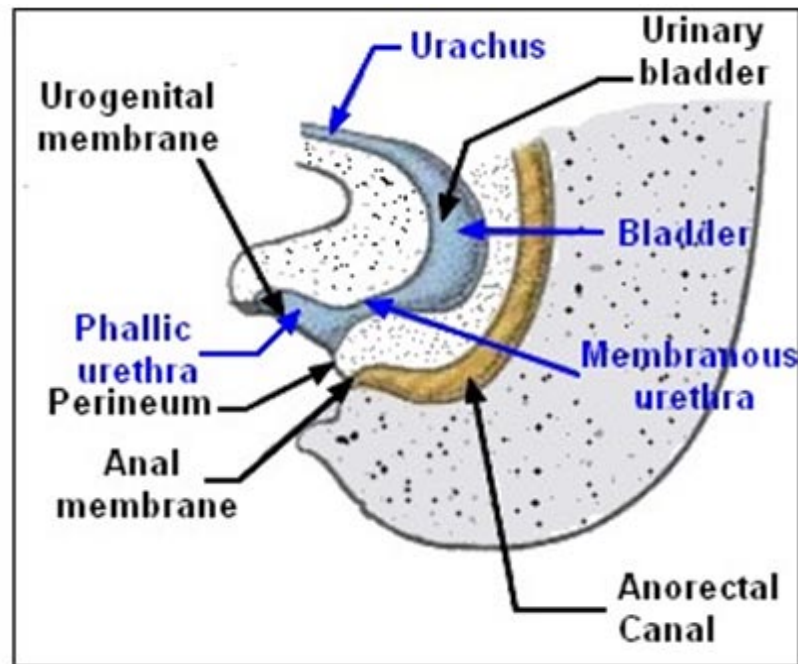
Figure 5: The four parts of urogenital sinus (in blue) - Urachus, Bladder, Membranous urethra and Phallic urethra. Modified from Langman and Sadler (1995).

i)   The most cranial part, the urachus, is contiguous with the allantois stalk. It is later obliterated and becomes a thick fibrous cord, the median umbilical ligament.

ii) The part where the ureters and urogenital sinus merges becomes the urinary bladder.

iii) The pelvic part of the urogenital sinus gives rise to prostate and membranous urethra.

iv) The distal part becomes the urethra.

1.3 BEEC pathogenesis

1.3.1 Anatomical and Embryological Theories

There are several theories concerning the development of bladder exstrophy during embryogenesis. The most conventional one is thought to be due to the failure of the cloacal membrane being reinforced by ingrowths of mesoderm (Muecke, 1964). Another popular theory is the abnormal overdevelopment of the cloacal membrane,

which prevents medial migration of the mesenchymal tissue and proper lower abdominal wall development (Marshall and Muecke, 1968). Other theories involve the development of the cloacal membrane and mesenchymal tissues (Ambrose and O'Brien, 1974; Bruch et al., 1996; Mildenberger et al., 1988; Patten and Barry, 1952). While the embryological process of bladder development and the disease pathogenesis is well-described, the molecular regulation of the development is poorly understood. Further explanations became possible with the advent of developmental biology and molecular genetics.

1.3.2 Environmental Factors

The study and identification of environmental factors contributing to BEEC is problematic. Due to the rarity of the phenotype, prospective population-based studies are not available to identify possible environmental factors. Therefore, information on phenotype expression and environmental risk factors is either limited or inconclusive (1987; Caton et al., 2007; Tang et al., 2006). Household exposure to pesticides during 0-12 weeks of gestation has been associated with classic bladder exstrophy in babies (Martin et al., 2009). Maternal smoking and medical radiation during the first trimester have also been associated with severe exstrophy phenotype (Reutter et al., 2011).

1.3.3 Genetic Explanations

Evidence for a genetic component in BEEC includes twin studies, which reveal that monozygotic twins are more prone (62%) to both having BEEC than dizygotic twins (11%) (Reutter et al., 2007b). In addition, the occurrence of BEEC increases 500 fold among siblings and offspring (Shapiro et al., 1984). Chromosomal anomalies have been reported in a small number of patients, shown in Table 1. There is also evidence of high risk genetic loci associated with this disease (Table 2). BEEC is significantly more

prevalent in children born to older mothers (Boyadjiev et al., 2004) and within Caucasian populations (Nelson et al., 2005). It was found that males were more common in both the epispadias group (M/F, 2.2, 29 patients) and the classic bladder-exstrophy group (M/F 1.8, 164 patients), but in the cloacal exstrophy group the sex ratio was close to unity (1.1, 15 patients). There was a statistically significant association with advanced parental age (P < 0.001). Birth weight, gestational age and maternal reproductive history did not appear to be significantly different to those in the general population (Boyadjiev et al., 2004).

Table 1: Reported chromosomal anomalies in BEEC patients.

Chromosomal Location	Type of mutation/genotype	Phenotype	Reference
-	47,XXY	E	Raboch (1975)
9p	Duplication [dup(9p)] of the short arm (47,XY)	E	Chipail et al. (1976)
-	Trisomy (47,XXX)	CE	Lin et al. (1993)
4p?-p?	Deletion in the short arm of chromosome 4 (46,XY)	E	Nicholls and Duffy (1998)
21	Duplication of chromosome 21 [dup(21)] (47,XX)	CE	Husmann and Vandersteen (1999)
-	45,X0/46,XX mosaicism	CE	
-	Trisomy (47,XXX)	CE	
-	Diploid/Tetraploid/t(1;6) mosaicism [in fibroblasts: 16% (3 cells) 92,XXXX; 11% (2 cells) 46,XX,t(1;6)(p32;q13); 73% (14 cells) 46,XX]	CE with hypomelanosis of Ito	Leonard and Tomkins (2002)
9q34.1-qter Deletion	*De novo* unbalanced translocation between chromosome 9q and Yq. 46,Xder(Y)t(Y;9)(q11.23;q34.1)–del(Y)(q11.2),der(9)t(Y;9)	CE	Thauvin-Robinet et al. (2004)
3q12.2-13.2	*De novo* deletion (46,XY)	CE	Kosaki et al. (2005)

Chromosomal Location	Type of mutation/genotype	Phenotype	Reference
t(8;9)(p11.2;q13)	Translocation between 8p11.2 and 9q13 (46,XY)	CBE	Boyadjiev et al. (2005)
t(2;9)(q13;q32)	Translocation between 2q13 and 9q32 (46,XY)	CBE	Ludwig et al. (2005)
22q11.21	*De novo* microduplication	CBE	Draaken et al. (2010)
1q43q44	*De novo* 10.4 Mb deletion	CBE with absencet phallus	Zaki et al. (2012)
19p13.12	*De novo* 0.9 Mb microduplication	CBE	Draaken et al. (2013)

Table 2: Genetic loci associated with BEEC pathogenesis.

Chromosomal Location	Description	Reference
4q31.21-22	Harbor genes for an autosomal recessive form of BEEC	Reutter et al. (2010)
19q13.31-41		
2p22.1-p21		
2p25.2-p25.1		
4q23-q32.3	Evidence for possible risk/modifying loci Classic Bladder exstrophy on chromosomes	Ludwig et al. (2009)
7q21.3-q33		
7q34-q36.1		
14q31.1-q32.2		
19q13.33-q13.43		

1.3.4 Molecular Pathogenesis

Epithelial-Mesenchymal Interaction

According to Baskin and co-workers, epithelial-mesenchymal interaction is essential in bladder mesenchymal (smooth muscle) development (Baskin et al., 1996a; Baskin et al., 1996b), *i.e.* a diffusible signal from the urothelium is responsible for inducing mesenchyme development (Figure 6A and 6B). The mesenchyme further develops into

the detrusor muscle. Haraguchi and colleagues have identified *Shh* (Sonic Hedgehog) gene as a diffusible signal involved in the external genitalia development (Haraguchi et al., 2007). In the bladder, multiple studies have demonstrated that *Shh* is involved in development (Figure 3) through the *Shh-Ptc1-Gli2-Bmp4* pathway (Cheng et al., 2008; Shiroyanagi et al., 2007).
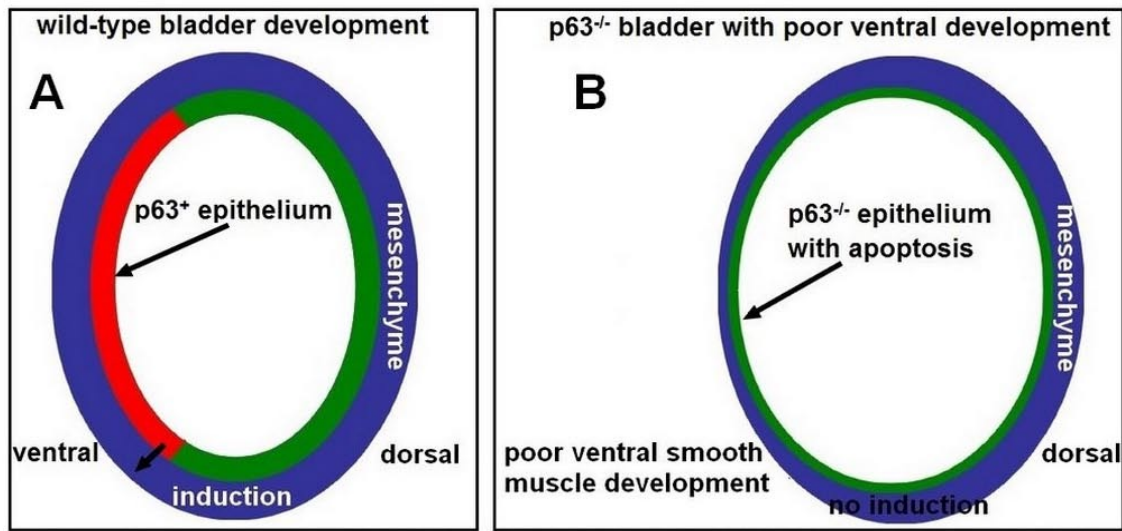
Figure 6: Epithelial-mesenchymal interaction in bladder development. A: *p63* is normally expressed in ventral bladder urothelium and is required for differentiation and production of signals required for inducing the normal developmental programme in adjacent mesenchyme. B: p63$^{-/-}$ bladder where ventral urothelium undergoes apoptosis. Absence of the epithelial signal results in the lack of mesenchymal induction and a lack of smooth muscle development. The figure is modified from Cheng et al. (2008).

Cheng et al. (2006) showed that the apoptotic ventral bladder epithelium in *p63*$^{-/-}$ mice may lead to reduced mesenchymal induction. The reduction in mesenchymal cell proliferation and failure of smooth muscle formation in turn may result in ventral midline defects, *i.e.* BEEC.

TP63 is a BEEC Candidate Gene

TP63 is important for basic cellular functions during the development of the ectoderm and its derivatives. It is essential for cell proliferation (Ihrie et al., 2005; Laurikkala et

10

al., 2006; Senoo et al., 2007; Yang et al., 1998). It has previously been shown that *Tp63* is a candidate gene for BEEC (Cheng et al., 2006). It is also a master regulator of stratified epithelial development (Koster et al., 2004), and is considered to be the molecular switch for the initiation of epithelial stratification. *P63* is a member of the *p53* tumour suppressor family expressed in all stratified epithelia (Barbieri and Pietenpol, 2006; Blanpain and Fuchs, 2007; Di Como et al., 2002; Ihrie and Attardi, 2005; Ihrie et al., 2005; McKeon, 2004). Carroll et al. (2006) observed that knockdown of *p63* down-regulates the cell adhesion-associated genes and conversely, over-expression of *p63* (*TAp63* or *ΔNp63*) up-regulates the adhesion complexes (Carroll et al., 2006). These findings implicate *p63* as a key regulator of cellular adhesion in stratified epithelial tissues.
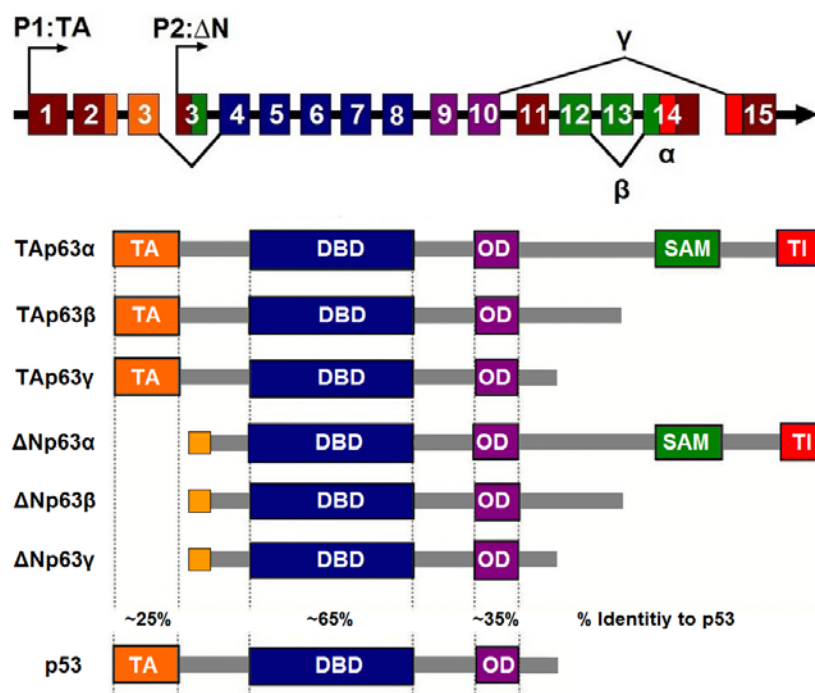
Figure 7: Structure of *TP63* gene and p63 protein. P1:TA and P2:ΔN are the two promoter for *TAP63* and *ΔNP63* isoform, respectively. TA: Transactivation Domain, DBD: DNA Binding Domain, OD: Oligomerization Domain, SAM: Sterile Alfa Motif, TI: Transcription Inhibitory Domain. Identity with p53 in TA, DBD and OD domain is shown in the bottom of the figure (% Identity to p53). The figure is modified from Yang and McKeon (2000).

Human *p63* has 15 exons with 2 promoters, *TAp63* and *ΔNp63,* located upstream of exon 1 and exon 3 respectively (Figure 7). Depending on the promoter there are two distinct N-terminal isoforms formed, i) *TAP63* (full-length) and *ΔNP63* (N-terminally truncated). *TAp63* isoforms are pro-apoptotic and the truncated *ΔNp63* is anti-apoptotic; these antagonise each other through competition for target genes (Koster and Roop, 2004; Yang et al., 1998). Due to alternative splicing at the 3' end, *TAP63* transcripts have three different C-terminal isoforms, *i.e. TAp63α, TAp63β* and *TAp63γ*. Likewise, the truncated *ΔNp63* also forms *ΔNp63α, ΔNp63β* and *ΔNp63γ* (Figure 7). P63 protein comprises DNA binding domain (DBD), oligomerization domain (OD) and N-terminal trans-activation (TA) domain (only TA isoforms). The alpha (α) isoforms of both TA and ΔN also possess a sterile alpha motif (SAM) and a transcription inhibitory domain (TI) (Figure 7) (Yang and McKeon, 2000).

Genetic Models of BEEC

Due to the rarity of BEEC in humans, a number of animal models were developed to explain its pathogenesis. One model is in the sheep, where foetal lambs underwent *in utero* surgical creation of classic bladder exstrophy (Slaughenhoupt et al., 1996). Another model is the chick embryo, where the fertilized eggs were windowed externally to inject suramin and/or trypan blue drug to induce cloacal exstrophy (Manner and Kluth, 2003). These models do not satisfactorily explain the complete BEEC spectrum reported in children, especially the pelvic bone anomaly. However, a better explanation of the pelvic abnormality was offered using a rabbit embryo model (Beaudoin et al., 2004). The above BEEC models generally explain anatomical and/or embryological aspects of the condition, but not the molecular basis.

A murine *p63*[-/-] knock-out model of BEEC was developed to give a better understanding the molecular conditions of this disease. Compared to wild type (Figure 8A), *p63*[-/-]

mouse embryos exhibit ventral midline defects (Figure 8B) identical to those of human BEEC, including bladder exstrophy, separation of pubic bones and split external genitalia (Cheng et al., 2006). The anti-apoptotic isoform of *p63*, *ΔNp63*, is predominantly expressed in foetal murine bladder epithelium. Compared to wild type mouse model (Figure 8C), *p63*$^{-/-}$ murine mutants lack stratified skin, and also have poorly stratified bladder epithelium (Cheng et al., 2006) (Figure 8D). They have a short tail, truncated limbs, and cleft palate. They die shortly after birth, presumably due to fluid loss as a consequence of poorly stratified skin epithelium (Mills et al., 1999; Yang et al., 1999).
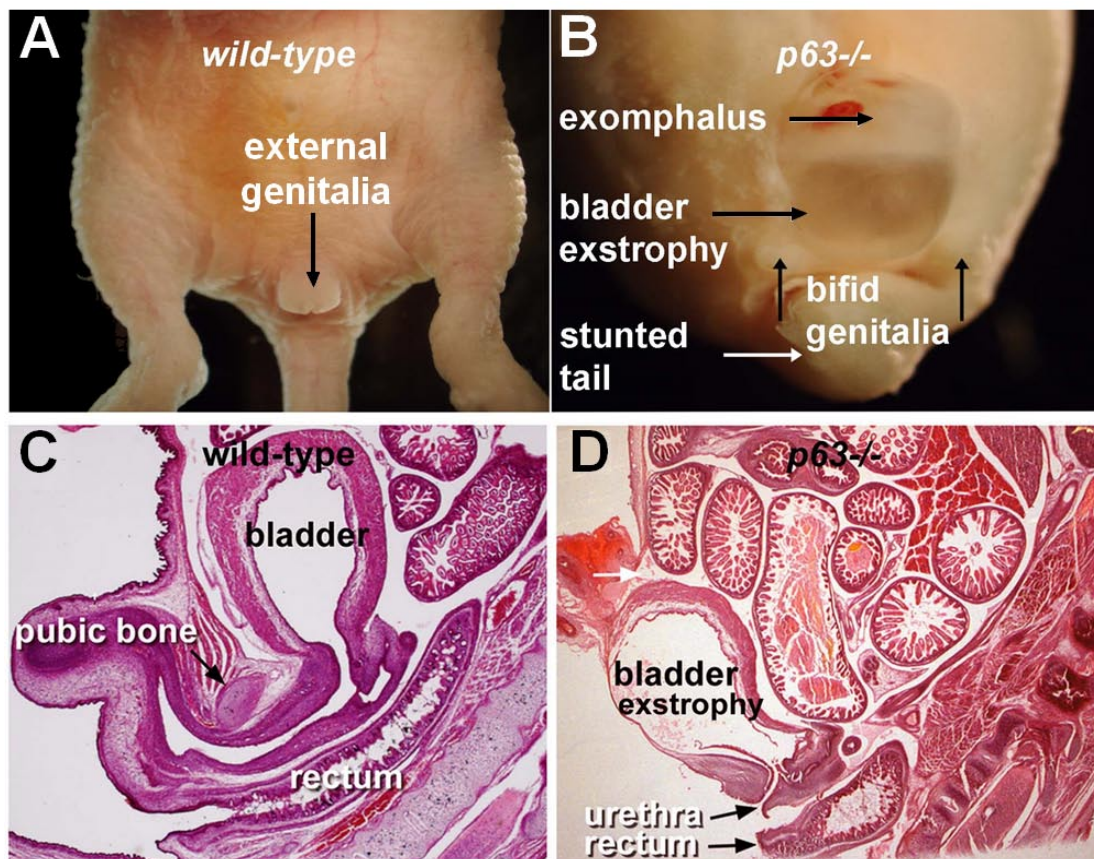
Figure 8: A: Wild-type E18 mouse embryo. B: *p63*$^{-/-}$ mouse embryo with bladder exstrophy, exomphalos (umbilical hernia), bifid genitalia (epispadias), stunted tail and absence of hind limbs. C: Histology of E18 wild-type foetus. D: *p63*$^{-/-}$ foetus with thin ventral bladder wall (Cheng et al., 2006). Arrows indicate the position of variation between wild-type and *p63*$^{-/-}$ mouse. The figure is modified from Cheng et al. 2006.

The dominant isoform, *ΔNp63,* is expressed throughout bladder organogenesis, especially along the genital tubercle and the ventral urothelium. Without *ΔNp63*, the ventral bladder epithelium is neither stratified nor differentiated. Moreover, the apoptotic activity of ventral urothelium of *p63$^{-/-}$* bladder is markedly increased. Loss of *ΔNp63* expression in the knock out model results an increase in apoptosis, decrease in cell proliferation and a loss of smooth muscle development in ventral bladder, *i.e.* BEEC. Based on this genetic model, *p63* is considered to be a BEEC candidate gene.

Involvement of PERP and DSP in BEEC

Qi et al. (2011) carried out a Genome-Wide Expression Profiling (GWEP) study in normal and exstrophic human bladder tissues, as well as human and mouse embryologic bladder precursor tissues. Their study has identified 162 genes differentially expressed in both embryonic and postnatal human samples, which may be candidate BEEC genes. 30% of these genes are directly associated with desmosome (cell-cell adhesion complex) structure/function or cytoskeletal assembly. This finding points to desmosomal and/or cytoskeletal deregulation as an etiologic factor for BEEC (Qi et al., 2011).

Based on the GWEP study, *PERP, SYNPO2* and the genes of *Wnt* pathway were reported to be candidate BEEC genes. *Perp* (*p53* Effector Related to *PMP22*) is a critical component of the desmosome, and is also required for maintaining the integrity of the stratified epithelia (Ihrie and Attardi, 2005; Ihrie et al., 2006; Ihrie et al., 2005; Ihrie et al., 2003; Marques et al., 2005; Marques et al., 2006). Moreover, whole-mount *in situ* hybridization studies have confirmed prominent *Perp* expression in bladder tissue (Ihrie et al., 2005). Interestingly, other studies show that *PERP* is a direct downstream target of *P63*; and that *PERP* and desmosome function together (Ihrie and Attardi, 2005; Ihrie et al., 2006; Ihrie et al., 2005). Therefore, it is possible that the

interrelation of *TP63*, *PERP* and the Desmosomal complex constitutes a P63>PERP>Desmosome pathway which might be involved in BEEC pathogenesis (Mahfuz et al., 2013b).

Additional Genes and Pathways in BEEC

A GWEP study showed that *Synpo2* and genes of WNT pathway may also play a role in BEEC pathogenesis (Qi et al., 2011). *Synpo2* is an actin-associated protein involved in actin-based motility in renal podocytes, which have a high capacity for protein synthesis and post-translational modifications (Lowik et al., 2009; Mundel et al., 1997). Extensive studies into the mechanisms of *Synpo2* may provide valuable information about the pathogenesis of this disease. The WNT protein family consists of 19 highly conserved secreted signalling proteins. WNT proteins have many functions, including the regulation of cell-to-cell interactions during embryogenesis. Abnormal WNT signalling and/or mutations in these genes can cause a range of birth defects and other diseases, including cancer (Angers and Moon, 2009; Logan and Nusse, 2004). Recent developments in WNT signal transduction research showed that it is essential for embryonic morphogenesis, and WNT signalling is essential for many critical developmental processes (Gao et al., 2011; Heisenberg et al., 2000; Qian et al., 2007; Rauch et al., 1997). Therefore, in depth research of WNT genes in BEEC patient may lead to promising discovery in disease progression.

It is evident that desmosomes are crucial components for epithelial stratification. Therefore, mutation screening of desmosomal components, such as desmoglein and desmocolin (found in extracellular core of desmosome), and also the plakoglobin and plakophilins (found in outer dense plaque of desmosome) may reveal more information on BEEC pathogenesis. These desmosomal components facilitate contact and adherence to neighbouring cells. In addition, establishing a functional relation between

*p63>Perp>*desmosome and WNT pathway could reveal more information on the development of BEEC.

1.4 *TP63 (P63)* Related Anomalies

Tp63 is expressed specifically in embryonic ectoderm and in the basal regenerative layers of epithelial tissues in the adult. Complete abrogation of *P63* gene function from a model animal reveals a developmental syndrome with ectodermal dysplasia, limb malformation and orofacial clefting as the key characteristic phenotype (Celli et al., 1999; Ianakiev et al., 2000; McGrath et al., 2001; van Bokhoven et al., 2001; van Bokhoven et al., 1999).

Heterozygous mutations in human *TP63* exhibits different combinations of ectodermal dysplasia , limb malformation and orofacial clefting, in five different human syndromes, such as i) Ectodactyly, Ectodermal dysplasia and Cleft lip/palate (EEC) syndrome, ii) Limb mammary (LMS) syndrome, iii) Acro-Dermato-Ungual-Lacrimal-Tooth (ADULT) syndrome, iv) Ankyloblepharon, Ectodermal Dysplasia, Clefting (AEC) syndrome and v) Rapp-Hodgkin (RHS) syndrome. These are the allelic *P63* conditions. In addition, non-syndromic single malformation conditions such as Split Hand Foot Malformation (SHFM4) and Non-Syndromic Cleft-Lip (NSCL) are also caused by *TP63* mutations.

In *P63*-related disorders, ectodermal dysplasia is one of the three major phenotypes, where abnormal growth and development of tissues and structures occurs. Ectodermal derivatives such as skin, hair, teeth, nail, as well as exocrine glands such as sweat and sebaceous, are usually abnormally developed. Lacrimal, salivary and mammary glands are also affected in some *P63* disorders (Figure 9) (Rinne et al., 2007). Limb malformation is the second major *P63* syndromic feature, with malformed hands and

feet having a severe median cleft in the palm and/or in the sole. These clefts usually occur as ectrodactyly (lacking one or more digits) and/or syndactyly (fusion of toes or fingers). The third hallmark of *P63* disorders is orofacial clefting, often found as cleft lip (CL) and/or cleft palate (CP). CL and/or CP are usually observed as a part of complex syndrome, where other organs are also affected (Figure 9) (Rinne et al., 2007).
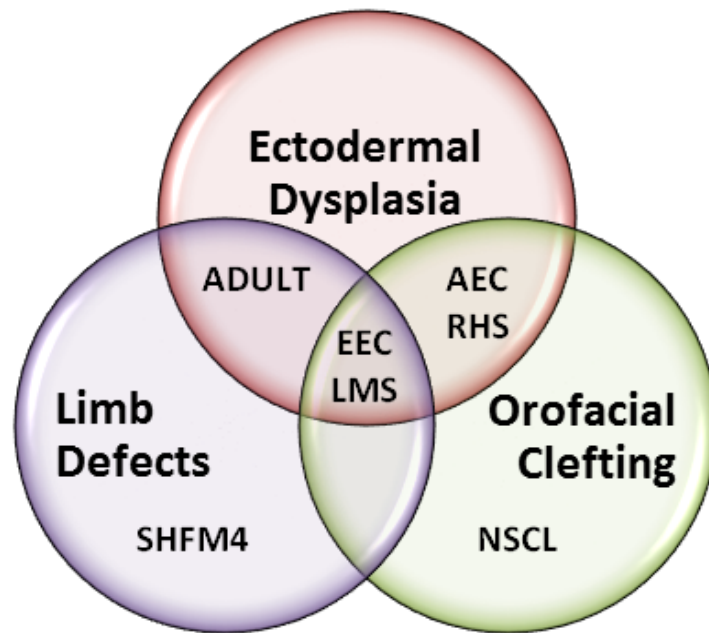
Figure 9: Various combinations of *TP63* phenotype showing allelic and non-syndromic P63 conditions. EEC: Ectodactyly, Ectodermal Dysplasia and Cleft lip/palate syndrome, LMS: Limb Mammary syndrome, ADULT: Acro-Dermato-Ungual-Lacrimal-Tooth syndrome, AEC: Ankyloblepharon, Ectodermal Dysplasia, Clefting syndrome and RHS: Rapp-Hodgkin syndrome, SHFM4: Split Hand Foot Malformation and NSCL: Non-Syndromic Cleft-Lip. The figure is adapted from Rinne et al. (2007).

1.4.1 Allelic P63 Conditions

Ectodactyly, Ectodermal Dysplasia and Cleft lip/palate (EEC) Syndrome

EEC syndrome is mainly characterized by one or more features of ectodermal dysplasia, including defects in hair, skin, nails, teeth and glands. EEC patients have light coloured, sparse hair. Sometimes absence of eyelashes, eyebrows and alopecia can be observed.

The epidermis is very dry and itchy with usually dystrophic nails. Dental alterations like hypodontia or anodontia have been reported in some cases (Arte, 2001; Bani et al., 2010; Chranowska et al., 1990). In EEC teeth are very susceptible to infection and caries, due to enamel defect and/or salivary gland malfunction. Lacrimal duct stenosis contributes to keratitis, which is common in EEC patients. Genitourinary malformations are also part of EEC. However, EEC severity can be highly variable, and is often related to the nature of the mutation (Rinne et al., 2006).

TP63 gene mutations account for most EEC cases, primarily a point mutation in the DNA binding domain of the *P63* gene. Only two mutations; and one insertion (1572 InsA) or a point mutation (L563P) have been found in the sterile alpha motif domain (SAM) (Celli et al., 1999; Rinne et al., 2006). However, five more frequently mutated amino acids in the EEC population are R204, R227, R279, R280 and R304; all located in CpG islands. These five mutations are called *P63* arginine hotspot mutations, which together explain almost 90% of the EEC syndrome patients (Rinne et al., 2006; van Bokhoven and Brunner, 2002). These five, and probably also other DNA binding domain mutations that are found in EEC syndrome, appear to impair *P63* protein binding to DNA (Celli et al., 1999).

Limb Mammary (LMS) Syndrome

LMS phenotype resembles EEC syndrome, but the ectodermal defects are much less prominent. Defects in hand and/or feet are found, but there are no hair or skin anomalies present in LMS patients. The more consistent features are nipple and/or mammary gland hypoplasia. Lacrimal duct obstructions are also frequently observed, and orofacial clefting is always limited to the palate in LMS patients (Rinne et al., 2007; Rinne et al., 2006).

LMS was the first TP63 syndrome linked to chromosome region 3q27 (van Bokhoven and Brunner, 2002; van Bokhoven et al., 2001; van Bokhoven et al., 1999). In LMS, *TP63* mutations are located in the N- and C- terminus of the gene. To date two missense mutations have been reported. One is found in exon 4, causing a G76W substitution in the ΔN-specific putative transactivation domain, and the other, S90W, is located between the TA and DBD domains (van Bokhoven et al., 1999). Two deletions were also found in C-terminus of the gene, specifically in exon 13 and 14 with a TT and an AA deletion respectively (van Bokhoven and Brunner, 2002; van Bokhoven et al., 2001). These deletions are predicted to cause a frame shift and a premature stop codon respectively. A stop mutation in the transcription factor inhibitory domain (TI) (K632X) has been identified in a sporadic LMS patient (Rinne et al., 2006).

Acro-Dermato-Ungual-Lacrimal-Tooth (ADULT) Syndrome

ADULT and LMS syndromes are clinically very similar, as in both cases mammary hypoplasia is one of the major key phenotypes. However, phenotypical differences were observed in a large patient cohort in ADULT syndrome, with absence of orofacial clefting and constant presence of ectodermal dyplasia (Rinne et al., 2007; Rinne et al., 2006).

ADULT syndrome is usually caused by a point mutation affecting amino acid R298 in exon 8 at the end of the DBD domain. This mutation is called the ADULT hot spot mutation, changing arginine into either a glutamine or a glycine. Two other mutations are located in the N-terminus: N6H mutation affects only the ΔN-isoforms, and in another isolated patient a missense mutation G134D is located just front of the DBD in exon 4 (Amiel et al., 2001; Slavotinek et al., 2005).

Ankyloblepharon, Ectodermal Dysplasia, Clefting (AEC) Syndrome

AEC syndrome is also known as Hay-Wells syndrome. I was first reported by R.J. Hay and R.S. Wells in 1976 (Hay and Wells, 1976). The principal feature of AEC syndrome at birth is the presence of skin erosion, which resembles a second degree burn (Tsutsui et al., 2003) and occurrence of ankyloblepharon, a partial or complete fusion of eyelids. These features are very rare in other EEC-like syndromes. Limb malformations are absent but CL/CP, (predominantly CP) is present. The other ectodermal dysplasia syndrome features (*i.e.* teeth, nails and hair defects) are constant in AEC. Lacrimal duct obstruction (50%), hearing impairment (40%) and mammary hypoplasia and hypohydrosis (13%) are also found (Rinne et al., 2007; Rinne et al., 2006).

Missense mutations in the SAM domain of P63 were initially detected in eight families by McGrath and colleagues (McGrath et al., 2001). These mutations are predicted to impair protein-protein interaction of the SAM domain. In a single family Barrow and colleagues found a mutation that causes a splicing defect, removing exon 11. This mutation affects the C-terminus of both *TP63 alpha* and *beta* isotypes (Barrow et al., 2002).

Rapp-Hodgkin (RHS) Syndrome

The features of RHS syndrome are very much similar to AEC syndrome, except absence of ankyloblepharon and a milder skin phenotype in RHS (Bertola et al., 2004; Rinne et al., 2006). Other ED symptoms, such as orofacial clefting and the near absence of limb malformations, are similar to AEC. The strong overlap between AEC and RHS suggests that they are variable manifestations of the same clinical entity (Bertola et al., 2004; Rinne et al., 2007).

TP63 mutations in RHS syndrome are located in the C-terminus. They are either point mutations in the SAM domain, or deletions in either the SAM or TI domains (Bougeard et al., 2003; Chan et al., 2005; Dianzani et al., 2003; Kannu et al., 2006; Kantaputra et al., 2003; Payne et al., 2005; Shotelersuk et al., 2005; Sorasio et al., 2006; Tsutsui et al., 2003). As the SAM domain of P63 is involved in protein –protein interactions, mutations in this region probably disrupt P63 function by hampering protein interactions.

1.4.2 Non-Syndromic P63 Conditions

Split Hand/Foot Malformation Type 4 (SHFM4)

Split Hand/Foot Malformation (SHFM) has been linked to several chromosomal loci. According to Ianakiev et al., 2000, about 10% of SHFM patients have a *TP63* mutation, which is mainly referred to as SHFM4 (Ianakiev et al., 2000). Several SHFM mutations have been reported to date, which are distributed along the *TP63* gene, such as a splice-site mutation in front of exon 4 (3'ss intron 4), missense mutations (R58C, K193E, K194E, R280C/H) in the DNA binding domain, and stop-mutations (Q634X, Q639X) in the TI domain (Ianakiev et al., 2000; van Bokhoven et al., 2001; Zenteno et al., 2005). However, it is still unclear how these widely dispersed mutations causes the limb defect, but it has been are reported that several of these mutations cause alteration in P63 protein activation and stability (Ghioni et al., 2005; Huang et al., 2004).

Non-Syndromic Cleft Lip/Palate (NSCL)

A single missense variant (R313G) was the first *TP63* variant associated with NSCL. This mutation is also observed in one sporadic EEC syndrome (Leoyklang et al., 2006). In addition to T*P63,* mutations and/or polymorphisms in *IRF6*, *MSX1* and *PVRL* have also been associated with syndromic and non-syndromic forms of orofacial clefting

(Ghassibe et al., 2005; Jumlongras et al., 2001; Lidral et al., 1998; Sozen et al., 2001; Zucchero et al., 2004).43-49. Leoyklang and colleagues also identified silent changes in *TP63*, such as N87N, S90L, L248L, H406H, and D564H, which did not cause any amino acid changes. It is quite possible that such variants in *TP63* can act as a risk factor in facial clefting, due to multifactorial origin of the condition (Rinne et al., 2007).

1.5. BEEC-Associated Anomalies

1.5.1 Urological Anomalies

About one third of all BEEC patients suffer from urological malformations in addition to bladder defect, which can be present in both males and females. The associated urological anomalies, however, are predominantly found in patients with cloacal exstrophy (Gearhart, 2001).

Ureteropelvic Junction (UPJ) Obstruction

UPJ obstructions are the most common form of congenital blockage in the urinary tract. The obstruction is at the point where the ureter joins the renal pelvis, which develops prenatally during the formation of the kidney. This defect is mostly considered as a functional obstruction, originating from anomalies in the smooth muscle of the pelvis and ureter (Mendelsohn, 2004). In UPJ obstruction, hydronephrosis (accumulation of urine within the kidney) occurs, where the kidney produces urine at a rate that exceeds the amount of urine able to drain out of the renal pelvis into the ureter. Although encountered less frequently in adults, UPJ obstruction may occur as a result of kidney stones.

Ectopic Pelvic Kidney

In 1936, Thomas and Barton defined ectopic kidney as a congenitally displaced kidney which never occupied a normal position. They also mentioned ectopic pelvic kidney as one that is fixed within the bony pelvis or across the spine. In such conditions the kidney derives its blood supply from the adjoining large vessels, such as iliac arteries (Thomas and Barton, 1936). Possible complications of an ectopic kidney include problems with urine drainage from that kidney. Abnormal urine flow and the placement of the ectopic kidney can lead to various problems such as, i) urinary tract infection due to urine accumulation in the kidney and ureter, ii) formation of urinary stones from substances found in the urine, such as calcium and oxalate, iii) damage to the kidney may occur due to vesicoureteral reflux, an abnormal flow of urine from the bladder to the upper urinary tract (ureter or kidney) and lastly iv) an ectopic kidney may be susceptible to blunt trauma, as it resides in the lower abdomen or pelvis.

Horseshoe Kidney

Horseshoe kidneys are the most common type of renal fusion anomaly. In normal individuals both kidneys are located in the back of the abdominal cavity on either side of the body. During development of the foetus, the kidneys move to their normal position and are protected by the ribs. But, in an individual with horseshoe kidney, as they transfer from the pelvic area to under the ribs, they fuse together at the lower end or base. Such fusion causes the kidneys to form a 'U' shaped organ which gives it the name "horseshoe kidney". The most common symptoms in children and adults with this defect include abdominal pain, nausea, kidney stones and urinary tract infections (Urologyhealth.org, 2014b).

Renal Agenesis

Renal agenesis is a congenital birth defect where babies are born without one or both kidneys. If the condition is bilateral (absence of both kidneys, traditionally known as the classic Potter syndrome), it is fatal to the newborn, whereas if unilateral (absence of one kidney), patients can have a normal life expectancy (Radiopaedia.org, 2014b).

The aetiology in many cases of renal agenesis is unknown (Mishra, 2007), but is thought to be multi-factorial. Embryologically, renal agenesis results from a failure of the proper development of the metanephros (the future definitive adult kidney) resulting in complete absence of a renal structure (Radiopaedia.org, 2014b).

Megaureter

The ureters are tube-like structures in the body with an average diameter of 3-5 mm in a new born. In case of megaureter the ureter is greater than 10 mm in diameter. Certain conditions can cause this abnormal widening, such as an abnormality of the ureter itself (primary), or from conditions related to the bladder obstruction and/development (secondary). To date, however, there are no genetic links to this disease (Urologyhealth.org, 2014c).

Ectopic Ureter

Ectopic ureter is a congenital renal anomaly that occurs as a result of abnormal caudal migration of the ureteral bud during its insertion to the urinary bladder (Berrocal et al., 2002). It is thought to occur in 1 in every 500 persons and is usually asymptomatic in males. It is more common in females, where the F:M ratio is 10:1 (Radiopaedia.org, 2014a).

Normally the ureter drains via the internal ureteral orifice at the trigone of the urinary bladder, but an ectopic ureter fails to connect properly to the bladder and drains

somewhere outside the bladder. In females, the ectopic ureter usually drains into the urethra or even the vagina, while in boys, it usually drains into the urethra near the prostate or into the genital duct system (Urologyhealth.org, 2014a).

Ureterocele

Ureterocele is a birth defect that affects the kidney, ureter and bladder. In this condition the portion of the ureter closest to the bladder swells up, because the ureteral opening in to the bladder is very small and the swollen area prevents urine from moving freely into the bladder. As a result of this, urine collects in the ureter and stretches its walls with a possibility of rupture (MedlinePlus, 2014).

1.5.2 Spinal and Orthopedic Anomalies

The incidence of spinal anomalies widely varies within the EEC spectrum. Among spinal anomalies, spina bifida, myelomeningocele and sacral hypoplasia were reported in BEEC patients (Cadeddu et al., 1997). Various spinal anomalies occur in about 7% of cases of classic bladder exstrophy (Ebert et al., 2009).

Spina bifida is one of the most common malformations, which results from failure of fusion of the caudal neural tube. The causes of this disorder are heterogeneous and include chromosome abnormalities, single gene disorders, and teratogenic exposures. However, the cause is not known in most cases and is sometimes associated with other abnormalities (Mitchell et al., 2004).

Myelomeningocele is a type of spina bifida in which the bones of the spine do not completely form, resulting in an incomplete spinal canal. This causes the spinal cord and meninges (the tissues covering the spinal cord) to protrude from the child's back. Such a defect was also reported in BEEC patients (A.D.A.M. Medical Encyclopedia, 2014).

Skeletal and limb anomalies *i.e.* clubfoot deformities, absence of feet, tibial or fibular deformities, and hip dislocations are commonly seen in cloacal exstrophy (Gearhart, 2001).

1.5.3 Gastrointestinal Anomalies

Gasstrointestinal anomalies are predominantly seen in cloacal exstrophy in the form of omphalocele or exomphalos (10-88% cases) of varying size (Ebert et al., 2009). Such an anomaly seldom presents in classic bladder exstrophy. Gearhart and co-workers reported that gastrointestinal malrotation and short bowel syndrome is present in about 46% and 25% cloacal exstrophy patient, respectively (Gearhart, 2001). On rare occasions, duodenal atresia (congenital absence or complete closure of a portion of the lumen of the duodenum) and small bowel deletion have been described in cloacal exstrophy (Gearhart, 2001; Schober et al., 2002).

1.5.4 Gynecological Anomalies

In addition to the split external female genitalia in BEEC, in most cases the cervix inserts low down at the superior vaginal wall (Ebert et al., 2009; Gearhart, 2001; Woodhouse and Hinsch, 1997). Müllerian anomalies are quite common in cloacal exstrophy, as duplication of the vagina and uterus, as well as vaginal agenesis have been reported (Gearhart, 2001).

1.5.5 Epidermolysis Bullosa

Epidermolysis bullosa (EB) is a congenital connective tissue disorder causing blisters in the skin and mucosal membranes. It is a result of a defect in anchoring between the epidermis and dermis, resulting in friction and skin fragility. In a case report Moretti and colleagues showed that this disorder is also associated with BEEC (Moretti et al., 1995).

1.6 Summary

Although several theories exists that attempt to explain the embryological association of the development of BEEC, none describe the molecular aspects of its pathogenesis. With the development of cutting edge technologies in developmental biology and molecular genetics, it is now possible to explain this congenital defect in different angle. The creation of the only genetic model (*p63* knockout mice), allowed researchers to identify the candidate gene *TP63* for human BEEC. Genome wide expression studies further suggested a promising contribution of *PERP* and *DSP* gene in the pathogenesis of BEEC. In addition, the close relationship among *TP63*, *PERP* and *DSP* gene suggests a role for the TP63>PERP>Desmosome pathway in BEEC aetiology (Figure 10).
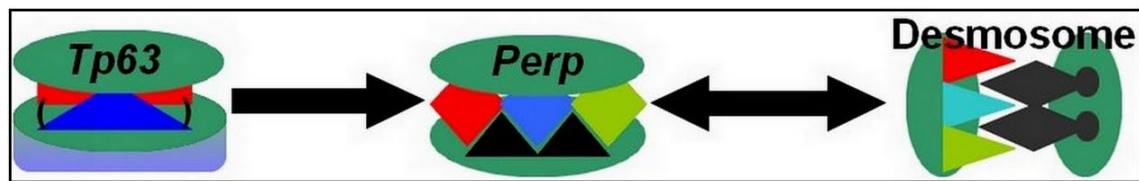
Figure 10: Proposed TP63>PERP>Desmosome pathway in BEEC pathogenesis.

1.7 Hypothesis

Precise regulation of the *TP63>Perp>*Desmosome pathway is required for the proper development and differentiation of the human bladder urothelium during embryogenesis. Perturbation of this pathway may increase the risk of BEEC.

1.8 Aims and objectives

Aim 1: To perform mutation screening and expression analysis of T*P63* in BEEC patients.

Aim 2: To perform mutation screening of *PERP* and *DSP* in BEEC patients.

Aim 3: To identify regulatory elements (enhancers) involved in bladder development.

Chapter 2

Methodology

Chapter 2: Methodology

2.1 Introduction

The methods of experiment are a systematic and scientific approach to research, in which the researcher manipulates one or more variables, and controls and measures any change in other variables (Blakstad, 2008). To achieve my research goal I have used a range of techniques for detecting genetic variation and studying gene regulation.

2.1.1 Mutation Detection

Since the discovery of the role of DNA in development, scientists have tried to link changes in DNA sequence with disease. This has influenced the detection, diagnosis and treatment of human health conditions. The study of genetic mutations reveals the normal functions of genes and proteins, the causes of many diseases, and the variability of responses among individuals. There are many different ways that DNA sequence can be changed, resulting in different types of mutation such as substitution, insertion, deletion and frame-shift.

Substitution

A substitution is a mutation that exchanges one base for another (i.e., a change in a single "chemical letter" such as switching an A to a G). Such a substitution could change a codon to one that encodes a different amino acid (a missense change) and cause a change in the protein produced (Figure 11). For example, sickle cell anaemia is caused by a substitution in the beta-haemoglobin gene, which alters a single amino acid in the protein produced.

It is also possible that change results in a codon that encodes the same amino acid (silent mutation) and causes no change in the protein produced. The most severe type is

one that changes an amino-acid-coding codon to a single "stop" codon (a nonsense mutation) and results in an incomplete protein. This can have serious effects on human health since the incomplete protein probably will be absent or function poorly.

Substitution
CATTAGGAGACT ←—Reference—→ CATTAGGAGACT **Deletion**
CATTGGGAGACT ←— Mutated —→ CATTAAGACT

A   **C**

Insertion
CATTAGGAGACT ←—Reference—→ CAT TAG GAG ACT **Frameshift**
CATTAGTAGGAGACT ←— Mutated —→ ATT AGG AGA CT

B   **D**
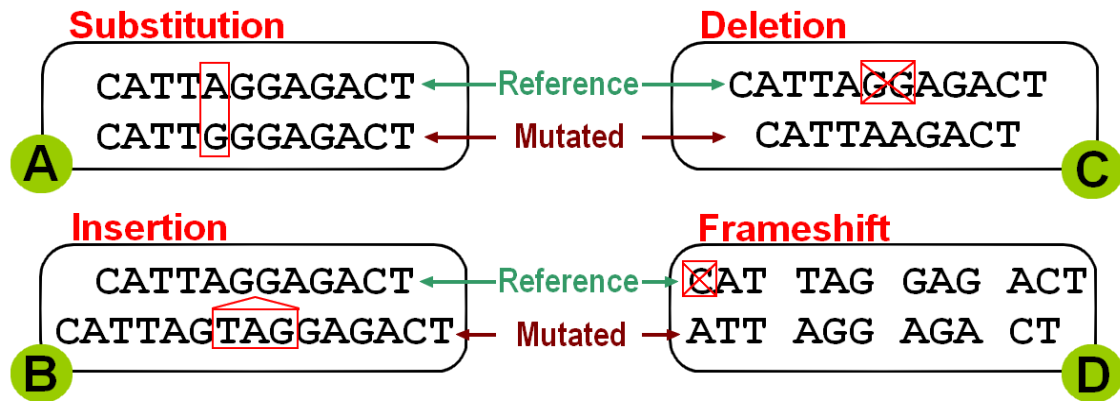
Figure 11: Different types of mutations. Upper line of the each panel (A, B, C and D) is reference sequence and lower line is mutated sequence. A. Substitution mutation where nucleotide 'A' is substitute with 'G'. B. Insertion mutation where nucleotide 'TAG' is inserted in to DNA sequence. C. Deletion mutation where nucleotide 'GG' is deleted from the reference sequence. D. Frame-shift mutation causing incorrect amino acid sequence which codes for different protein or no protein at all.

Insertion and Deletions

Insertions are mutations in which extra base pairs are inserted into a new place in the DNA. Figure 11 shows an insertion of 'TAG' into the DNA strand. On the other hand, deletions are mutations in which a section of DNA is lost, or deleted. Figure 11 shows a deletion of 'GG' from the DNA.

Frame-shift

Since protein-coding DNA is divided into codons of three bases, insertions and deletions can alter the reading frame of a gene so that its message is no longer correctly parsed. These changes are called frameshifts. Frameshift mutations at the DNA level,

causing the codons to be parsed incorrectly, usually lead to a premature stop codon and a truncated protein (Figure 11).

2.1.2 Gene Regulation

Gene regulation is the process of turning genes on or off. This process ensures that the appropriate genes are expressed at the proper times, from early developmental stages in the foetus to adulthood. Gene regulation is significantly more complex in eukaryotes than in prokaryotes for a number of reasons. First, the genome being regulated is larger. Another source of complexity in eukaryotic gene regulation is the many different cell types present in most eukaryotes. Liver and pancreatic cells, for example, differ dramatically in the genes that are highly expressed (Berg et al., 2002).

During early development, cells begin to take on specific functions regulated by different factors. Gene regulation can also help an organism respond to its environment. Regulation of gene expression can be accomplished by a variety of mechanisms, including chemically modifying genes and using regulatory proteins to turn genes on or off. The expression of mammalian genes is regulated primarily at the level of initiation of transcription. The regulatory structure of mammalian genes consists of the sequence coding for a protein, a proximal upstream promoter sequence which binds the general (basal) transcription factors and one or more distant enhancer sequences which binds the inducible transcription factors (Figure 12) (Beyersmann, 2000).
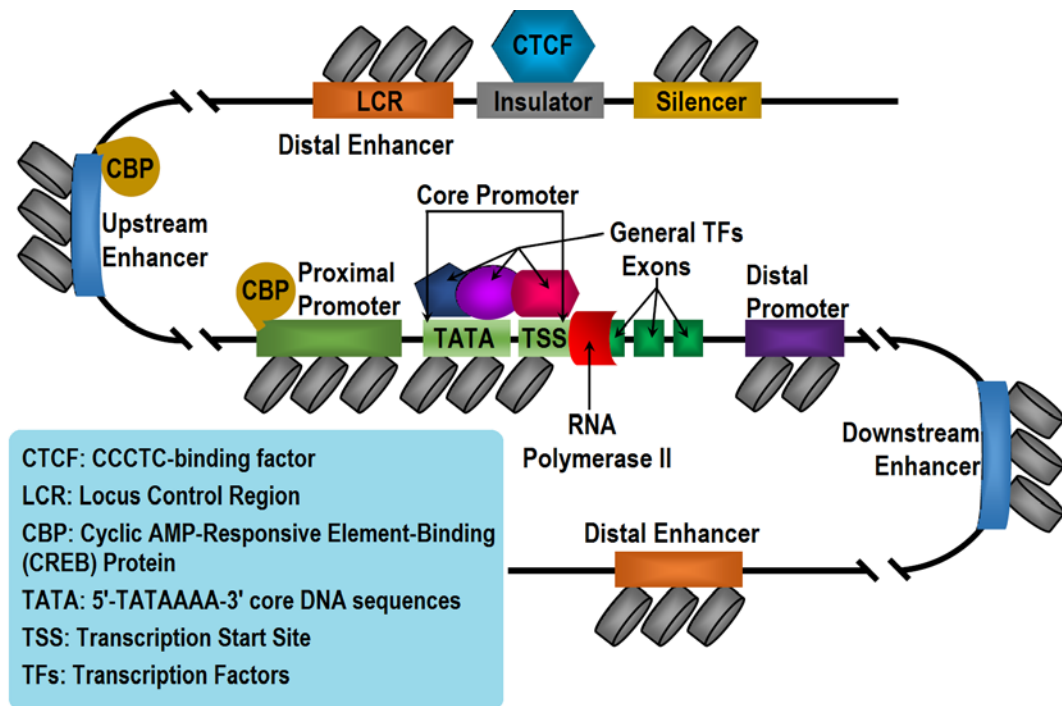
Figure 12: Schematic diagram of mammalian gene regulation showing position of promoter, enhancer and transcription. Adapted fromOng and Corces (2011), Nat. Rev. Genet.

Epigenetics

Epigenetics is one of the most rapidly expanding fields in biology. Epigenetics involves any processes that can alter the way a cell interacts with DNA, without changing the sequence of molecular building blocks of DNA, and leads to modifications that are heritable (Weinhold, 2006; Williams, 2013). Epigenetic processes are natural and essential to many organism functions, but if they occur improperly, can have major adverse health and behavioral effects. Epigenetic regulation or changes consists of chemical flags, or markers, on genes which includes methylation, acetylation, phosphorylation, ubiquitylation, and sumolyation. They are involved in regulating key cellular process such as gene transcription, DNA replication, and DNA repair. These changes can influence a cell's access to a gene, turning it on or off for specific functions (Portela and Esteller, 2010; Weinhold, 2006; Williams, 2013). The best known and

most studied epigenetic regulations are DNA methylation and chromatin (histone) modifications.

DNA Methylation

DNA methylation (Robertson, 2005) is the addition or removal of a methyl group (CH_3), to the C-5 position of the cytosine ring of DNA by DNA methyltransferases (DNMTs). It is regarded as a key player in epigenetic silencing of transcription. DNA methylation may co-ordinately regulate the chromatin status via the interaction of DNMTs with other modifications such as acetylation, phosphorylation, ubiquitylation, and sumolyation, and with components of the machinery mediating those marks (Jin et al., 2011; Robertson, 2005).

Chromatin Modification

Chromatin is the complex of proteins (histones) associated with DNA. Five major histones are present in chromatin. Four (H2A, H2B, H3, and H4), associate with one another as the nucleosome; the other histone is called H1. Histones have strong basic properties because a quarter of the amino acid residues in each histone are either arginine or lysine (Berg et al., 2002). The DNA is tightly bound around the nucleosome to fit into the nucleus. There are several substances which can modify the chromatin, such as acetyl groups (the process called acetylation), enzymes, and some forms of RNA such as microRNAs and small interfering RNAs. Chromatin structure is altered by such modifications to influence gene expression. In general, tightly folded chromatin tends to be shut down, or not expressed, while more open chromatin is functional, or expressed (Portela and Esteller, 2010; Weinhold, 2006). An example of chromatin remodelling is shown in the Figure 13.
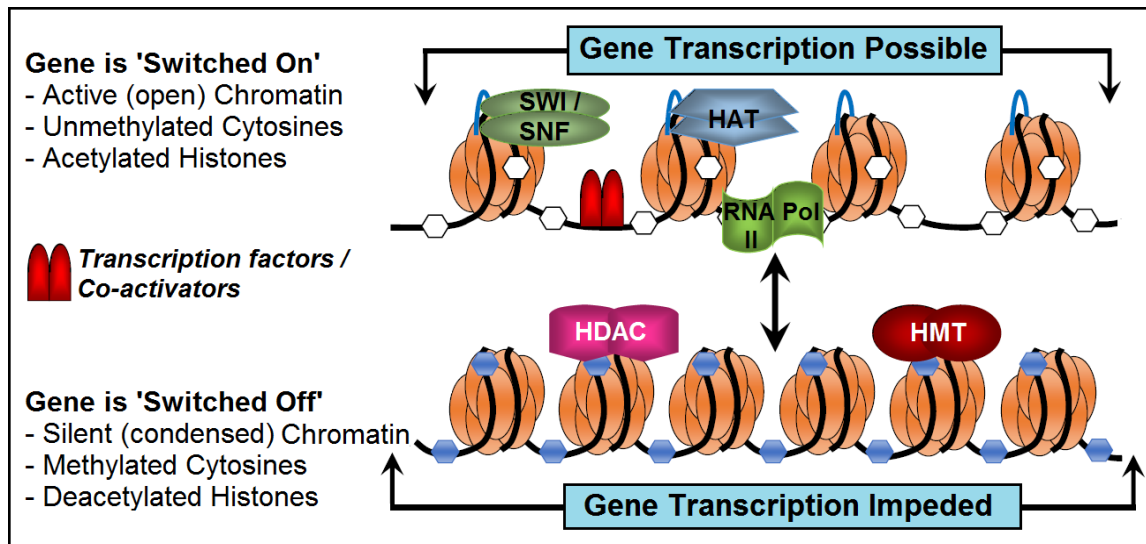
Figure 13: Chromatin remodelling complexes in the regulation of gene transcription. Chromatin is loosely packaged in the presence of acetylated histones (HAT mediated) and absence of methylase (HMT) activity (unmethylated cytosines are shown in white hexagon). A very well-known chromatin remodeler complex, SWI/SNF, opens up DNA region where transcription initiation proteins *i.e.* RNA Polymerase II, transcription factors and co-activators bind to turn on gene transcription. In the absence of SWI/SNF, chromatin remains tightly crammed to one another. Additional methylation by HMT (methylated cytosines are shown in blue hexagon) and deacetylation by HDAC proteins condenses DNA around histones and thus, make DNA inaccessible for binding by RNA Pol II and other activators, leading to gene silencing.

Promoters

A promoter can be defined as a combination of short sequence elements located upstream of a gene to which RNA polymerase binds, in order to initiate transcription of a gene. Expression of genes is initiated by the binding of transcription factors to the promoter (Juven-Gershon et al., 2008; Smale and Kadonaga, 2003; Strachan and Read, 1999). There are two major types of promoters found, focused and dispersed. In focused core promoters, transcription initiates at a single site or in a cluster of sites in a narrow region of several nucleotides. Dispersed core promoters are typically found in CpG islands in vertebrates and usually yield multiple weak start sites over a region of 50 to 100 nucleotides (Juven-Gershon et al., 2008; Smale and Kadonaga, 2003). In addition,

core promoters may contain many different sequence motifs, such as the TATA box, BRE, Inr, MTE, DPE, DCE, and XCPE1, which specify different mechanisms of transcription and responses to enhancers. Thus, the core promoter is a sophisticated gateway to transcription that determines which signals will lead to transcription initiation (Juven-Gershon et al., 2008; Kadonaga, 2012; Muller et al., 2007; Nikolov and Burley, 1997; Sandelin et al., 2007; Smale and Kadonaga, 2003).

Enhancers

A large part of gene transcription is regulated by enhancers, which are gene-distal DNA sequences containing multiple binding sites for a variety of transcription factors. They are often associated with specific histone modifications (Griffiths et al., 2003). These histones marks may be modified later in development to alter patterns of gene expression and cell differentiation choices. These marks may contribute to the repertoire of epigenetic mechanisms responsible for cellular memory, and determine the timing of transcription factor accessibility to the enhancer (Ong and Corces, 2011). Enhancers can activate transcription independent of their location, distance or orientation with respect to the promoters of genes (Banerji et al., 1981). In some instances, they can even activate transcription of genes located in a different chromosome (Geyer et al., 1990; Lomvardas et al., 2006).

Transcription Factors

Transcription factors (TFs) are proteins that bind to specific DNA sequences, thereby controlling the rate of transcription of genetic information from DNA to messenger RNA (Karin, 1990; Latchman, 1997). They are also known as sequence-specific DNA-binding factors. An essential feature of TFs is that they consist of one or more DNA-binding domains (DBDs) (Figure 14), which attach to specific sequences of DNA adjacent to the genes that they regulate (Mitchell and Tjian, 1989; Ptashne and Gann,

1997). Additional proteins such as co-activators, chromatin remodelers, and histone modifiers (acetylases, deacetylases, kinases, and methylases) lack DNA-binding domains. For this reason they are not classified as transcription factors, but do play vital roles in gene regulation (Brivanlou and Darnell, 2002).

Transcription factors may perform alone or with other proteins in a complex. In eukaryotes, RNA polymerase II transcribes messenger RNAs and several small nuclear RNAs. Like RNA polymerases I and III, RNA polymerase II cannot act alone. It requires five general transcription factors (TFIIB, TFIID, TFIIE, TFIIF, and TFIIH) to assemble on promoter DNA with RNA polymerase II for accurate transcriptional initiation (Lee and Young, 2000; Nikolov and Burley, 1997; Roeder, 1996). Another group of accessory factors, transcriptional activators and co-activators, regulate the rate of RNA synthesis from each gene in response to various developmental and environmental signals (Nikolov and Burley, 1997).

Figure 14: Schematic diagram of a transcription factor showing a DNA-binding domain (DBD), Signal-sensing domain (SSD), and a Transactivation domain (TAD). The order and the number of domains may differ in various types of transcription factors. In addition, the transactivation and signal-sensing functions are frequently contained within the same domain.

2.2 Methods Used in Detecting Genetic Variants

There are different methods that researchers use for detecting changes in DNA. Some look for large deletions and/or duplications, and others look for point mutations. WE have used several techniques to detect the sequence variation in selected genes.

2.2.1 High Resolution Melt (HRM) Analysis

High resolution melting is a versatile tool for detecting changes in a DNA sequence. This technique is unique in its simplicity, low cost, sensitivity and specificity (Aguirre-Lamban et al., 2010; Erali and Wittwer, 2010). With the advancement in fluorescent dyes, instruments and software for DNA melting analysis, the HRM technique opens up a new horizon for identifying sequence variants (Erali and Wittwer, 2010; Wittwer, 2009).

Until recently, mutation detection has relied heavily on the use of gel-based methods. Such methods can be both time consuming and difficult to design. Therefore, especially in the research/diagnostic/clinical setting, nongel-based systems are very important. An example is high-resolution melt (HRM) curve analysis (Tindall et al., 2009). HRM is a closed-tube method that incorporates a saturating dye during DNA amplification followed by a monitoring of the change in fluorescence as the DNA duplex is denatured by an increasing temperature (Polakova et al., 2008; Tindall et al., 2009).

Incorporation of saturated dye in HRM analysis facilitates the detection of heteroduplexes without any additions, processing or separation steps (Wittwer et al., 2003). Homozygous or hemizygous variants are detected by prior mixing with wild-type DNA. Heterozygous variants are identified by a change in the melt curve shape. This is achieved by plotting the fluorescence difference between melt curves compared to the reference (usually the wild-type) profile (Figure 15) (Erali and Wittwer, 2010).

HRM Difference Curves
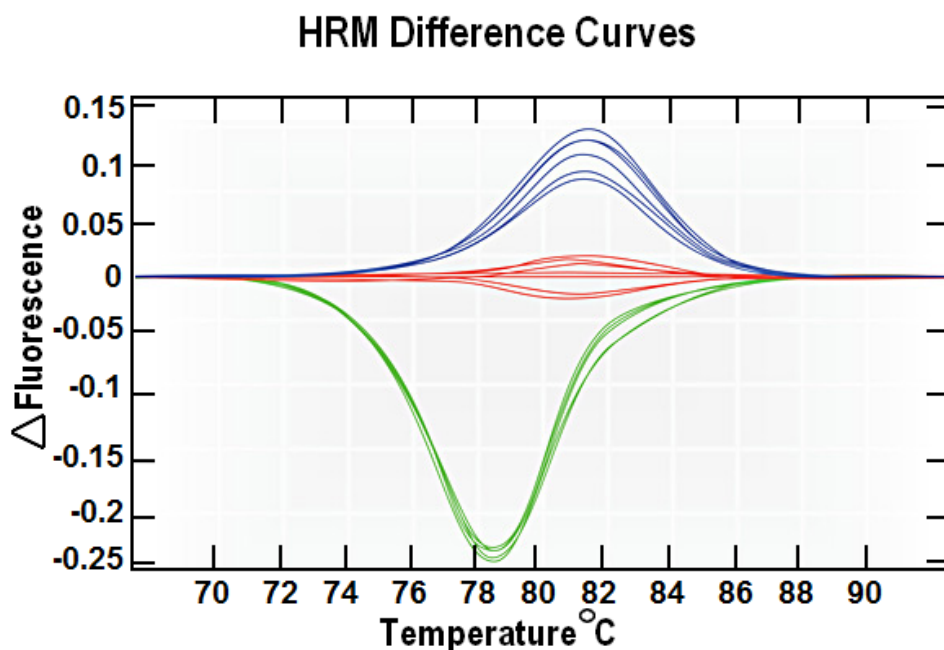
Figure 15: HRM difference curves showing melting temperature (x axis) plotted against fluorescence (y axis). In the above figure, HRM curves in 'Red' is considered as normal curve which is identical to reference sequence while the curves in 'Blue' and 'Green' are the deviation from the normal HRM curve considered as aberrant HRM curves which contains one or more sequence variant compare to reference.

2.2.2 Sanger Sequencing

Sanger sequencing/Chain-termination method was the most widely used sequencing method for approximately 25 years. This method was developed by Frederick Sanger and colleagues in 1977. This DNA sequencing approach is based on the selective incorporation of chain-terminating di-deoxynucleotides by DNA polymerase during *in vitro* DNA replication (Sanger and Coulson, 1975; Sanger et al., 1977). The method requires a single-stranded DNA template, a DNA primer, a DNA polymerase, normal deoxynucleoside triphosphates (dNTPs), and modified di-deoxynucleotide triphosphates (ddNTPs) (Figure 16). Due to the lack of a 3'-OH group in ddNTPs, the phosphodiester bond between two nucleotides cannot form. As a result of this, DNA polymerase is unable to extend the DNA sequence when a modified ddNTP is incorporated.

37
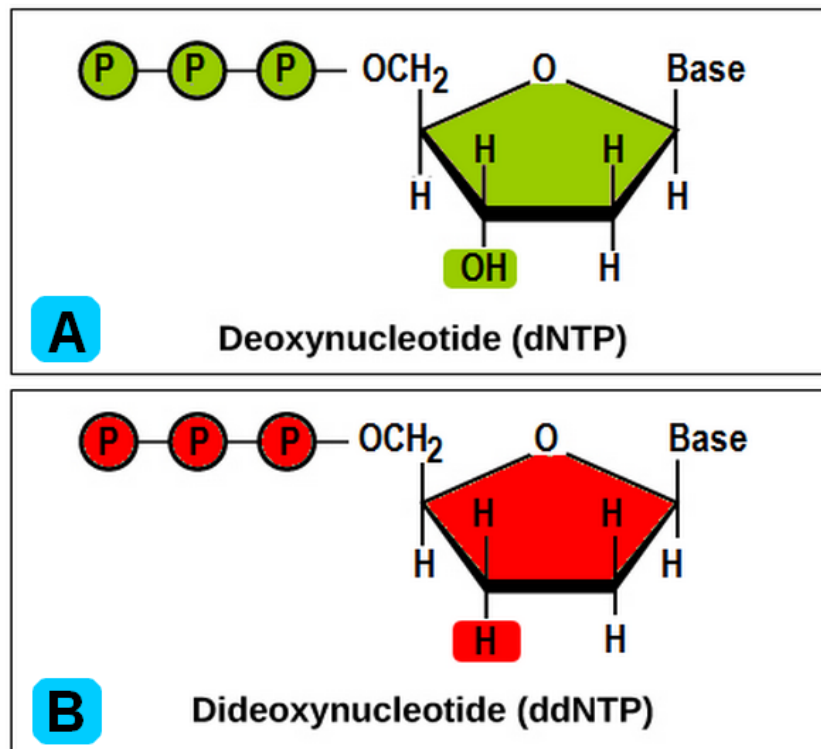
Figure 16: Chemical structure of normal and modified dNTPs used in the chain termination method. A. Normal dNTP (Deoxynucleotide) containing 3'OH group. B. Modified dNTP (Dideoxynucleotide) lacking 3'OH group which impede formation of phosphodiester bond.

The classical chain termination method requires four separate sequencing reactions, containing all four of the standard deoxynucleotides (dATP, dGTP, dCTP and dTTP) and the DNA polymerase (Figure 17). To each reaction is added only one of the four dideoxynucleotides (ddATP, ddGTP, ddCTP, or ddTTP). Four separate reactions are needed in this process to test all four ddNTPs. Following rounds of template DNA extension from the bound primer, the resulting DNA fragments are heat denatured and separated by size using gel electrophoresis.
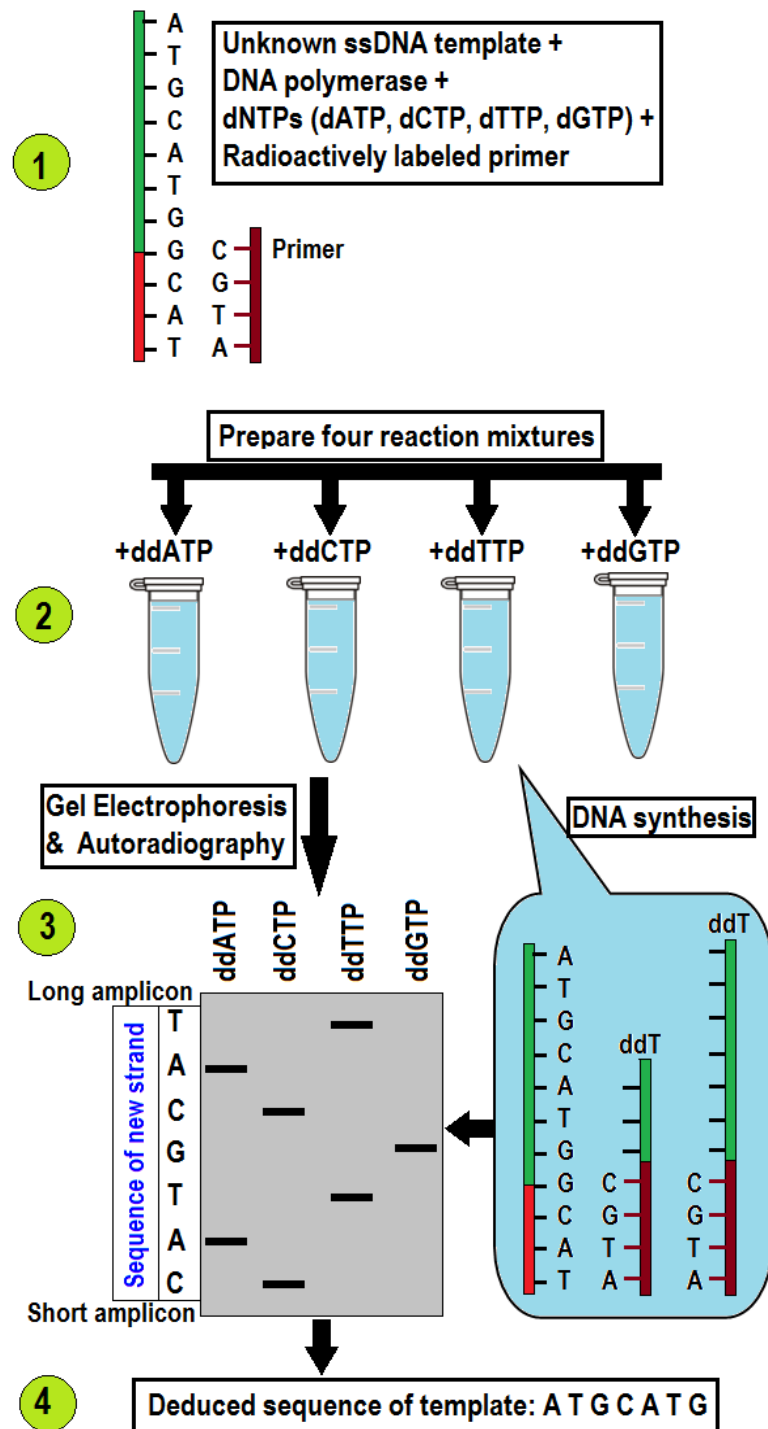
Figure 17: Chain termination method of DNA sequencing. Step 1. A master mix (MM) is prepared using radioactively labelled primer, normal dNTPs, DNA polymerase and unknown DNA template. Step 2. Four separate reaction mixture is prepared by mixing 4 different ddNTPs to the MM. Step 3. DNA synthesis occurred following rounds of template DNA extension. Amplified DNA was then separated by gel electrophoresis and autoradiography. Step 4. Based on the amplicon size sequence of the unknown DNA template is deduced.

A technical variation of Sanger sequencing is the dye terminator sequencing, where each of the four dideoxynucleotide chain terminators is labelled with a different fluorescent chromatophore (Smith et al., 1985; Smith et al., 1986). These fluorophore are covalently attached to the oligonucleotide primer used in enzymatic DNA sequence analysis. For each specific bases (A, C, G and T), a different coloured fluorophore is used, each of which emit light at different wavelengths (Smith et al., 1986). The reaction mixtures are combined and co-electrophoresed down a single polyacrylamide gel tube. The separated fluorescent bands of DNA are detected near the bottom of the tube, and the sequence information is acquired directly by computer. This method of sequencing is partially automated, faster and economical, which permits sequencing in a single reaction, rather than four reactions as in the classic chain termination method.

Modern DNA sequencing is performed in a DNA sequencer. It can sequence up to 384 DNA samples in a single run. Modern day DNA sequencers follow the same basic principle as in the classic chain termination method, except it is fully automated. However, sequencing reactions by thermo-cycling, clean-up and re-suspension in a buffer solution are performed separately before loading onto the sequencer. The DNA fragments are separated by capillary electrophoresis, and the sequence trace is detected by a laser detection method (Figure 18). The trace is then captured and analysed in computer (Hood et al., 1992).
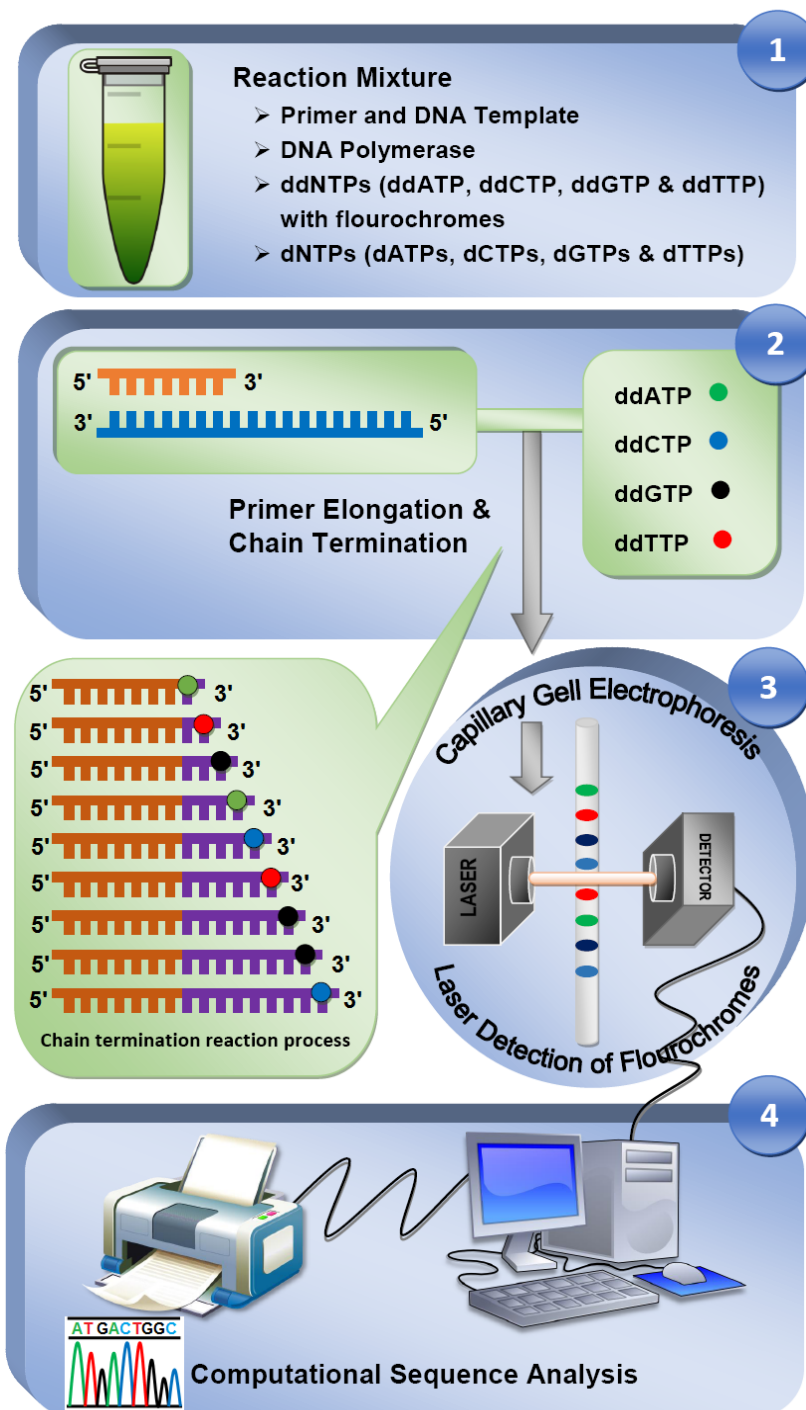
Figure 18: Automated dye terminator sequencing with laser detection and computational analysis. Step 1. MM is prepared using unknown DNA template, primers, DNA polymerase, dNTPs and ddNTPs. Step 2. DNA synthesis is performed by primer elongation and chain termination in a DNA sequencer. Step 3. Sequence of the template DNA is determined by capillary electrophoresis and laser detection of corresponding nucleotide. Step 4. Sequence trace is then captured and analysed in computer.

2.2.3 Next Generation Sequencing

Over the past 2-3 decades, DNA sequencing by the Sanger method (Sanger et al., 1977) has been regarded as the gold standard for the identification of mutations in human genetic disease. Sanger sequencing of a target region is an accurate and cost effective way to obtain a definitive molecular diagnosis for single gene disorders with known mutation hot spots (Zhang et al., 2014). However, it is costly and time consuming when it comes to identifying the candidate gene of a disorder or in case of whole genome sequencing (Zhang et al., 2014). Recently, the development of high throughput, massively parallel sequencing (MPS), or next generation sequencing (NGS), technology has revolutionized the molecular diagnosis and research of human genetic diseases. Platforms have been developed by several companies, including Roche Applied Science (454 Genome Sequencer FLX [GS FLX] System), Illumina (Genome Analyzer [GA] II), Life Technologies (Sequencing by Oligonucleotide Ligation and Detection [SOLiD™]) and Helicos BioSciences (HeliScope™ Single Molecule Sequencer) (Mardis, 2008; Shendure and Ji, 2008; Su et al., 2011). The ability to generate enormous amount of sequence data in a short time at an affordable cost makes NGS ideal for a wide range of applications, from sequencing a group of candidate genes and exome sequencing to the entire human genome (Morozova and Marra, 2008; Su et al., 2011; Zhang et al., 2014). Compared to Sanger sequencing, NGS technologies yield shorter read lengths; however, despite this drawback, they have greatly facilitated genome sequencing (Marguerat et al., 2008; Nowrousian, 2010). NGS technology can now be regarded as the molecular microscope in virtually every field of biomedical sciences (Buermans and den Dunnen, 2014). Also, next-generation sequencing is triggering a number of new assays and applications which has greatly advanced the knowledge of genome function (Figure 19) (Wold and Myers, 2008). The principle behind these

alternative applications is that the complex DNA or RNA samples are directly sequenced to determine their content. Such evolving applications are termed the "Sequence Census Method' (Marguerat et al., 2008; Wold and Myers, 2008).
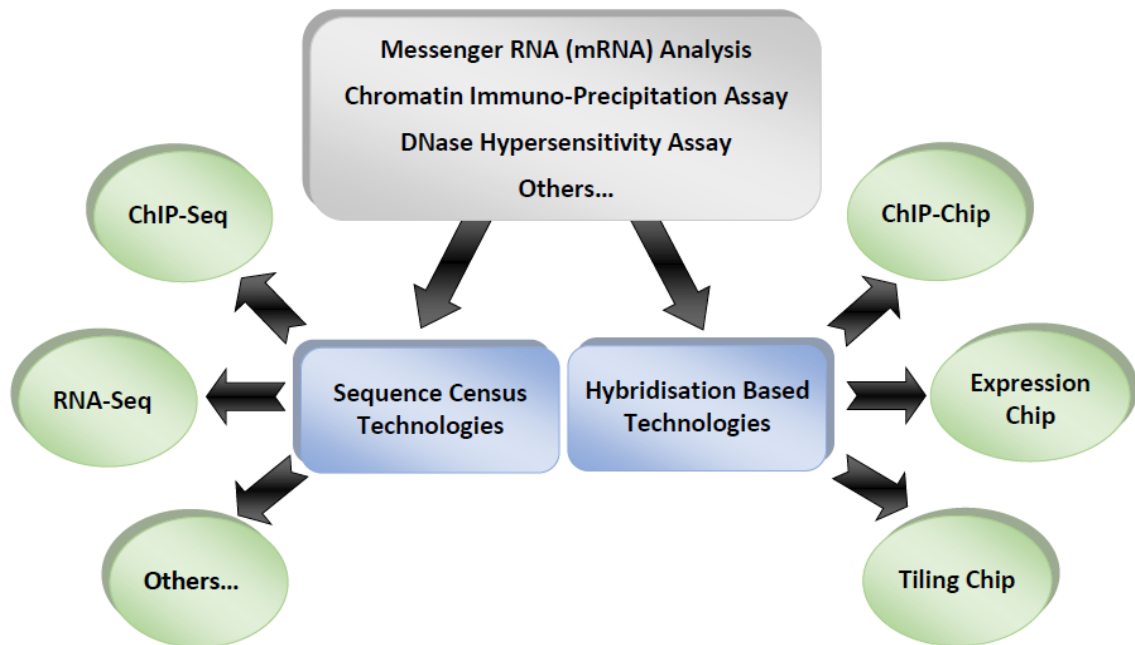
Figure 19: Sequence census method showing applications evolved from NGS technologies to facilitate the study of gene regulation and expression.

With availability of reference genomes; locations of generated short sequence reads can easily be mapped (except for repeated regions). After mapping millions of sequences, reads are aligned to determine their genomic distribution (Figure 20) (Buermans and den Dunnen, 2014; Metzker, 2010; Nowrousian, 2010).
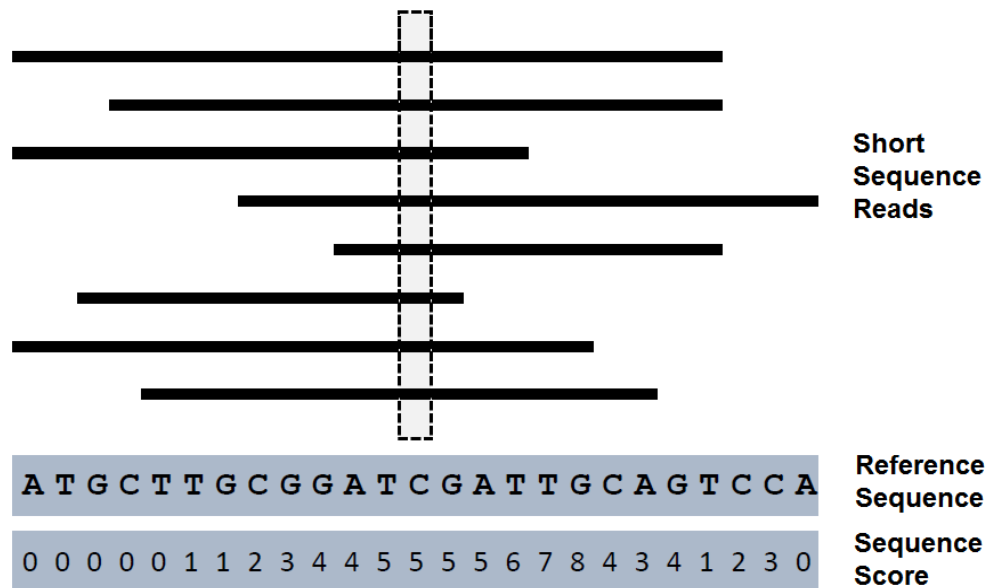
Figure 20: An example of sequence census data applied to cDNA, allowing genome-wide measurements of transcript levels (RNA-Seq). The sequence score defines the number of times each base of the reference genome sequence is covered by a sequence read.

2.3 Methods Used in Studying Gene Regulation

We have used several important methods in studying gene regulation, which are described below.

2.3.1 Cell Culture

Cell culture techniques have allowed scientists to analyse living cells for experimental studies and many types of biological assays. For a successful cell culture, aseptic techniques were used to maintain and preserve cultures in all basic steps of cell culture. In our study, we have used Human Urothelial Cell Line (Sv-Huc1) from American Type Culture Collection (ATCC). The cells were grown in Dulbecco's Modified Eagle Medium (DMEM). Sub-cultures were maintained according to standard cell culture protocol (Mather and Roberts, 1998)

2.3.2 Chromatin Immuno-Precipitation (ChIP)

Many crucial cellular functions (e.g. transcription, replication and recombination, repair, segregation, chromosomal stability, cell cycle progression, and epigenetic silencing) are essentially dependent on the association between proteins and DNA. ChIP is a powerful and versatile technique in the study of DNA-protein interactions within the milieu of the cell (Das et al., 2004). This technique has been extensively used to map the localization of histones, histone variants, transcription factors, or chromatin modifying enzymes on the genome or on a given locus. The basic steps in this technique are fixation, fragmentation, immune-precipitation, and analysis of the immune-precipitated DNA (Collas, 2010; Das et al., 2004).

To identify DNA-protein interactions, cells are cross-linked by adding 1% formaldehyde, which is stopped by the addition of glycine (125 mM). Fragmentation (usually by sonication) of chromatin to an average length of 200-700 bp is carried out in a Covaris using appropriate cell lysis buffer. Chromatin Immuno-Precipitation (IP) of the target protein is performed using magnetic beads coated with corresponding antibody. The IP sample is then reverse cross-linked to unbind the DNA-proteins crosslink from the magnetic beads. Finally, the DNA is purified by magnetic bead purification to get ultrapure DNA, and quantified for further analysis (Figure 21).
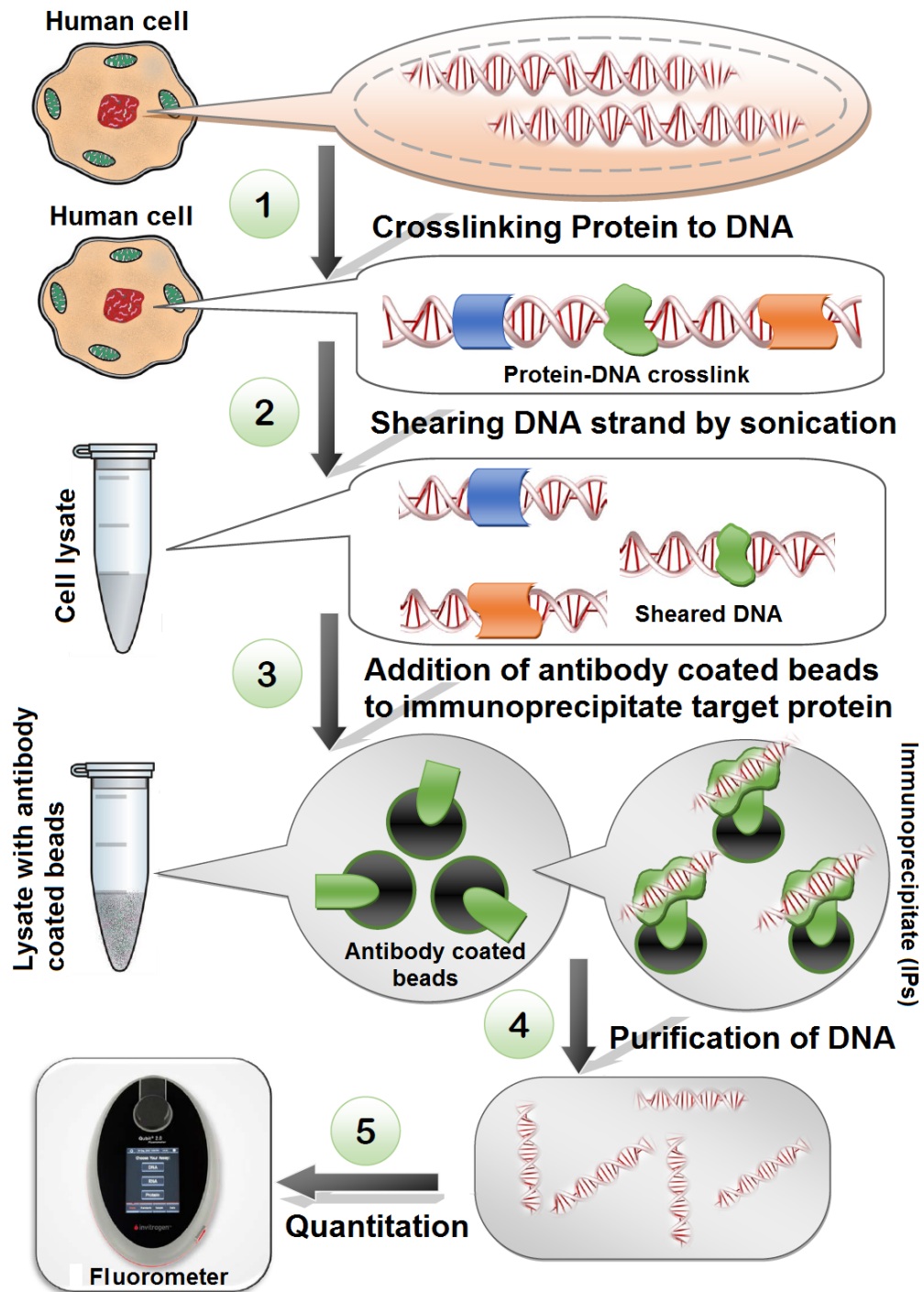
Figure 21: Basic steps of chromatin immuno-precipitation method (from cross linking to quantitation). Step 1. Protein-DNA cross linking is performed using formaldehyde in a cell suspension. Step 2. Cross linked cells are lysed using lysis buffer, and chromatin is sheared by sonication. Step 3. Antibody coated magnetic beads are added to the sheared chromatin. Step 4. Protein-DNA reverse cross-linking is performed and DNA is purified using magnetic purification. Step 5. Concentration of ultra-pure DNA is measured using fluorometer.

2.3.3 Real Time PCR (QPCR)

QPCR is a technique that allows quantification of a DNA sequence using relatively small quantity of tissues (Wong and Medrano, 2005). It is a refinement of the original polymerase chain reaction (PCR) developed by Kary Mullis and co-workers (Saiki et al., 1985). In contrast to standard PCR, in real-time PCR the amount of product formed is monitored during the course of the reaction. For this, it requires a fluorescent reporter that binds to the product formed and reports its presence. The reporter generates a fluorescence signal that reflects the amount of product formed and it is proportional to the number of amplification cycles required to obtain a particular amount of DNA (Higuchi et al., 1992; Kubista et al., 2006).

Table 3: List of fluorophores used in Real Time PCR

Fluorophores for Real Time PCR	
Intrinsically strong fluorescence, such as fluorescein and rhodamine derivatives-	Change fluorescence properties upon binding with nucleic acids-
• Hydrolysis probes, popularly called Taqman probes (Holland et al., 1991)	• LightUp probes (Svanvik et al., 2000)
• Molecular Beacons (Tyagi et al., 1998; Tyagi and Kramer, 1996)	• AllGlo probes (AlleLogic Biosciences, 2004)
• Hybridization probes (Caplin et al.)	• Displacement probes (Li et al., 2002)
• Lion probes	• Simple probes

Both non-specific labels and sequence-specific probes are available as reporters. Among non-specific labels, asymmetric cyanine dyes such as SYBR Green I and BEBO have become more popular. They bind to the minor groove of DNA and their fluorescence increases with the amount of double stranded product formed (Bengtsson et al., 2003; Zipper et al., 2004). Labelled primers and probes, also called fluorophores, are of two kinds: either with an intrinsically strong fluorescence, such as fluorescein and

rhodamine derivatives, or being able to change their fluorescence properties upon binding to nucleic acids (Table 3) (Kubista et al., 2006).
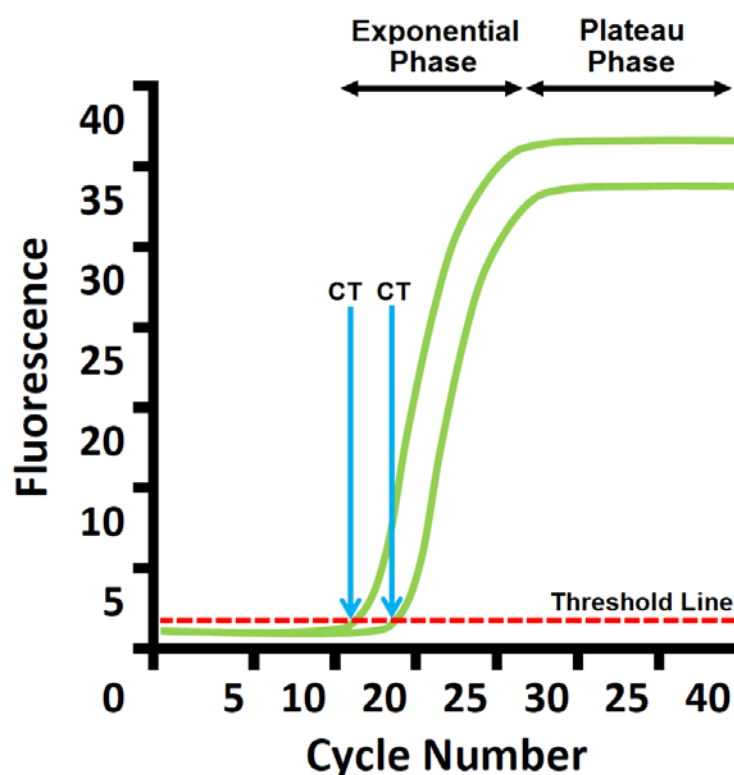
Figure 22: Real-time PCR response curves. A threshold level is set sufficiently above background and the number of cycles required reaching threshold, CT. The figure is modified from Kubista et al. (2006).

In a typical real-time PCR experiment all response curves saturate at the same level (Initial and Plateau phase). Therefore, it can be difficult to determine the initial amounts of target molecules in the samples by end-point PCR measurements; they only distinguish a positive from a negative sample. However, the response curves are separated and are expected to be parallel in the exponential phase of the reaction (Figure 22). This phase reflects the difference in their initial amounts of template molecules and allows quantification by comparing the number of amplification cycles required for each sample to reach a particular threshold fluorescence signal level, called the CT value (Higuchi et al., 1992; Kubista et al., 2006; Wong and Medrano, 2005). Different

instrument software use different methods and algorithms to select the threshold, and most also let the user set it manually.
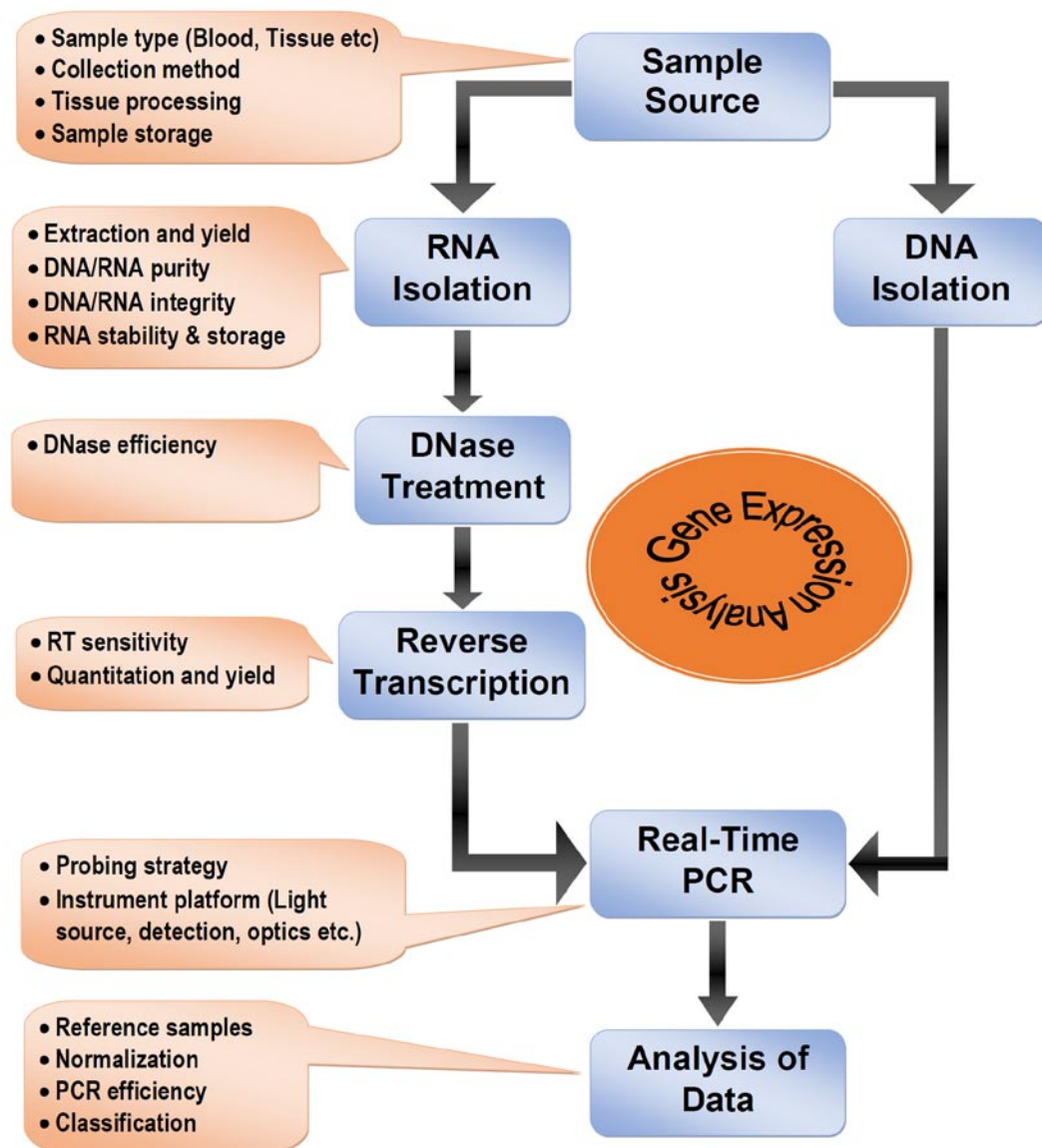
Figure 23: Flow diagram of RNA and DNA analysis from biological samples. Sources of variation are indicated in the balloons. Modified from Kubista et al. (2006).

For sensitive and accurate quantification of a gene by real-time PCR, it is necessary to convert the mRNA in the sample to cDNA by reverse transcription (RT). The RT step is critical since the amount of cDNA produced must accurately reflect the input amount of the mRNA (Kubista et al., 2006). Different steps of performing a gene expression measurement are show in the Figure 23.

49

2.3.4 ChIP-Sequencing

ChIP-Seq is a powerful and versatile tool. It combines Chromatin Immuno-Precipitation with next-generation sequencing (Barski et al., 2007; Johnson et al., 2007). ChIP-Seq requires a high-quality genome sequence for read mapping. Using ChIP-Seq, protein binding can be mapped in a genome-wide manner with extremely high resolution. The immuno-precipitated DNA from ChIP is used as the input for NGS (Figure 24). The methods involved in NGS are generally- template preparation, sequencing and imaging, and data analysis. Because the data from NGS are short sequence reads, ChIP-seq offers a rapid analysis pipeline. It also facilitates detection of mutations in binding-site sequences, which may directly support any observed changes in protein binding and gene regulation (Mardis, 2007; Mardis, 2008).
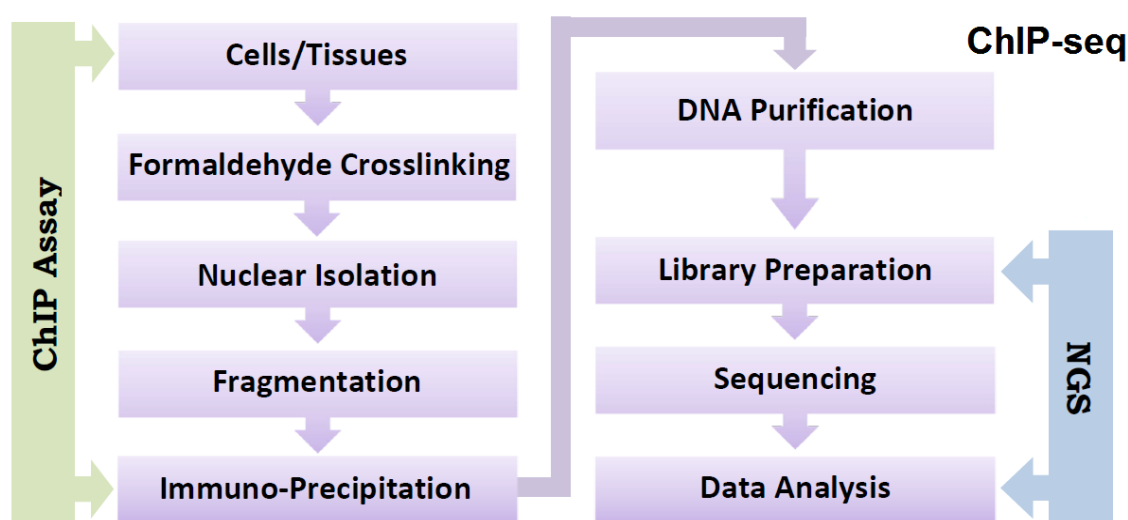
Figure 24: Overview of ChIP-seq workflow from harvesting to data analysis.

During data analysis peak calling and motif search is performed by software like HOMER (Heinz et al., 2010), Bowtie (Langmead and Salzberg, 2012), MACS (Liu, 2014) etc. From the analysed data it is possible to compare between different variables like patient and normal individual, treated and non-treated samples etc. Figure 25 shows an example of predicted outcome of ChIP-Seq assay.
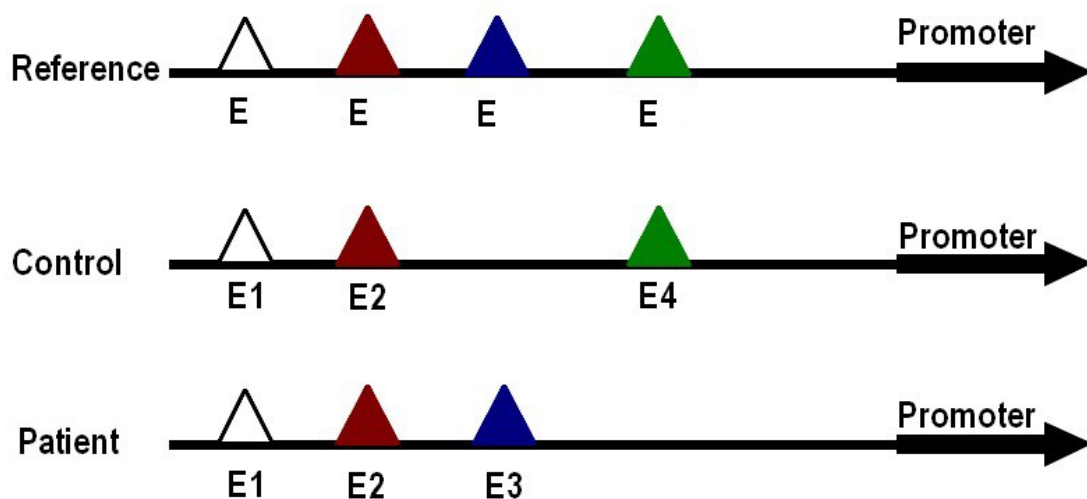
'E' denotes enhancers

Figure 25: Example of predicted outcome of ChIP-Seq assay in the control and patients. In the patient, the E3 peak is found to be active, while it is absent from the control. Therefore, analysing this particular region in the patient may give insight to its function. Again, the patient lacks E4 peak while it is functional in the control. If we examine the absence of this E4 peak in the patient, it may lead to new directions in the pathogenesis of a particular disease.

2.4 Software and Web Tools

2.4.1 Light Scanner Primer Design (LSPD)

LightScanner Primer Design (LSPD) software is a tool for rapid creation of high quality PCR assays for Hi-Res Melting® analysis.Multiple file formats, such as European Molecular Biology Laboratory (EMBL), FASTA, GenBank, or regular text files containing exons in upper case, can be imported into LSPD, allowing successful generation of HRM primers. The software also allows sequences to be copied and pasted directly from an Internet browser or other source. To ensure automatic exon designation within the software, it requires the exons to be in upper case. All sequences imported must contain at least 200 bases upstream and downstream of each exon for the software to recognize each exon appropriately. The suggested sequence source for

common genomes is UCSC Genome Browser (http://www.genome.ucsc.edu/cgi-bin/hgGateway).

2.4.2 Primer 3 Web Tool

Primer3 is a free open source online tool to design and analyse primers for PCR and real time PCR experiments. Reliable primer design is crucial for successful PCR. For the past two decades the Primer3 tool has been widely used for primer design in different field of biomedical sciences, often in high-throughput genomics applications. In designing PCR primers, the most challenging task is to design a primer pair devoid of secondary structures like primer dimer or hairpin formations. The Primer3 web tool is equipped with an enhanced core code of thermodynamic model, which facilitates the prediction of melting temperature and reduces the probability of forming such a secondary structure (Untergasser et al., 2012). It also gives precise control over primer placement in the genome, and allows different primer parameters to be saved and reused (Untergasser et al., 2012; Ye et al., 2012).

2.4.3 *In Silico* PCR

In silico PCR refers to a virtual PCR executed by a computer program with an input of a pair or a batch of primers against an intended genome (Yu and Zhang, 2011). It is a very useful and efficient method to determine the primer specificity in a diverse range of PCR applications. The most commonly used and popular web-based PCR tools are, (i) "In silico PCR" from UCSC (University of California, Santa Cruz) Genome Browser (http://genome.cse.ucsc.edu/)(5); (ii) "SNPCheck" from the National Genetics Reference Laboratory (Manchester, http://www.ngrl.org.uk/Manchester/-projects/snpcheck); and (iii) primer-BLAST from the National Centre for

Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/tools/primer-blast/) (Yu and Zhang, 2011).

2.4.4 uMelt- HRM Curve Prediction Tool

The uMelt is a flexible web-based tool for predicting DNA melting curves and denaturation profiles of PCR products. In this web tool, users need to input an amplicon sequence and a set of parameters for thermodynamic and experimental output. It also allows control over Mg^{++} concentrations and DMSO% used in the PCR reaction mix and also a temperature range to predict a melting curve (Figure 27) (Dwight et al., 2011).

Figure 26: uMelt web interface for curve prediction. A. Showing the predicted melting curve of the input amplicon, B. PCR amplicon input sequence, C. Advance control window to set parameters like MG^{++} and DMSO (%) concentration and also temperature range, D. Dissociation curve of input amplicon sequence.

Chapter 3

Study of TP63

PART B: Suggested Declaration for Thesis Chapter

Monash University

Declaration for Thesis Chapter 3

Declaration by candidate

In the case of Chapter 3 (Based on the article: Insertion/deletion polymorphisms in the DeltaNp63 promoter are a risk factor for bladder exstrophy epispadias complex), the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Performed experiments, Analyzed data and Contributed to the manuscript	10%

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Wilkins, S.	Conceived and designed the experiments, Performed the experiments, Analyzed the data	
Zhang, K. W.	Performed the Experiments, Analyzed the data	
Quantin, R.	Performed the Experiments, Analyzed the data	
D'Cruz, N.	Performed the Experiments	
Hutson, J.	Contributed reagents/materials/analysis tools	
Ee, M.	Contributed reagents/materials/analysis tools	
Bagli, D.	Contributed reagents/materials/analysis tools	
Aitken, K.	Contributed reagents/materials/analysis tools	
Fong, F. N.	Performed the Experiments, Analyzed the data	
Ng, P. K.	Performed the Experiments, Analyzed the data	
Tsui, S. K.	Performed the Experiments, Analyzed the data	
Fung, W. Y.	Performed the Experiments, Analyzed the data	
Banu, T.	Contributed reagents/materials/analysis tools	
Thakre, A.	Contributed reagents/materials/analysis tools	
Johar, K.	Performed the Experiments	
Jaureguizar, E.	Contributed reagents/materials/analysis tools	
Li, L.	Contributed reagents/materials/analysis tools	
Cheng, W.*	Conceived and designed the experiments	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work*.

		Date:
Candidate's Signature		**30/12/2014**

		Date:
Main Supervisor's Signature		**30/12/2014**

*Note: Where the responsible author is not the candidate's main supervisor, the main supervisor should consult with the responsible author to agree on the respective contributions of the authors.

PART B: Suggested Declaration for Thesis Chapter

Monash University

Declaration for Thesis Chapter 3

Declaration by candidate

In the case of Chapter 3 (Based on the article: No TAP63 promoter mutation is detected in bladder exstrophy-epispadias complex patients), the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Designed experiment, Analyzed data, Contributed to final manuscript	30%

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Darling, T.	Drafted and prepared the manuscript, Performed experiments, Analyzed data	60%
White, S.	Coordination and drafted the manuscript	
Cheng, W.*	Conceived the idea, Coordination and drafting of the manuscript	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work*.

Candidate's Signature		Date: 30/12/2014

Main Supervisor's Signature		Date: 30/12/2014

*Note: Where the responsible author is not the candidate's main supervisor, the main supervisor should consult with the responsible author to agree on the respective contributions of the authors.

Chapter 3: *TP63* Study in Human BEEC

3.1 Introduction

TP63 has been one of the few genes associated with Bladder Extrophy-Epispadias Complex ((BEEC; MIM600057) pathogenesis. As described in Chapter 1, *TP63* codes for two different isoforms, the pro-apoptotic *TAP63* and anti-apoptotic *ΔNP63*, through alternate promoter usage. These transcriptional regulatory proteins act in opposition to activate or repress the expression of downstream target genes, of which over five thousand target sites have been identified (Yang et al., 2006).

The study of the only genetic model of BEEC, the murine *p63* knockout model, has established that *p63* is a candidate gene for this disease. We previously showed that *p63* is expressed in bladder epithelium throughout its development, and that *ΔNp63* is the predominant isoform. In the murine $p63^{-/-}$ model, *ΔNp63* expression is reduced or absent and ventral bladder epithelial development is abnormal because of an increase in ventral bladder epithelial apoptosis (Cheng et al., 2006). In addition, $p63^{-/-}$ null mice also have phenotypes identical to the associated anomalies of BEEC patients (Table 4).

Table 4: The associated anomalies of BEEC patients and the p63-/- mouse phenotype.

Associated anomalies of BEEC patients	p63-/- knockout mice phenotype
Exomphalos (Quiroz-Guerrero et al., 2009)	Exomphalos (Cheng et al., 2006)
Ano-rectal malformations (Boemers et al., 1994; Singh et al., 2011)	Ano-rectal malformations (Ince et al., 2002)
Spina bifida, sacral hypoplasia, myelo-meningocele (Cadeddu et al., 1997)	Stunted tail (Mills et al., 1999)
Cleft palate, median cleft face syndrome (Celli et al., 1999; Rinne et al., 2006)	Cleft palate (Yang et al., 1999)
Epidermolysis bullosa (Moretti et al., 1995)	Non-stratified skin epidermis (Yang et al., 1999)
Absence of feet, tibial deformity (Celli et al., 1999)	Absent or stunted limb buds (Mills et al., 1999)

In human BEEC patients, down-regulation of *ΔNP63* and up-regulation of *TAP63* isoform expression is observed in bladder tissues of neonates (Ching et al., 2010; Wilkins et al., 2012). Previously, Ching and colleagues did not find any *TP63* exon mutations in human BEEC cases (Ching et al., 2010). The study of Ching et al. (2010) strongly supports the involvement of *TP63* in embryonic urogenital and ventro-caudal development in humans. It also showed that *TP63* is highly dysregulated in human BEEC.

This chapter intended to explain and establish whether the findings of $p63^{-/-}$ knockout model can be translated into a better understanding of human BEEC. This chapter is mainly based on our published article on analysis of *ΔNP63* and *TAP63* promoter region (Darling et al., 2013; Wilkins et al., 2012) where these two *P63* isoform were screened for potential sequence variants in human BEEC samples.

3.2 Material and Methods

3.2.1 The BEEC Research Patient Cohort

The Royal Children's Hospital (RCH) Melbourne Research Ethics Committee reviewed and approved the study. Informed consent of patients and the controls were obtained from the parents/guardians. Collaborators from overseas centres also had the study approved by their respective institutional ethic committees prior to sample collection.

3.2.2 Collected Samples

Buccal Swab: Buccal swab DNA samples (n = 112) were collected from different ethnic regions (India, Bangladesh, China, Australia, Spain, Canada and USA). DNA from the exfoliated cells was collected by rubbing the inside of the mouth on each side with a buccal swab. The samples were then placed in a storage medium and posted to

MIMR-PHI Institute of Medical Research via normal airmail, as the swab DNA is stable in the medium at room temperature for 2 months.

Tissue Sample: Six patient and one control bladder tissue samples were collected from Royal Children's Hospital, Melbourne. Samples were collected during reconstructive surgery of BEEC patients and were kept in the -80°C and later transported from Royal Children's to MIMR-PHI Institute of Medical Research in dry ice. All samples were codified and subject identities kept confidential.

3.2.3 Extraction of Genomic DNA

Isolation of Genomic DNA from buccal swab was carried out using the BuccalAMP DNA Storage and Extraction Kit (Epicentre®) as per manufacturer's protocol. The DNA concentration of each DNA sample was determined using a Nanodrop® ND-1000 spectrophotometer.

For the extraction of genomic DNA from human bladder tissue, approximately 10 mg of tissue was excised from the sample, and placed in a Ceramic Bead (1.4 mm) tubes for homogenization. Homogenization was performed in a Precellys®24 lyser/homogenizer (Bertin Technologies). The homogenization was done in 6500 rpm in two 20 min sessions with a 5 sec interval in between session. Genomic DNA was than extracted from the homogenized tissue following manufacturer's protocol of QIAamp DNA Mini Kit (Qiagen). To ensure RNA free genomic DNA, RNase A treatment was also performed following the kit protocol. Genomic DNA was quantified using NanoDrop 1000 Spectrophotometer (Thermo Scientific).

3.2.4 Quantitative Real-Time PCR

RNA was extracted from bladder urothelial tissue samples from BEEC patients. Normal bladder urothelium from cystectomy (cancer) specimens were used as controls. 20 mg

tissue was excised from both normal and patient tissue samples. Tissue samples were disrupted and homogenised using Ceramic Bead (1.4 mm) Tubes (Gene Works), mixed with 10 μl of β-mercaptoethanol (Sigma-Aldrich Pty. Ltd, Castle Hill, NSW, Australia) in buffer. Homogenisation was performed in a Precellys®24 lyser/homogenizer (Bertin Technologies). Homogenization was done in 6500 rpm in two 20 min sessions with a 5 sec interval in between sessions. The lysates were centrifuged and RNA was extracted with the RNeasy Mini Kit (Qiagen Pty. Ltd.) following manufacturer's protocol. cDNA synthesis was accomplished using High Capacity cDNA Synthesis Kit (Applied Biosystems) for downstream applications. Real-time qPCR was carried out using SYBR Green Master Mix (Applied Biosystem) on 7900HTk Fast Real-Time PCR (Applied Biosystem) machine, using published primer sequences (Chilosi et al., 2003) for *TAP63* and *ΔNP63*. Relative expression was analysed using Applied Biosystem's RQ Manager 1.2 software. ANOVA of RQ values were analysed using GraphPad Prism 5 software.

3.2.5 PCR and Sequencing of *ΔNP63* Promoter

Prior to sequencing, PCR amplification was performed using GoTaq Green PCR mater mix (Promega) following manufacturer's protocol. Four pairs of primers (ΔNP63PP1 to PP4) were used in PCR as described in Wilkins et al. (2012) to cover the 2785 bp promoter region (Figure 27). Primers pairs are listed in Appendix Table I. The extent of the *ΔNP63* promoter was determined by aligning and comparing the 5′ upstream regions of exon 3 of *TP63* sequences from humans, mouse and pig. The most conserved region 2785 bp (22696 to +89) was selected as the putative promoter region of *ΔNP63*.
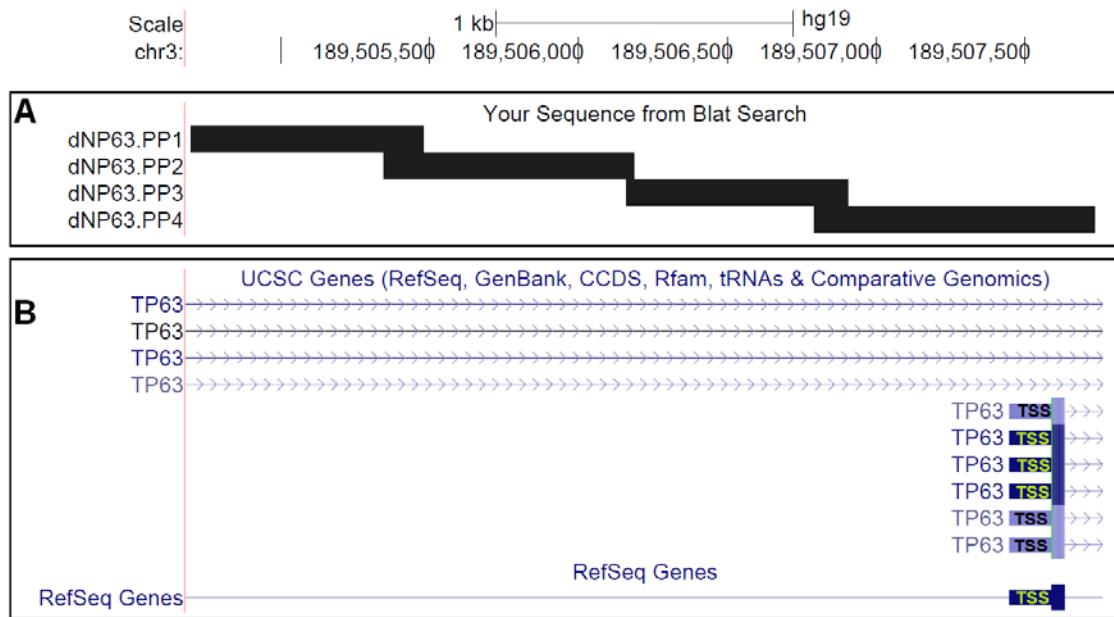
Figure 27: Position of the *in silico* PCR amplifiedproductof *ΔNP63* promoter regions. A. PCR products (Black Bars) were amplified by dNP63.PP1-PP4 sequencing primers. These primers are also indicating the estimated extent of *ΔNP63* promoter region (2785 bp). BLAT search was used to determine the amplicon position. B. *ΔNP63* gene transcription start site (TSS). These images are derived from the UCSC Genome Browser (http://genome.ucsc.edu).

The PCR products were purified using the QIAquick PCR Purification Kit (Qiagen) following manufacturers protocol. Amplified PCR products were then quantified and used as templates for direct DNA sequencing. Sequencing was carried out in both directions on an automated ABI Prism 3100 Genetic Analyzer, using the Big Dye Terminator kit (Applied Biosystems). Chromatographs were analyzed using Sequence Scanner software (Applied Biosystems) and compared with reference sequences of TP63 using BLAST (Altschul et al., 1990) from NCBI (http://www.ncbi.nlm.nih.gov/) and BLAT (Kent, 2002) from UCSC Genome Browser (http://genome.ucsc.edu/).

3.2.6 HRM and Sequencing of *TAP63* Promoter

Eight HRM primer pairs (TAP63.PP1 to PP8) were designed to cover the most conserved region of the putative *TAP63* promoter HRM and sequencing (Appendix

Table II). The conserved area was based on the multiple alignments of 100 vertebrate species (Figure 28) (Pollard et al., 2010; Siepel et al., 2005). An area 2.5 kb upstream of the *TAP63* transcription start site (TSS) was considered as promoter region, and analysed using the University of California, Santa Cruz (UCSC) genome browser (http://genome.ucsc.edu/cgi-bin/hgGateway) (Kent et al., 2002). A total of 112 BEEC buccal DNA samples were subjected to the screening process. The screening success ranged between 48 and 67 samples across each of the regions R1, R5, R6, R7 and R8 in the *TAP63* promoter (Figure 28).

HRM analysis was carried out to identify any potential sequence variants within the chosen regions of the promoter (Figure 28). HRM was performed in 10 μl reaction volumes, consisting of HRM Master Mix (TrendBio), 5 μM each of forward and reverse primer, and 25 ng of genomic DNA sample from BEEC patients. PCR reactions were carried out under the following conditions; an initial hold at 95°C for 2 min, followed by 45 cycles of 94°C for 30 sec and 58°C for 30 sec, in 96 well plate (Skirted- Black & White form Biorad). Amplified products were analysed using a LightScanner HR96 machine (Idaho Technologies, USA) with a melting range of 70–98 °C and a hold temperature at 67°C. Curve analysis was performed using the LightScanner HRM analysis software (Version 2.0.0.1331) (Idaho Tech.). The presence of a single PCR product of expected size was confirmed with agarose (1%) gel electrophoresis.
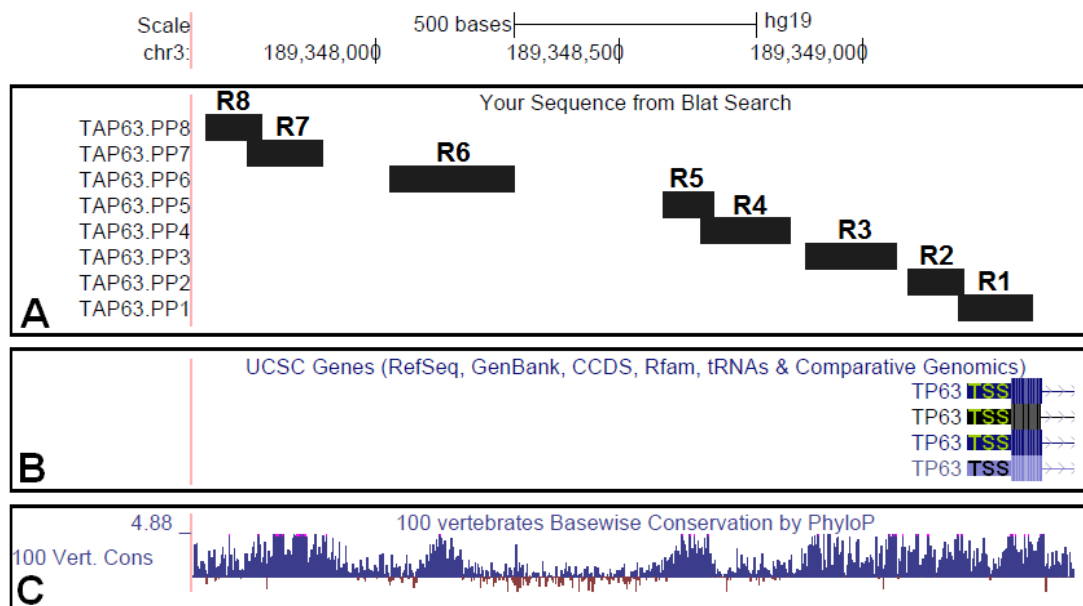
Figure 28: Position of *TAP63in silico* PCR product for HRM and sequencing from UCSC genome browser. A. Primers are indicated as TAP63.PP1-PP8 (region R1-R8) from BLAT search. B. The transcription start site (TSS) of the *TAP63* isoform. C. Promoter region was selected based on the most conserved region upstream of TSS as indicated in UCSC 100 vertebrates' conservation. The base wise conservation on 100 vertebrate species was carried out by PhyloP (Pollard et al., 2010). Image Source: UCSC Genome Browser (http://genome.ucsc.edu).

Potential sequence variants were confirmed through Sanger sequencing on an Applied Biosystems automated ABI Prism 3100 Genetic Analyzer and the Big Dye Terminator kit (Applied Biosystems). Both forward and reverse sequence was carried out using primers listed in Table 3. Sequence chromatograph trace files were viewed using the Sequence Scanner (Applied Biosystems). Variant calling was done visually, and sequences were aligned to the genome using BLAST (Altschul et al., 1990) and BLAT (Kent, 2002).

3.3 Results

3.3.1 *ΔNP63* and *TAP63* expression in BEEC Patients

To establish if *TP63* plays a role in human bladder exstrophy, we compared the expression of both *TP63* isoform in bladder mucosa of a Caucasian BEEC patient (P1; <1 year old) (Figure 29). This patient showed decreased *ΔNP63* expression, whereas *TAP63* expression was increased compared with normal controls.
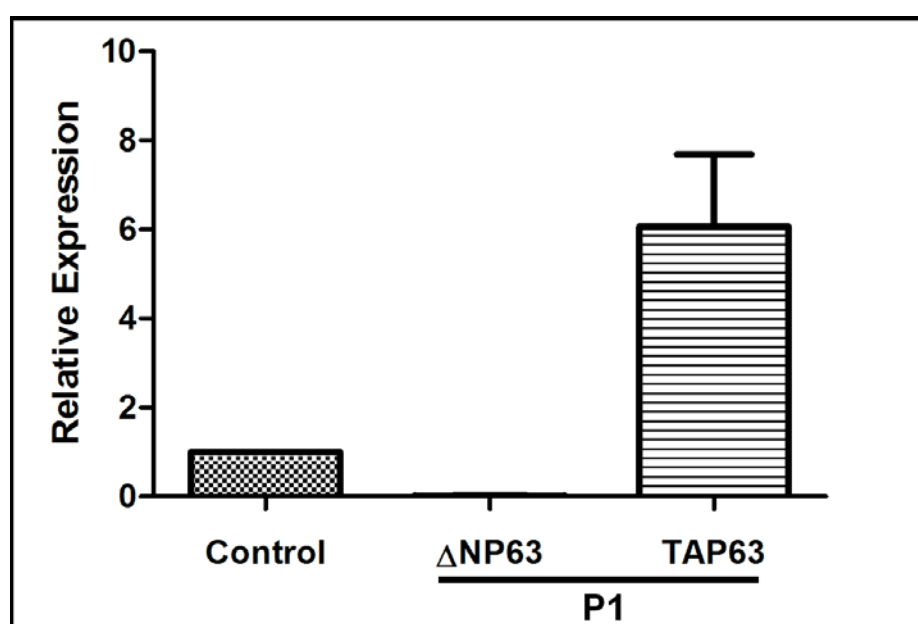
Figure 29: Real-time qPCR (n=6) showing reduced *ΔNP63* mRNA and increased *TAP63* mRNA expression in bladder mucosa from a Caucasian BEEC patient (<1 year old male) compared with normal control (normalized to 1). Error bar indicates standard error of mean (SEM) referring to repeats (n=6) of the experiments.

TAP63 expression was also checked on three additional BEEC patients (P2, 13 yr. male; P3, 14 yr. female; P4, 7 yr. male) on bladder mucosa tissue. Compared to the control samples, *TAP63* shows a variable expression pattern in patient samples (Figure 30). *ΔNP63* expression was also measured in the same patient samples (P2, P3 and P4) (Figure 31). It was found that the *ΔNP63* is expressed in higher level than that of *TAP63* in these three patients.
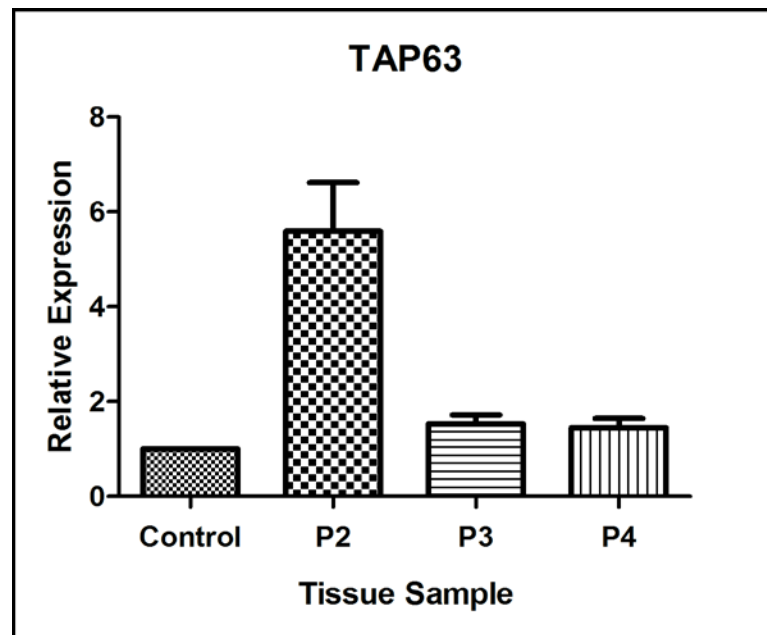
Figure 30: Real-time qPCR (n=3) showing *TAP63* mRNA expression in bladder mucosa from three Caucasian BEEC patients (P2, 13 yr. male; P3, 14 yr. female; P4, 7 yr. male) compared with a control sample (normalized to 1). Error bar indicates standard error of mean (SEM)
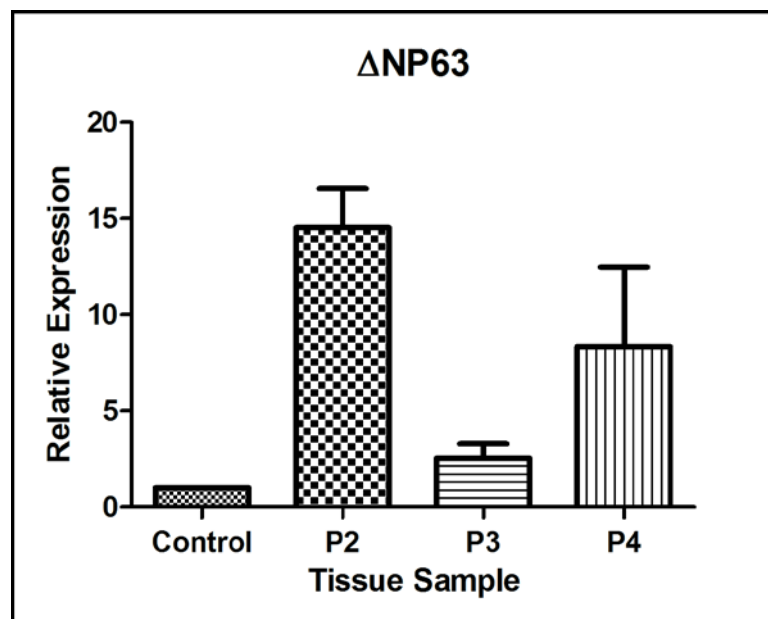
Figure 31: Real-time qPCR (n=3) showing *ΔNP63* mRNA expression in bladder mucosa from a three Caucasian BEEC patients (P2, 13 yr. male; P3, 14 yr. female; P4, 7 yr. male) compared with a control sample (normalized to 1). Error bar indicates standard error of mean (SEM).

3.3.2 Sequence Variation in the *ΔNP63* Promoter

To explain reduced *ΔNP63* expression (Ching et al., 2010; Wilkins et al., 2012), sequencing of the *ΔNP63* promoter (2700 nucleotides upstream of exon 3) was carried out in four BEEC patients. Three insertion/deletion polymorphisms (rs6148242, rs5855273 and ss#541028600) were identified in two patients (P3 and P4). The genotype of the four patients is shown in Table 5. Interestingly, expression analysis on these patients showed significant difference in expression pattern (Figure 30 and 31). The heterozygous 12 base-pair (bp) deletion (rs6148242) and 1bp insertion (rs5855273) was found in both patient P3 and P4. The heterozygous 4bp insertion (ss#541028600) was found only in P4, and P3 showed a homozygous 4bp insertion (Figure 32 and 33). Compared to the expression of *ΔNP63* in P4, it is possible that decreased level of *ΔNP63* expression in P3 may be a result of this homozygosity.

Table 5: Genotypes of BEEC patient samples used in real-time expression experiments.

Promoter Position	Genotype				Sequence Information	
	Patient 1 (P1)< 1 year old	Patient 2 (P2)≥ 13 year old	Patient 3 (P3)≥ 14 year old	Patient 4 (P4)≥ 7 year old	Reference Allele	SNP/ Indel Polymorph.
-2651	C/C	C/C	C/T	C/T	C	rs2138247
-2293 to -2282	TCCAGAA TCTT/TC CAGAATC TT	TCCAGAA TCTT/TC CAGAATC TT	TCCAGAA TCTT/-	TCCAGAA TCTT/-	TCCAGAAT CTT	rs6148242 (12 bp del.)
-1287	T/T	-/-	T/-	T/-	-	rs5855273 (1 bp ins.)
-1209	T/T	C/C	C/C	C/C	T	rs1464118 SNP
-71	AGAG/AG AG	AGAG/-	AGAG/AG AG	AGAG/-	-	ss#541028600 Novel 4 bp ins.
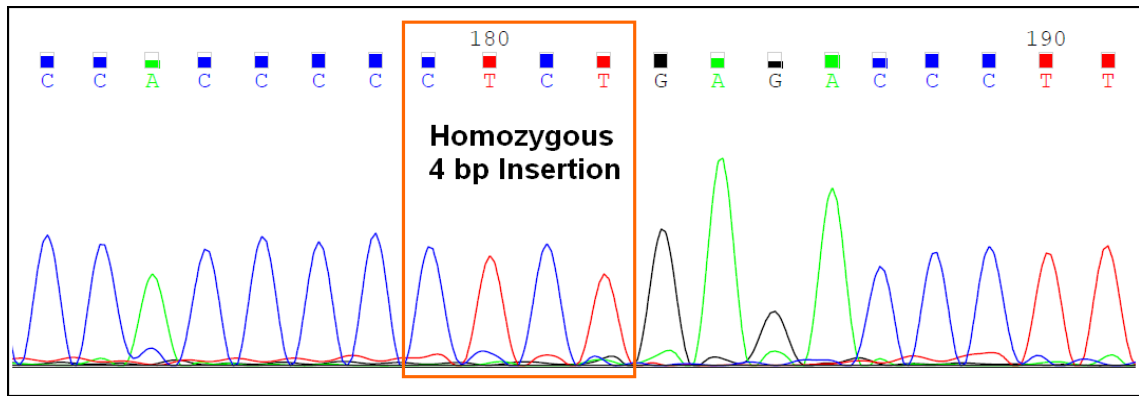
Figure 32: Sequence trace of the P3 patient showing homozygous 4bp insertion (Reverse Strand) in *ΔNP63* promoter region. Orange rectangle is indicating inserted nucleotide.

Figure 33: Alignment of *ΔNP63* chromatogram in P3 patient with reference sequence showing the position of homozygous 4bp insertion (Pink rectangle). 'Query' sequence is the chromatograph trace from the patient aligned with the 'Sbjct' *i.e.* reference sequence (NV_000003.12) from NCBI.

3.3.3 Sequence Variation in the *TAP63* Promoter

To establish possible explanation for the up-regulation *TAP63* in neonates and infants (Ching et al., 2010; Wilkins et al., 2012), *TAP63* promoter sequencing was carried out on the targeted regions (R1-R8) as described earlier (Figure 28, Appendix Table II). HRM analysis of R1 region on 48 samples identified a known common SNP (rs28673064, A/T) with a minor allele frequency of A = 0.485 (Figure 34A). Analysis R5 and R8 displayed uniform curve shapes across 60 and 56 patients respectively (Figure 34B).
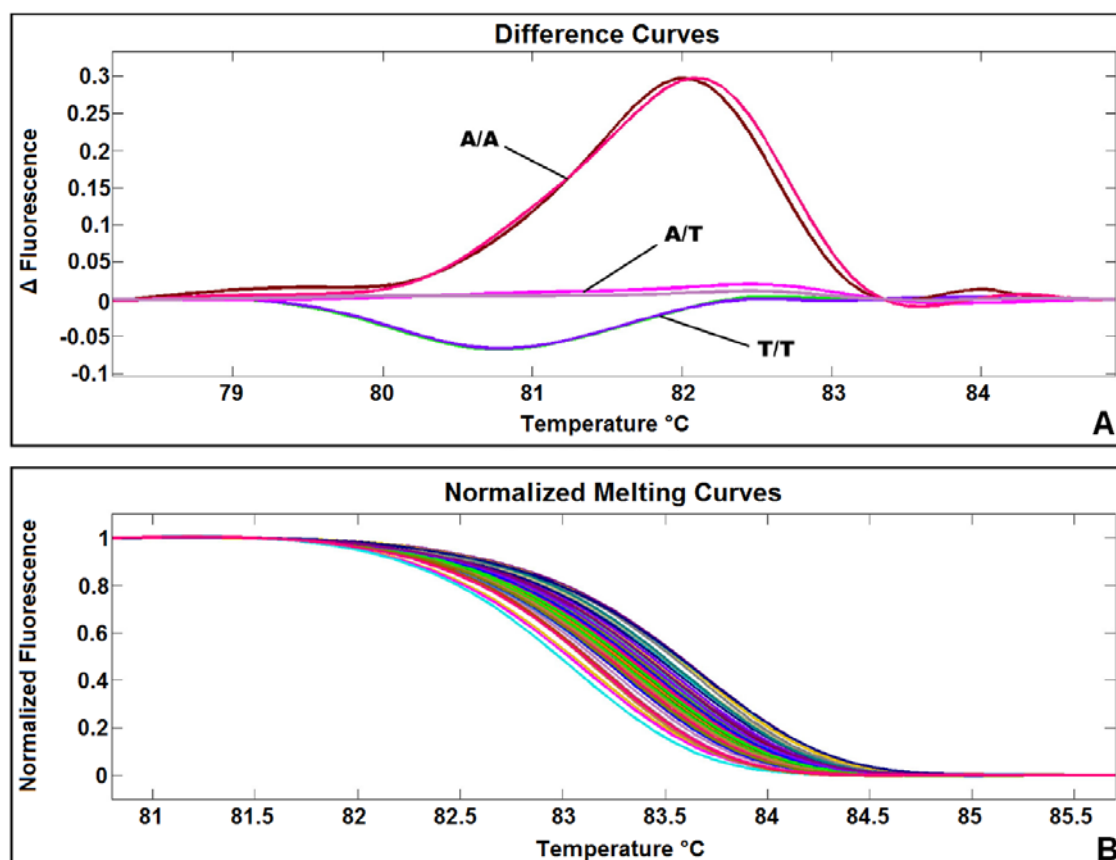
Figure 34: HRM analysis of *TAP63* promoter region. A. SNP variant (rs28673064, A/T) with difference curves in R1 region showing heterozygous and homozygous variant. B. Normalized curves of R5 region (60 patient samples) showing uniform curve pattern. In both section (A and B) fluorescence ('y' axis) was plotted against temperature ('x' axis) as described in Chapter 2: Methodology to generate a melting profile.

Sequencing of these regions did not identify any sequence variation as aligned to the current standard UCSC human reference genome sequence (February 2009 GRCh37/hg19). Although, R6 showed variable curve patterns, following no variation was detected. HRM and sequencing analysis of R7 region have identified another known heterozygous SNP (rs138456383, A/G) with minor allele frequency (MAF) G=0.004 (1000 Genomes). However, no novel sequence variant within the selected *TAP63* promoter region was detected (Darling et al., 2013) that could contribute to the BEEC pathogenesis.

3.3.4 Bioinformatics Analysis of the *ΔNP63* Promoter Region

Analysis of transcription factor binding (TFB) sites in the *ΔNP63* promoter using MATCH software, predicted that the 12 bp in/del might affect binding of GATA3, GATA6, GFI1, LEF1, TCF1 and SOX10, whereas the 4 bp in/del may affect binding of EGR1 and CBF transcription factors. Bioinformatics analysis using UCSC genome browser also revealed some important TFB sites in this promoter region related to gene transcription and development, such as FOS, JUND, E2F4 and POLR2A (Figure 35). Major functions of these transcription factors and HGNC (HUGO Gene Nomenclature Committee) ID are listed in the Table 6.

The presence of a 4bp in/del in the POLR2A binding region in the P3 and P4 patients may be the region responsible for variable expression compared to the P2 patient, because it is the largest subunit of RNA polymerase II, the polymerase responsible for synthesizing messenger RNA in eukaryotes (Mita et al., 1995). Further studies would be needed to verify these interactions.
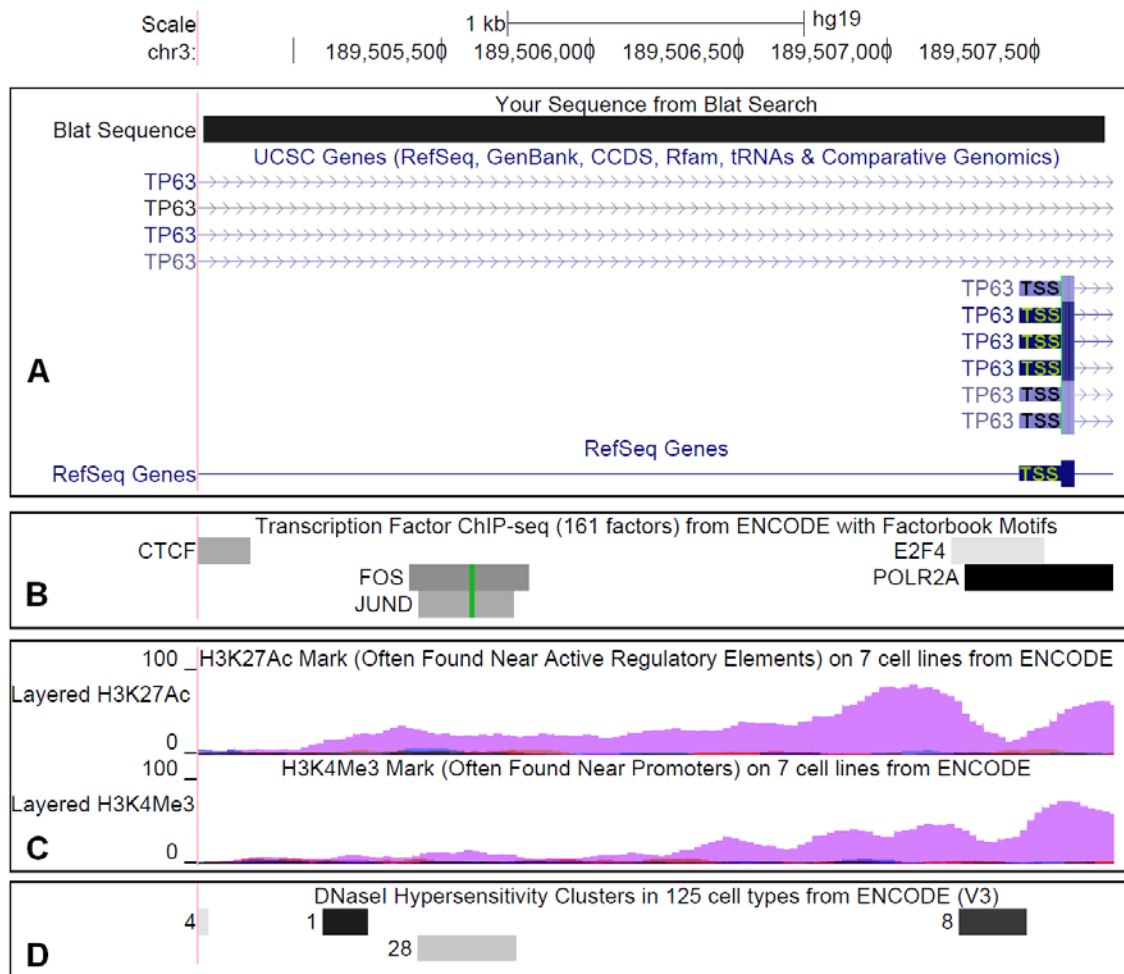
Figure 35: Bioinformatics analysis of *ΔNP63* promoter region using UCSC genome browser. A. BLAT sequence of targeted *ΔNP63* promoter in *P63* gene, also showing TSS of the respective *ΔNP63* isoform. B. Region showing transcription factor binding motifs in the promoter, based on data for 161 transcription factors profiled as part of the ENCODE project (ENCODE Project Consortium, 2012). C. Enriched region of histone marks (H3K27ac and H3K4me3) found near promoter region, as found in the 7 cell lines by ENCODE. Each layered colour code denotes different cell line and their enrichment e.g. the pink colour denotes enrichment of respective histones in Normal Human Lung Fibroblast (NHLF) cell line. D. DNaseI hypersensitivity clusters from ENCODE. The number on the left side e.g. 28 means that it was found in 28 cell types out of 125 cell type examined. Image Source: UCSC Genome Browser (http://genome.ucsc.edu).

Table 6: Transcription factor binding sites in *ΔNP63* promoter region.

Transcription Factors	Functions	References	HGNC ID
E2F4 (E2F Transcription Factor 4, P107/P130-Binding)	• Plays a crucial role in the control of cell cycle. • Also control the action of tumor suppressor proteins. • Have an important role in the suppression of proliferation-associated genes.	Ginsberg et al. (1994); Sardet et al. (1995)	HGNC:3118
EGR1 (Early Growth Response 1)	• Functions as a transcriptional regulator. • Activates the transcription of target genes required for mitogenesis and differentiation.	Sukhatme et al. (1988)	HGNC:3238
FOS (FBJ Murine Osteosarcoma Viral Oncogene Homolog)	• Functional component of the AP1 transcription factor complex • Identified as regulators of cell proliferation, differentiation, and transformation. • Sometimes, it has also been associated with apoptotic cell death.	Monje et al. (2005); Bossis et al. (2005); Konsman and Blomqvist (2005)	HGNC:3796
GATA3 (GATA Binding Protein 3)	• Important regulator of T cell development. • Plays an important role in endothelial cell biology. • Facilitate core promoter sequence-specific DNA binding.	Joulin et al. (1991); Yamashita et al. (2004)	HGNC:4172
GATA6 (GATA Binding Protein 6)	• Important for regulating terminal differentiation and/or proliferation of cells. • Play an important role in organogenesis during vertebrate development.	Suzuki et al. (1996); Morrisey et al. (1996)	HGNC:4174
JUND (Jun D Proto-	• Functional component of the AP1 transcription factor complex.	Nomura et al. (1990); Berger	HGNC:6206

Transcription Factors	Functions	References	HGNC ID
Oncogene)	● Proposed to protect cells from p53-dependent senescence and apoptosis.	and Shaul (1991)	
LEF1 (Lymphoid Enhancer Binding Factor 1)	● Have a role during skin development, hair cell differentiation and follicle morphogenesis. ● Bind to functionally important site in the T-cell receptor-alpha enhancer, thereby conferring maximal enhancer activity.	Milatovich et al. (1991); Jamora et al. (2003)	HGNC:6551
POLR2A [Polymerase (RNA) II (DNA Directed) Polypeptide A, 220kDa]	● This DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA. ● It synthesizes mRNA precursors and many functional non-coding RNAs.	Ni et al. (2011); Mita et al. (1995)	HGNC:9187
SOX10 [SRY (Sex Determining Region Y)-Box 10]	● Involved in the regulation of embryonic development and in the determination of the cell fate. ● Act as a transcriptional activator after forming protein complex with other proteins.	Pingault et al. (1998); Bondurand et al. (1999); Huber et al. (2003)	HGNC:11190

3.3.5 Bioinformatics Analysis of the *TAP63* Promoter Region

Bioinformatics analysis of the highly conserved *TAP63* promoter region identified the presence of multiple epigenetic markers of transcriptional regulation within highly conserved areas of the genome. Histone markers, DNase1 hypersensitivity regions, CpG islands and RNA polymerase binding sites indicative of transcriptional regulatory regions were identified. Putative binding sequences for the transcription factors MYC, FOXA1, RELA, MEF2A, EP300 and TBP were also found within 2 kb of the *TAP63* transcription start site (Figure 36, Table 7).
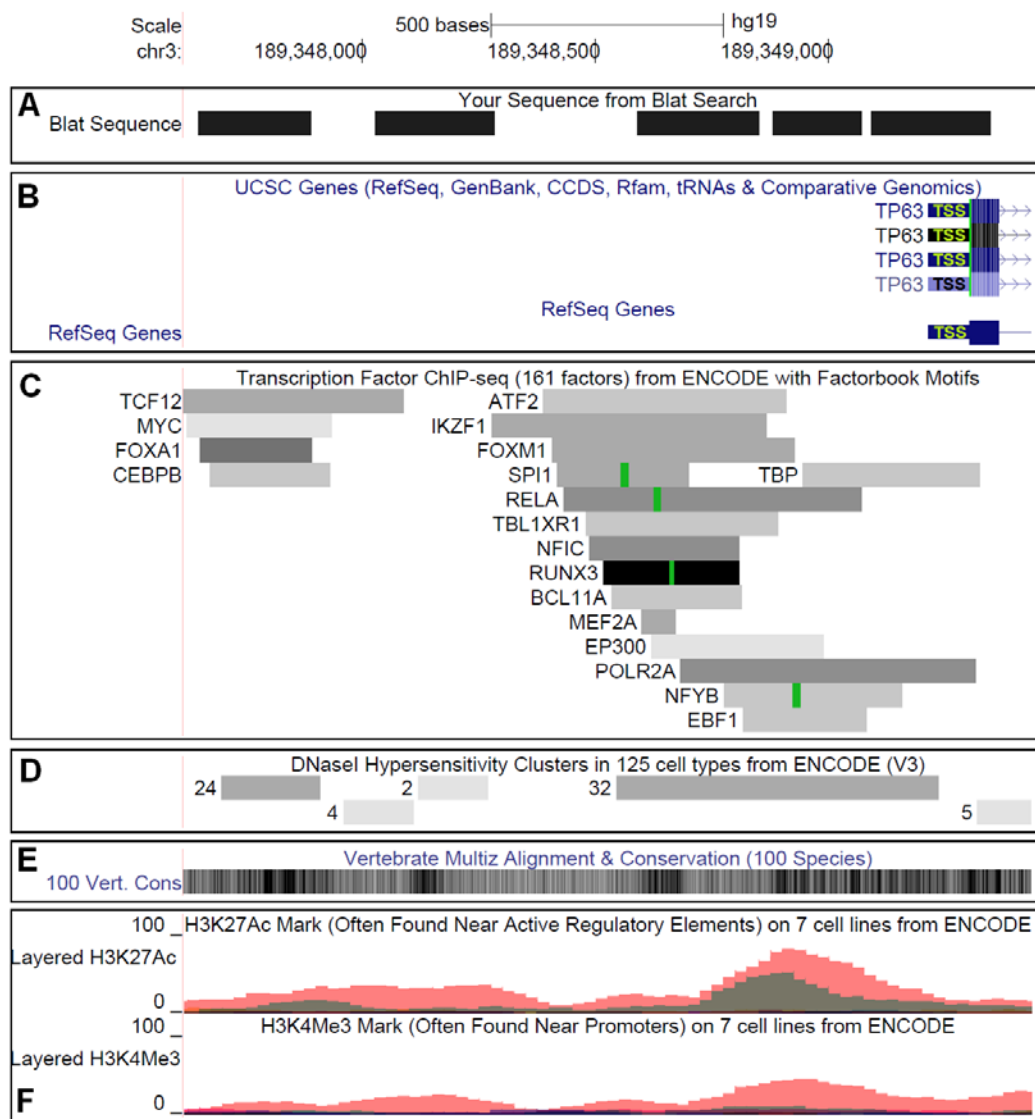
Figure 36: Bioinformatics analysis of *TAP63* promoter region. A. BLAT sequence of targeted regions in *TAP63* promoter flanked by designed primers. B. Reference sequence of respective isoform with TSS. C. Transcription factor binding motifs on the promoter, based on data for 161 transcription factors profiled as part of the ENCODE project (ENCODE Project Consortium, 2012). D. ENCODE DNase1 hypersensitivity regions. The number on the left side e.g. 32 means that it was found in 32 cell types out of 125 cell type examined (ENCODE Project Consortium, 2012). E. Highly conserved areas of the promoter region from 100 vertebrate species. The targeted regions of the promoter were selected based on these areas. F. Enriched region of histone marks (H3K27ac and H3K4me3) found near regulatory and promoter region, as found in the 7 cell lines by ENCODE (ENCODE Project Consortium, 2012). Each layered colour code denotes different cell line and their enrichment.

Table 7: Transcription factor binding sites in *TAP63* promoter region.

Transcription Factors	Functions	References	HGNC ID
EP300 (E1A Binding Protein P300)	Functions as histone acetyltransferase that regulates transcription via chromatin remodeling. Important in the processes of cell proliferation and differentiation and also have role in epithelial cancer.	Eckner et al. (1994)	HGNC:3373
FOXA1 (Forkhead Box A1)	Involved in embryonic development and establishment of tissue-specific gene expression Regulates gene expression in differentiated tissues and also interact with chromatin	Bingle and Gowan (1996); Mincheva et al. (1997)	HGNC:5021
MEF2A (Myocyte Enhancer Factor 2A)	Involved in several cellular processes, including muscle development, neuronal differentiation, cell growth control, and apoptosis.	Yu et al. (1992)	HGNC:6993
MYC (V-Myc Avian Myelo-cytomatosis Viral Oncogene Homolog)	Plays a role in cell cycle progression, apoptosis and cellular transformation. Activates the transcription of growth-related genes.	Dominguez-Sola et al. (2007); Persson and Leder (1984)	HGNC:7553
POLR2A	Described in Chapter Table 6		HGNC:9187
RELA (V-Rel Avian Reticulo-endotheliosis Viral Oncogene Homolog A)	Homodimeric/Heterodimeric complex of ubiquitous transcription factor NFĸB is formed by RELA. Activates NFĸB signal transduction events related to tumorigenesis, inflammation, and immunity, Also initiate NFĸB for differentiation, cell growth, and apoptosis.	Nolan et al. (1991)	HGNC:9955
TBP (TATA Box Binding Protein)	Functions at the core of the DNA-binding multiprotein factor TFIID to initiate transcription by RNA polymerase II.	Kao et al. (1990); Nakamura et al. (2001)	HGNC:11588

3.4. Discussion

3.4.1 *ΔNP63* Isoform of *TP63*

Wilkins and co-workers previously showed that *ΔNP63* is predominantly expressed in normal human tissue (foreskin) whereas *TAP63* is expressed at a lower level. This study found in a one year old patient that the *ΔNP63* expression was down-regulated while *TAP63* expression is up-regulated. This finding corroborates the findings of Wilkins and colleagues where they found the same expression pattern in a neonate patient (2 days old) (Wilkins et al., 2012). However, in three older patients (two male patients, 13 and 7 year old; and a 14 year old female), *TAP63* and *ΔNP63* showed variable expression pattern compared to the normal control. Such expression of *TAP63* and *ΔNP63* in older patients may be due to reconstructive surgery, where they have undergone Mitroffanof procedures (using appendix as a conduit for bladder catheterization) and bladder augmentation (using small bowel patch to increase bladder volume). Also, the irritation from small bowel mucus secretion and various degree of cystitis (from bacteria introduced by catheters) may affect *TP63* expression in the urothelia sample. We and others found that the *ΔNP63* expression in some BEEC patients is significantly reduced in bladder urothelia (Ching et al., 2010; Wilkins et al., 2012). However, such post-natal expression does not conclusively represent the expression pattern to that of the embryological development. The low sample size of post-natal bladder urothelia analysed in this study is not enough to draw a statistically convincing conclusion, but it support a possible involvement of *ΔNP63* in BEEC pathogenesis.

Sequence analysis of the *ΔNP63* promoter in four BEEC patient samples revealed three in/del polymorphisms (rs6148242, 12bp deletion; rs5855273, 1bp insertion; and ss#541028600, 4bp insertion). Wilkins et. al. (2012) showed in their study that these polymorphisms were significantly associated with increased risk of BEEC. They also

reported that these in/del polymorphisms may significantly reduce the transcriptional efficiency compared to the control lacking these polymorphisms (Wilkins et al., 2012). The bioinformatics analysis of this promoter showed the binding of GATA3, GATA6, LEF1, TCF1 and SOX10 may be affected by the 12bp deletion.

It has been reported that BEEC patients have 700 fold increased chance of developing bladder cancer (Ricketts et al., 1991). A major portion of BEEC patient (60%) suffers from bladder carcinoma in the fourth and fifth decades of life (de Riese and Warmbold, 1986; Facchini et al., 1987; Sharma et al., 2013). Liang et. al., (2014) showed that GATA-3 expression was present in 70% (72/103) of conventional bladder Urothelial Carcinomas. GATA-3 is a member of zinc finger transcription factors. Its function is also linked to renal development (Muroya et al., 2001) and disease associated with GATA-3 is renal dysplasia which is one of the BEEC associated anomalies. Like p63, the relation of GATA-3 to development and tumorogenesis suggests a potential role in BEEC pathogenesis.

Like GATA3, GATA6 is also a member of zinc finger transcription factors. It is involved in the regulation of cellular differentiation and organogenesis during vertebrate development at early embryogenesis (Morrisey et al., 1998). Morrisey et. al. (1996) reported that GATA6 may have unique function(s) during the development of urogenital systems (Morrisey et al., 1996). Qiao and colleagues reported two sequence variations within the GATA6 gene promoter may contribute to the development indirect inguinal hernia (IIH) as a risk factor by changing the GATA6 levels (Qiao et al., 2014). It is possible that the 12bp deletion in GATA-6 binding motif on *ΔNP63* promoter result in an abnormal development of bladder during organogenesis leading to BEEC. However, detail study in large patient cohort is necessary confirm such interactions.

Bioinformatics analysis also predicted a LEF1 transcription factor binding site in the 12bp deletion position. It has been reported by Behrens and colleagues that β-catenin and LEF1 interaction provides a molecular mechanism for the transmission of signals from cell-adhesion components (Cadherin) to the cell nucleus (Behrens et al., 1996; Muller et al., 2002). Several articles showed that the association of the cadherin/catenin complex with the cytoskeleton is essential for tight cell–cell interaction (Knudsen et al., 1995; Rimm et al., 1995). Specially, the association with β-catenin is necessary for cadherin-mediated cell-to-cell adhesion (Peifer and Polakis, 2000; Polakis, 2000). Muller et. al. (2002) showed that β-catenin together with LEF1/TCF transcription factor drives the expression of the genes necessary for epithelial-mescenchymal transition. Interestingly, Baskin and co-workers found that the epithelial-mesenchymal interaction is essential in bladder mesenchymal (smooth muscle) development (Baskin et al., 1996a; Baskin et al., 1996b). So, it is tempting to speculate that disrupted LEF1 binding in *ΔNP63* promoter leads to abnormal epithelial-mesenchymal signals which eventually result in BEEC.

Bioinformatics analysis of the 4 bp insertion within the *ΔNP63* promoter predicted an EGR1 binding site. In order to maintain cellular homeostasis, EGR1 respond to a wide variety of stimuli such as many extracellular signalling molecules including hormones, neurotransmitters, growth and differentiation factors, and cytotoxic metabolites. The transcription factor then immediately exerts its effects in response to that stimuli (Thiel and Cibelli, 2002). Biophysical analysis conducted within the target gene promoters reveals that DNA sequence variations can tightly modulate the EGR1 binding and more detrimental to EGR1-DNA interaction (Mikles et al., 2014). It is possible that the sequence variation in the EGR1 binding site in *ΔNP63* promoter altered the

transcription of the respective gene. As a result of this, development of the bladder is compromised leading to BEEC.

FOS and JUN is the member of a basic lucine zipper (bZip) family. AP-1 is a sequence-specific transcription factor complex formed by JUN, FOS or ATF (activating transcription factor) subunits (Angel and Karin, 1991; Johnson and McKnight, 1989; Karin et al., 1997). These proteins are associated to form a variety of homodimers and heterodimers that bind to a common site in the DNA (Angel and Karin, 1991). AP-1 is vital to number of cellular processes including cell growth, proliferation, survival, apoptosis, differentiation and cell migration in eukaryotes (Johnson and McKnight, 1989; Karin et al., 1997). Sequence variation in the FOS and JUN binding site in *ΔNP63* promoter may interfere with the formation of AP-1 activator complex resulting in an impaired signal to drive the cellular processes in bladder development. However, more research is needed to confirm such speculation.

E2F4 is a member of the E2F protein family. DeGrogori and co-workers demonstrated that a large number of genes encoding proteins important for cell cycle regulation and DNA replication can be activated by the E2F proteins (DeGregori et al., 1997). Very recently, ChIP-seq analysis conducted by Gokhman et. al. (2013) reveals that the E2F4 is a master transcriptional regulator of histones (Gokhman et al., 2013). As histones are the building blocks of eukaryotic chromatin, and are essential for genome packaging, function and regulation, it is possible that variation in the E2F4 binding site impaired the regulation of the *ΔNP63* isoform resulted in altered function.

Disruption of a POLR2A binding site in the 4 bp insertion region of *ΔNP63* promoter may interfere with gene transcription e.g., tissue and development specific, stress response, etc. This protein interacts with the promoter regions of genes as well as with a

variety of elements and transcription factors to determine essentially all of the transcriptional parameters (Buratowski, 1994; Kornberg and Lorch, 1991).

3.4.2 *TAP63* isoform of *TP63*

Yang and co-workers have established that the *TAP63* isoform antagonizes *ΔNP63* and increased apoptosis (Yang et al., 1998; Yang and McKeon, 2000; Yang et al., 1999). We hypothesized therefore that the *TAP63* promoter region may contain sequence variants that contribute to BEEC pathogenesis. However, our study of the *TAP63* promoter region did not reveal any sequence variation (Darling et al., 2013), suggesting that the dysregulation in *TAP63* expression observed in the few patients (Ching et al., 2010; Wilkins et al., 2012) could not be attributed to sequence variants in the analysed regions of the *TAP63* promoter. In the absence of sequence variation in *TAP63* promoter, the dysregulation may be due to- i) ageing in the patient which lead to a potentially different *TAP63* expression in bladder, ii) exposure of bladder to exterior during early onset of the disease, iii) different surgical practices to reconstruct the urogenital system. It is also possible that the presence of transcription binding site in the promoter region related to development, cell growth, proliferation and apoptosis are disrupted. Bioinformatics analysis of *TAP63* promoter regions have identified some important transcription factor binding sites such as MYC, FOXA1, RELA, MEF2A, EP300 and TBP. Due to the rarity of the disease, fresh tissue sample from neonates are very difficult to obtain. It is therefore difficult to perform any statistical analysis of *TAP63* mRNA expression in BEEC patient to achieve a conclusive outcome. Study of mRNA expression in larger patient cohort may shade some light on the extent of *TAP63* dysregulation in BEEC patients. Also the regulation and interaction of different transcription factor may be able to give some promising insight.

Expression analysis of *TAP63* isoform in BEEC neonates and infants showed down regulation of this gene. It is possible that the presence of MYC transcription factor in the promoter may have such an effect, because it has been reported in several articles that MYC may also repress gene expression (Peukert et al., 1997; Wu et al., 1999).

It has been shown in several articles that members of the FOXA subfamily of transcription factors, especially FOXA1, are involved in normal urogenital development and differentiation (Mauney et al., 2010; Oottamasathien et al., 2007; Thomas et al., 2008; Varley et al., 2009). DeGraff and co-workers showed that the loss of this urothelial differentiation marker, FOXA1 is associated with bladder cancer (DeGraff et al., 2012). Recently, Wang et. al. (2013) reported FOXA1 has a potential function in regulating epithelial-to-mesenchymal transition (Wang et al., 2013a). Previously, Baskin and colleagues found that the epithelial-mesenchymal interaction is essential in bladder (smooth muscle) development (Baskin et al., 1996a; Baskin et al., 1996b). From these research, it is evident that FOXA1 is directly associated with urothelial development. Detailed study of this transcription factor in BEEC patient may reveal a possible mechanism behind the pathogenesis.

NFƙB (NF-kappa-B) is one of the most important and ubiquitous transcription factors (DiDonato et al., 2012). It is found in numerous cell types that express cytokines, chemokines, growth factors, cell adhesion molecules, and some acute phase protein (Baldwin, 1996). This transcription factor is usually held at an inactive state, inhibited by I-kappa-B proteins in the cytoplasm. To act as a transcriptional regulator NFƙB is bound with RELA or RELB to form a complex (Chen et al., 1999). Nenci et al. showed that transcription factor NFƙB is a master regulator of pro-inflammatory responses, and functions in gut epithelial cells to control epithelial integrity (Nenci et al., 2007).

EP300 is a histone acetyltransferase that regulates transcription via chromatin remodelling. This transcription factor is important in the regulation of cell proliferation and differentiation (Gayther et al., 2000). Visel and co-workers showed that *in vivo* mapping of p300 binding is very useful in the study of the role of tissue-specific enhancers (Visel et al., 2009). Genome wide binding study EP300 may facilitate the identification of regulatory enhancers and their associated activities in BEEC patient bladder sample.

POLR2A and TBP are mainly responsible for the initial start of transcription of respective gene (Ni et al., 2011; van Nuland et al., 2013). It has been shown by Savinikova and colleagues that variations in TBP binding sites are associated with several diseases such as α-, β- and δ-thalassemia, myocardial infarction and thrombophlebitis, changes in immune response, amyotrophic lateral sclerosis, lung cancer and haemophilia B Leyden (Savinkova et al., 2013). So, identification of sequence variation in such a site may help find out more about BEEC pathogenesis.

Several publications showed that sequence variation in promoter region can be a risk factor in numerous diseases such as pneumoconiosis (Ni et al., 2009), auto-immune diseases (Miceli-Richard et al., 2009), asthma (Burchard et al., 1999), and β-thalassemia (Kulozik et al., 1991). In Chinese populations, Sun and co-workers identified a six-nucleotide deletion (-652 6N del) variant in the CASP8 promoter that is a risk factor for multiple cancers such as lung, oesophagus, stomach, colorectal, breast and cervix (Sun et al., 2007). However, study of this mutation in European and US populations have not shown the same association to breast, prostate, and colorectal cancer (Frank et al., 2008; Haiman et al., 2008; Pittman et al., 2008). Such research corroborates with the finding of Wilkins et. al. (2012), where the prevalence and the role of in/del polymorphisms differed significantly between Caucasian and non-Caucasian ethnic populations. This

also suggests that the absence/presence of mutation in *TP63* may differ between the ethnic origins of the patient cohort. However, more research is required to confirm such concept.

3.5 Conclusion

Expression analysis of *TAP63* and *ΔNP63* showed variable expression pattern in an older patient. Being a predominant isoform, *ΔNP63* is expressed in higher level in older patient than *TAP63*. However, in neonates and infants the *ΔNP63* expression was reduced while *TAP63* was increased compared to control sample, which mimics the findings of knockout *Tp63* murine BEEC model (Cheng et al., 2006). Although sequencing of *TP63* exons in BEEC patients did not reveal any mutations, the above and published sequencing data of the *ΔNP63* promoter revealed three in/del polymorphisms which are associated with a statistically significant increase risk of BEEC (Wilkins et al., 2012). Screening of the *TAP63* promoter region for possible sequence variation did not reveal any significant variation which could contribute to the BEEC pathogenesis (Darling et al., 2013).

Chapter 4

Study of PERP and DSP

PART B: Suggested Declaration for Thesis Chapter
Monash University

Declaration for Thesis Chapter 4
Declaration by candidate

In the case of Chapter 4 (Based on the article: Mutation screening of *TP63* related target genes *PERP* and *DSP* in Bladder Exstrophy Epispadias Complex), the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Drafted and prepared the manuscript, Performed experiment, Analyzed data	85%

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
White, S.	Coordination and drafted the manuscript,	
Cheng, W.*	Conceived the idea, Coordination and drafting of the manuscript	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work*.

Candidate's Signature		**Date: 30/12/2014**

Main Supervisor's Signature		**Date: 30/12/2014**

*Note: Where the responsible author is not the candidate's main supervisor, the main supervisor should consult with the responsible author to agree on the respective contributions of the authors.

PART B: Suggested Declaration for Thesis Chapter

Monash University

Declaration for Thesis Chapter 4

Declaration by candidate

In the case of Chapter 4 (Based on the article: Identification of *Streptococcus parasanguinis* DNA contamination in human buccal DNA samples), the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Drafted and prepared the manuscript, Performed experiment, Analyzed data	85%

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
White, S.	Conceived the idea, Coordination and drafting of the manuscript	
Cheng, W.*	Coordination and drafting of the manuscript	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work*.

Candidate's Signature		Date: 30/12/2014

Main Supervisor's Signature		Date: 30/12/2014

*Note: Where the responsible author is not the candidate's main supervisor, the main supervisor should consult with the responsible author to agree on the respective contributions of the authors.

Chapter 4: *PERP* and *DSP* Study in Human BEEC

4.1 Introduction

Several studies have suggested involvement of a genetic component in BEEC pathogenesis (Boyadjiev et al., 2004; Caton et al., 2007; Nelson et al., 2005; Reutter et al., 2007a; Reutter et al., 2007b; Shapiro et al., 1984). The study of Cheng and co-workers first showed that the *TP63* (*P63*) gene is a candidate gene for BEEC, and is involved in the development of the ventral bladder (Cheng et al., 2006). In Chapter 3, expression studies and sequence variation analysis in the *TP63* gene (isoforms: *TAP63* and *ΔNP63*) promoters, revealed valuable information on BEEC pathogenesis. We have identified three in/del polymorphisms in the promoter of the predominant *TP63* isoform; *ΔNP63*, which were associated with increased risk of BEEC in humans (Wilkins et al., 2012).

A recent GWEP study identified 162 genes differentially expressed in both embryonic and postnatal human bladder samples, which are BEEC candidate genes. This study showed that *PERP* may have a possible role in BEEC (Qi et al., 2011). It also found that 30% of these differentially expressed genes are directly associated with desmosome (cell-cell adhesion complex) structure/function or cytoskeletal assembly. These finding points to desmosomal and/or cytoskeletal deregulation as an etiological factor for BEEC (Qi et al., 2011).

Perp was originally identified as an apoptosis-associated target of *p53*, and also a direct downstream target of *p63* (Ihrie et al., 2005). *PERP* has three exons and it has been demonstrated that *Perp* displays epithelial-specific expression during embryogenesis. It may also have an effect on bladder development. Desmosomes are complexes of linking

and adhesion proteins which form intracellular bridges. They provide multicellular structures and strength, and are essential for tissues with mechanical stress like myocardium, bladder, gastrointestinal mucosa, and skin (Delva et al., 2009; Getsios et al., 2004; Holthofer et al., 2007). Desmosomal proteins come from three different families including i) the cadherins (desmogleins and desmocolins) (Garrod et al., 2002; Kljuic et al., 2003; Whittock and Bower, 2003), ii) the armadillo proteins (plakoglobin, p120^{CTN} related proteins and plakophilins) (Bonne et al., 2003; Chen et al., 2002; Hatzfeld, 1999; Hatzfeld et al., 2003) and iii) the plakins (desmoplakin) (Bornslaeger et al., 1996; Gallicano et al., 1998). Among the desmosomal proteins, Desmoplakin (*Dsp*) is ubiquitous (Borrmann et al., 2000; Mueller and Franke, 1983), and has an important role in early embryogenesis (Gallicano et al., 1998).

4.1.1 Cellular Adhesion is regulated by *Perp* and Desmosome

The role of *Perp* and desmosome within the *p63* developmental program is to maintain epithelial integrity. This is further confirmed by down-regulation of *Perp* expression in some patients with ankyloblepharon ectodermal dysplasia and cleft lip/plate (AEC) syndrome (Beaudry et al., 2009). *Perp* protein works as a stabilizer of desmosomal complexes within cell–cell junctions. It can be regulated and activated by various *p63* isoforms (Ihrie et al., 2005). Therefore, it is possible that *Perp* mutations may occur in some *TP63*-related human disorders. *Perp* is involved in cellular adhesion, which also requires desmosome for its function. Desmosomes are responsible for anchoring the cells to each other, and desmoplakin plays an obligatory role in regulating keratin intermediate Filaments (Ifs) association to cytoskeletal system (Huen et al., 2002). Recently it was shown by Ihrie et al. (2005) that *Perp* null mice die post-natally within 10 days of birth of dehydration and develop blisters in stratified epithelia, including the skin and oral mucosa, indicating lack of cellular adhesion in the stratified epithelia. This

phenotype is similar to that of *p63* null mice, which die of dehydration shortly after birth. Also wound healing in *Perp* deficient mice takes longer than in controls, suggesting that *Perp* is necessary in cell-cell adhesion during wound closure (Beaudry et al., 2010).

Fuchs and Raghavan (2002) found that the proper cell-cell adhesion in stratified epithelium requires adherence junctions, desmosomes, and tight junctions. By utilizing Madin Darby canine kidney (MDCK) epithelial cells, a well-known cell line for studying intracellular junctions of polarized epithelial cells, Yeaman et al. (2004) found that *Perp* co-localises with both the desmosomal protein desmoplakin and the adherence junction protein E-cadherin. Further electron microscopic studies demonstrated the localization of *Perp* almost exclusively with Desmosomes. Moreover, *Perp* null (Koch et al., 1997) and desmosomal component deficient mice (Green and Gaudry, 2000) showed a similar phenotype of blistering. Furthermore, desmosomes in desmoplakin-null embryos do not attach to IFs. The electron microscopy study on morphometric difference of *Perp*$^{-/-}$ mice skin showed that the structures of desmosomes were perturbed in these specimens, suggesting a key role of both *Perp* and desmosome in stratified epithelial development.

4.1.2 Role of *PERP* and *DSP* in Bladder Development

In a study of three postnatal BEEC bladders (exstrophic bladder, EB) and three normal bladder (NB) samples, Qi et al. (2011) found that *PERP* is over-expressed in EB samples compared to NB, and present in all embryonic mesenchyme (EM) surrounding the urogenital sinus. Over-expression of *PERP* may be due to the absence, or reduced expression, of Δ*NP63* in BEEC patients. The *Perp* promoter is bound not only by *p53* but also by *p63*, indicating that *Perp* is responsive to *p63* signalling (Flores et al., 2002; Ihrie et al., 2003). The expression of *Perp* depends on direct activation by *p63*. *Perp*$^{-/-}$

knock-out mice showed that it plays an essential role in cell-cell adhesion by enabling desmosome function. Intriguingly, *Perp* specifically localizes to the desmosomes, where it extends through the plasma membrane to link with neighbouring cells. It is thought to be required for the proper trafficking and assembly of one or more desmosomal constituents during development (Bektas and Rubenstein, 2009). This was confirmed with a biochemical solubility assay, showing that, in the absence of *Perp*, desmosomal complexes assemble improperly (Ihrie and Attardi, 2005).

Desmoplakin is an essential element of functional desmosomes. It is encoded by the *Dsp* gene, which has 24 exons. *Dsp* fasten IFs (Intermediate Filaments) to desmosomal plaques i.e. plakoglobin and the plakophilins (Bornslaeger et al., 1996; Borrmann et al., 2000). This interaction finally forms an inner dense plaque which tethers the cytoskeletal network to the adhesion complex (Delva et al., 2009; Garrod and Chidgey, 2008; Kowalczyk et al., 1994). Moreover, *Perp* localizes to the desmosome, and the N-terminus (head) of *Dsp* is required for localization. Vasioukhin et al. (2001) reported that the association between *Dsp* and Keratin IFs is essential for both developing epidermis and adult stratified epidermal tissue. Jonkman et al. (2005) showed that loss of *Dsp* tail (C-terminal) leads to epidermolysis bullosa, which is an associated anomaly of BEEC. This tail domain is thought to regulate the binding of *Dsp* to intermediate filaments. In a GWEP study, Qi et al. (2011) showed that the two most under-expressed genes in BEEC are Desmin (DES) and Desmulin (DMN), which encode muscle specific intermediate filament (IF) proteins that directly interact with the C-terminal domain of Desmoplakin (*Dsp*). *Dsp* was found to be the sixth most over-expressed gene in exstrophic bladder samples (Qi et al., 2011). This study suggests a possible role of desmosomal activity in BEEC.

It is evident that *Perp* is required for stratified epithelial function *in vivo* (Ihrie and Attardi, 2005; Ihrie et al., 2006; Ihrie et al., 2005; Marques et al., 2006) and it involves *Dsp* linking IFs to the cytoskeletal network (Green and Gaudry, 2000). As with *p63*, *Perp* is also highly expressed in stratified epithelia, suggesting a possible involvement of *Perp* in BEEC. Based on the interactions of *TP63*, *PERP* and *DSP* in early embryogenesis, we previously speculated that these genes work in a common P63>PERP>DSP pathway in BBEC. Therefore, we hypothesized that the disruption of this pathway may result in a loss of epithelial integrity during bladder development and eventually lead to BEEC (Mahfuz et al., 2013b). This current study screened these two (*PERP* and *DSP*) candidate genes in order to gain a better understanding of BEEC pathogenesis.

4.2. Materials and Methods

4.2.1 DNA Samples

Buccal swab DNA samples were collected from India, Bangladesh, China, Australia, Spain, Canada and USA. All the samples were obtained with informed consent from the parents or guardians and ethics approval from the respective institutions (Australian samples, Royal Hospital for Children, Melbourne, and Number # HREC28140A). DNA isolation from buccal swabs was carried out using the BuccalAMP™DNA Storage and Extraction kit obtained from Epicentre® as per manufacturer's protocol. The DNA quantity and quality of each sample was determined using a Nanodrop® ND-1000 spectrophotometer.

4.2.2 Primer Design

Exon specific primers were designed using Light Scanner Primer Design Software (Version 1.0.r.84) and Primer 3 web tool (available at http://bioinfo.ut.ee/primer3/). An *in silico* PCR (available at http://genome.ucsc.edu/cgi-bin/hgPcr?command=start) was performed to confirm the exon specificity for each designed primer pair. All primer pairs were optimized using a Labnet multi-gene PCR Machine (Model: Multigene Gradient, Catalogue number TC 9600-G-230v). The temperature gradients used for optimization is shown in Table 8.

Table 8: Temperature gradient for optimizing the primers of *PERP* and *DSP*.

Temperature gradients in PCR					
T1 = 60 °C	T2 = 60.7 °C	T3 = 61.3 °C	T4 = 62.9 °C	T5 = 64.7 °C	T6 = 66.4 °C
T7 = 67.2 °C	T8 = 68.7 °C	T9 = 70.1 °C	T10 = 71.3 °C	T11 = 71.6 °C	T12 = 72 °C

For *PERP*, three primer pairs were designed for each of the 3 exons (Figure 37). Primer pairs for *PERP* exon 1 was designed such that it covered the part of the promoter with most conserved sequence, covering up to 750 bp upstream of the transcription start site (Figure 38). In *DSP*, a total of 30 primer pairs were designed for 24 exons (Figure 37). Two *DSP* promoter primers were also designed, covering up to 940 bp up-stream of the transcript start site (Figure 39). *PERP* and *DSP* primers are listed in the Appendix Table III and IV.
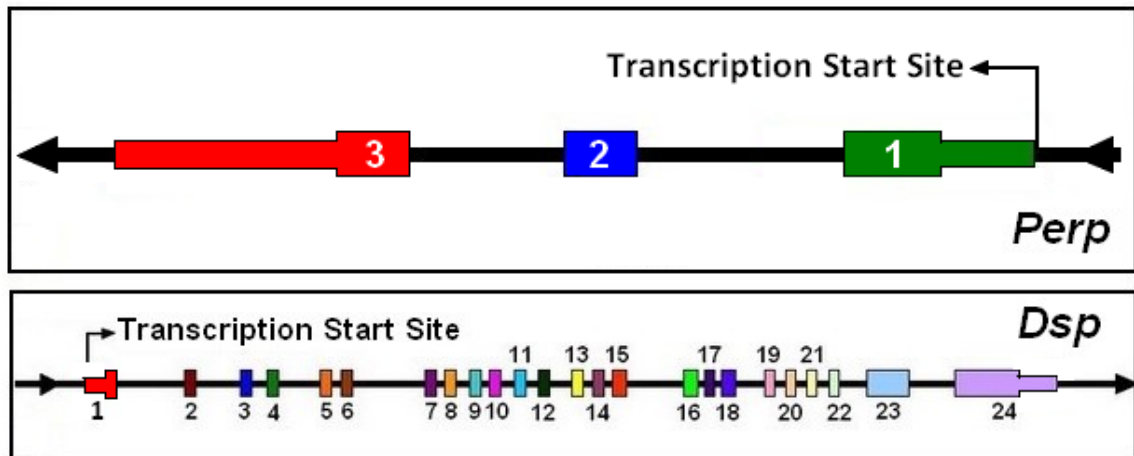
Figure 37: Schematic diagram of the *PERP* and *DSP* genes, showing different exons. The exons are indicated with numbers.
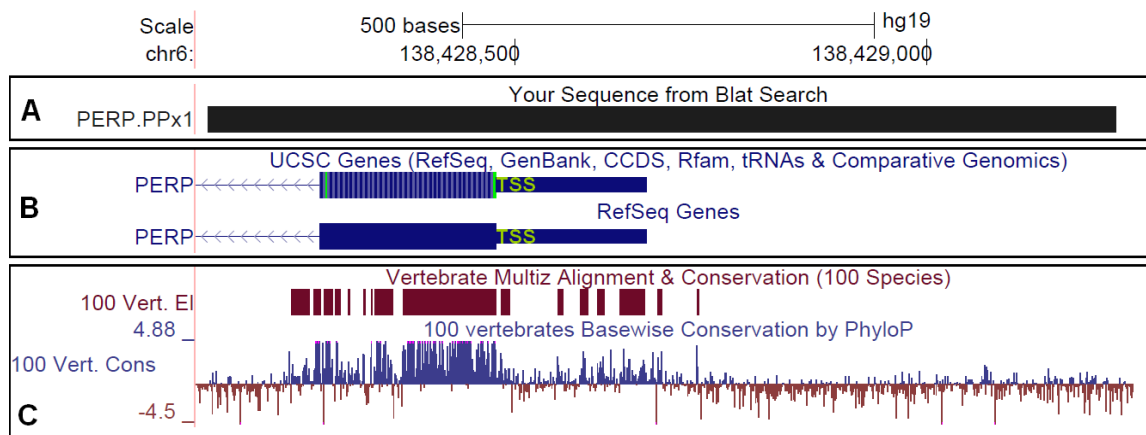
Figure 38: Primer design for screening of *PERP* promoter and exon 1. A. Amplified promoter and exon 1 sequence using PERP.PPx1 primer pair. B. Reference sequence from UCSC genome browser showing TSS. C. Vertebrate multiz alignment and basewise conservation found in 100 species (Kuhn et al., 2013).
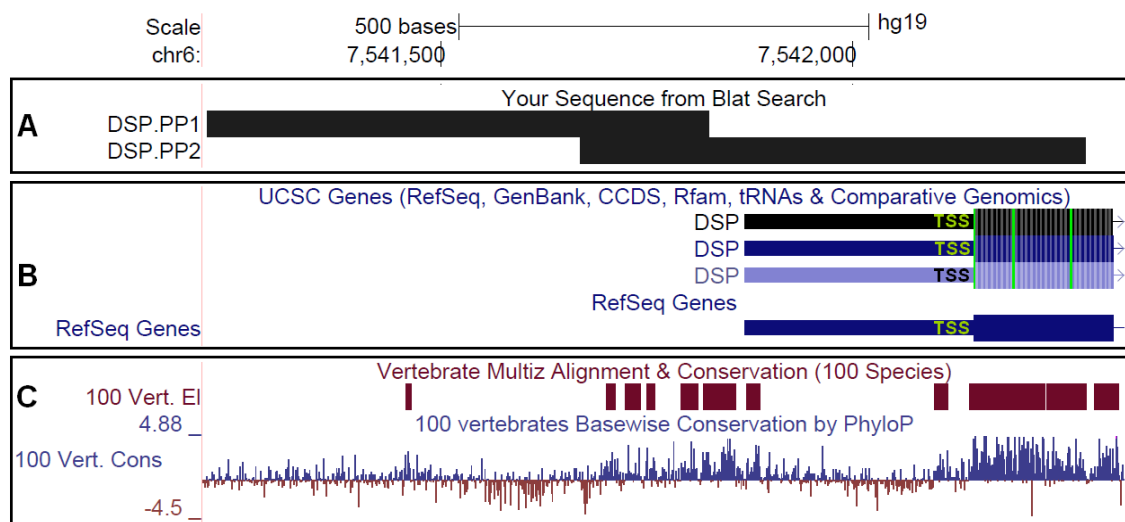
Figure 39: Primer design for screening of *DSP* promoter region. A. Amplified promoter sequence using DSP.PP1 and DSP.PP2 primer pairs. B. Reference sequence from UCSC genome browser showing TSS. C. Vertebrate multiz alignment and basewise conservation found in 100 species (Kuhn et al., 2013).

4.2.3 High Resolution Melting (HRM)

The majority of the *PERP* (exon 2) and *DSP* (exon 2-7, 9-17, 20 and 22) exons were screened using HRM analysis. HRM Curve prediction with designed primers was performed using uMelt software (Dwight et al., 2011). HRM was performed on DNA samples as previously described in chapter 2 using a LightScanner HR96 machine with a melting range of 70–96 °C. Curve analysis was performed using the LightScanner HRM analysis software (Version 2.0.0.1331). Normalized and difference curves were generated by plotting time ('x' axis) and fluorescence ('y' axis), and were used to identify potential sequence variants. A baseline was established from the most frequent curves on a difference plot upon which aberrant curves were determined and analysed for sequence variation. Gel Electrophoresis was conducted to check the amplification quality. The samples were then treated with ExoSAP-IT reagent following manufacturer's protocol (USB Corporation) to remove unused primers and enzyme, followed by Sanger sequencing.

4.2.4 Sanger Sequencing

PCR were performed with HotStart-IT Taq Master Mix (2X) in *PERP* exons 1 and 3, and *DSP* exons 1, 8, 18, 19, 21, 23 and 24, as a high GC ratio within the DNA sequence rendered them unsuitable for HRM analysis. The reaction condition was set according to HotStart-IT Taq Master Mix (2X) Protocol by USB corporation. The quality and quantity of amplified products were checked using gel electrophoresis.

Amplified samples were treated with ExoSAP-IT reagent following manufacturer's protocol (USB Corporation). Potential sequence variants were confirmed through Sanger sequencing on an Applied Biosystems 3130xl Genetic Analyzer at the Gandel Charitable Trust Sequencing Centre of Monash Health Translation Precinct, Melbourne, Australia. Sequence chromatograph trace files were viewed using the Sequence Scanner from Applied Biosystems. SNP identification was done visually, and sequences were aligned to the genome using NCBI-BLAST (Altschul et al., 1990) and UCSC-BLAT (Kent, 2002).

4.3. Results

4.3.1 Sequence Variation in *PERP* Exons and Promoter

PERP and *DSP* exons were screened for sequence variants in 22 BEEC buccal swab DNA samples. No novel sequence variants were identified in either gene that could contribute to BEEC pathogenesis. Two common SNPs (rs648396 and rs648802) in *PERP* exon 3 were identified in 13 patients (Figure 40).
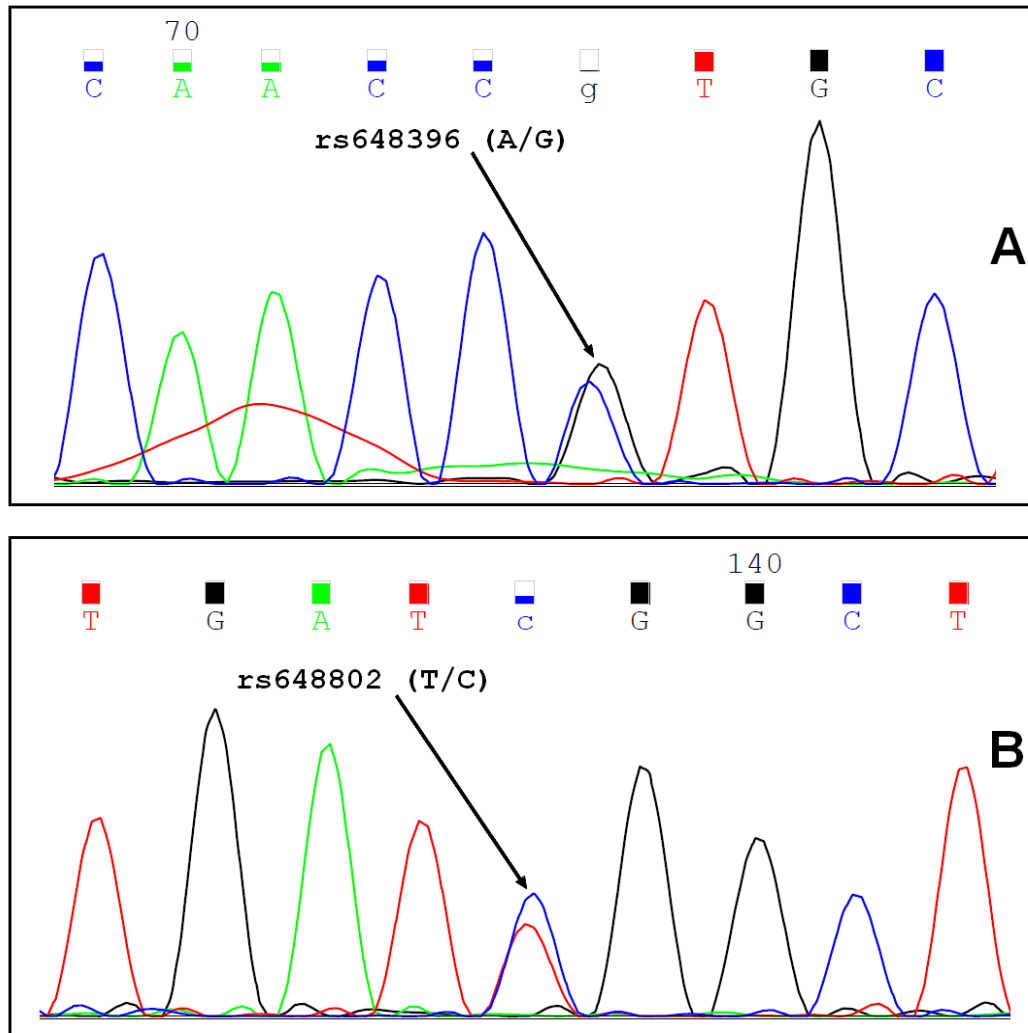
Figure 40: Sequence trace of *PERP* exon 3 showing two common synonymous variants. A. Common sequence variant rs648396 where A>G substitution occurred. B. Common sequence variant rs648802 where T>C substitution was found. The position of the variant in both A and B in '+' strand is indicated by an arrow.

The sequencing of the promoter identified one rare SNP (rs181052939, allele frequency<1%) in a single case. To check the presence of this SNP in larger patient cohort we used HRM. A primer pair was designed to cover this SNP and optimized. The *in silico* PCR result is shown in Figure 41. Upon HRM analysis, no new variants were identified.
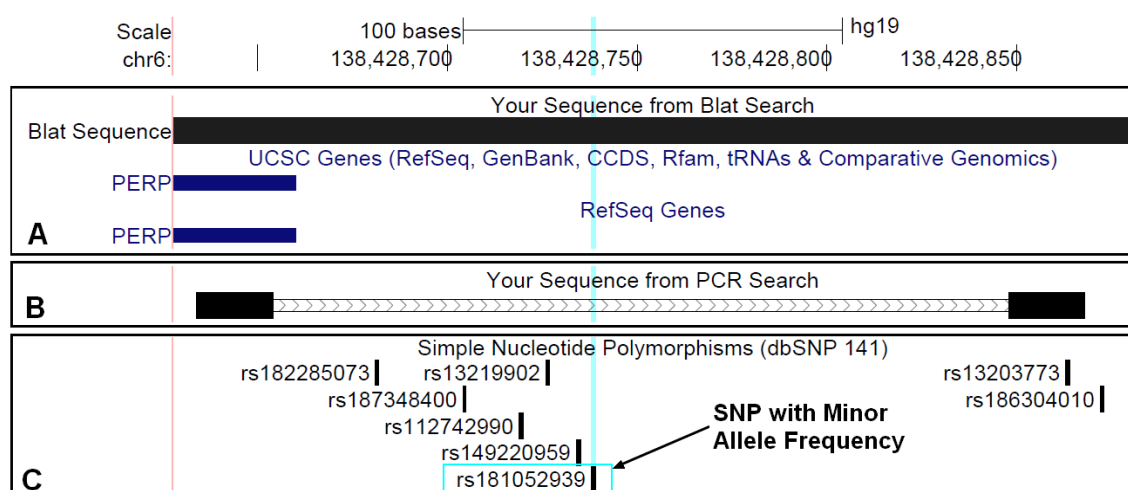
Figure 41: The region of the *PERP* promoter screened for sequence variation. A. BLAT sequence showing a part of *PERP* Promoter amplicon. B. *In silico* PCR of the PCR product rare variant in promoter region (allele frequency<1%). C. The rare variant in the *PERP* promoter region with minor allele frequency (MAF) according to dbSNP 141 and the highlighted vertical line shows the position of the variant in relation to the gene. MAF refers to the frequency at which the least common allele occurs in a given population. Image Source: UCSC Genome Browser (http://genome.ucsc.edu).

4.3.2 Bioinformatics Analysis of *PERP* Promoter

Bioinformatics analysis of *PERP* promoter region showed a number of transcription factor binding (TFB) sites of which are important for cell cycle regulation, transcription and cell proliferation and development (Figure 42). A list of important transcription factors and their functions are show in Table 9. DNase hypersensitivity site and layered chromatin marks such as H3K27ac and H3K4me3 are shown in Figure 43. The presence of the rare SNP (rs181052939) in a POL2 site may affect the transcription of the gene. More studies are needed to determine the exact nature of involvement of these TFBs in disease pathogenesis.

Table 9: Transcription factor on *PERP* promoter region regulating cell proliferation and differentiation (UCSC genome browser and GeneCard Database).

Transcription Factors	Functions	References	HGNC ID
CTCF (CCCTC-Binding Factor)	▪ Involved in transcriptional regulation by binding to chromatin insulators ▪ Prevents interaction between promoter and nearby enhancers and silencers. ▪ Plays a important role in chromatin remodelling and epigenetic regulation ▪ Plays an essential role in oocyte and preimplantation embryo development.	Filippova et al. (1996); Rubio et al. (2008)	HGNC:13723
EGR1	▪ Described in Chapter 3 in Table 6		HGNC:3238
EP300	▪ Described in Chapter 3 in Table 7		HGNC:3373
FOS	▪ Described in Chapter 3 in Table 6		HGNC:3796
FOSL1 (FOS-Like Antigen 1)	▪ Regulators of cell proliferation, differentiation, and transformation ▪ It can dimerize with proteins of the JUN family, thereby forming the transcription factor complex AP-1.	Matsui et al. (1990); Shaulian and Karin (2002)	HGNC:13718
FOSL2	▪ Same as FOSL1		HGNC:3798
FOXA1	▪ Described in Chapter 3 in Table 7		HGNC:5021
FOXA2 (Forkhead Box A2)	▪ Involve in embryonic development. ▪ Establishment of tissue-specific gene expression. ▪ Regulation of gene expression in differentiated tissues.	Kaestner et al. (1994); Mincheva et al. (1997)	HGNC:5022
GATA3	▪ Described in Chapter 3 in Table 6		HGNC:4172
HDAC2 (Histone Deacetylase 2)	▪ Histone deacetylases act via the formation of large multiprotein complexes, ▪ Cell growth arrest, differentiation and death, some developmental events	Betz et al. (1998); Randhawa et al. (1998)	HGNC:4853
HNF4A (Hepatocyte	▪ May play a role in development of the liver, kidney, and intestines.	Chartier et al. (1994); Battle	HGNC:5024

Transcription Factors	Functions	References	HGNC ID
Nuclear Factor 4, Alpha)	▪ Mutations in this gene have been associated with monogenic autosomal dominant non-insulin-dependent diabetes mellitus type I. ▪ Creation and maintenance of epithelial cell layers by regulating cell growth and adhesion between cells.	et al. (2006); Chandra et al. (2013)	
JUND	▪ Described in Chapter 3 in Table 6		HGNC:6206
MYC	▪ Described in Chapter 3 in Table 7		HGNC:7553
POLR2A	▪ Described in Chapter 3 in Table 6		HGNC:9187
TBP	▪ Described in Chapter 3 in Table 7		HGNC:11588
UBTF (Upstream Binding Transcription Factor, RNA Polymerase I)	▪ Plays a critical role in ribosomal RNA transcription as a key component of the pre-initiation complex, ▪ Mediates the recruitment of RNA polymerase I to rDNA promoter regions. ▪ It may also play important roles in chromatin remodelling and pre-rRNA processing	Matera et al. (1997)	HGNC:12511

Figure 42: Bioinformatics analysis of *PERP* promoter region. A. BLAT sequence showing the amplified region of *PERP* Promoter and Exon 1. B. Reference sequence showing 5′ UTR and exon 1 of *PERP* with transcription start site (TSS). C. Different transcription factor binding motifs found by ChIP-seq analysis of 161 factors from ENCODE project (ENCODE Project Consortium, 2012; Wang et al., 2013b). The green bar with an arrow on the transcription factor binding site denotes the consensus binding motif of the relevant transcription factor.

Figure 43: Bioinformatics analysis of *PERP* promoter region, showing DNaseI cluster and histone marks. A. BLAT sequence showing the targeted promoter region analyzed. B. Reference sequence showing 5′ UTR and exon 1 of *PERP* with TSS. C. DNaseI hypersensitivity clusters from ENCODE (ENCODE Project Consortium, 2012; Wang et al., 2013b). The number on the left side of DNaseI hypersensitivity sites e.g. 34 means that it was found in 34 cell types out of 125 cell types. D. Enrichment of H3K27ac and H3K4me3 histones as found in the 7 cell lines (ENCODE Project Consortium, 2012). Each layered colour code denotes different cell line and their enrichment e.g. the pink colour denotes enrichment of respective histones in Normal Human Lung Fibroblast (NHLF) cell line. Image Source: UCSC Genome Browser (http://genome.ucsc.edu).

4.3.3 Sequence Variation in *DSP* Exons and Promoter

DSP exons were screened for sequence variants in 22 BEEC buccal swab DNA samples. HRM and Sanger sequencing of *DSP* exons 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 21 and 22 did not identify any sequence variants. However, common SNPs were found in exon 20 (rs2064217 in 2 patients), exon 23 (rs28763969 and rs6929069 both in 2 patients) and exon 24 (rs2076300 in 2 patient, rs2744380 in 5 patients and rs11558731 in 8 patients) (Table 10). Also a common insertion was found in *DSP* exon 1 (rs17133512 in 9 patients). Figure 44 and 45 are showing HRM

difference curves and sequence trace of the common synonymous variant (rs2064217) in *DSP* exon 20. Figure 46 and 47 are showing the sequence trace of common SNPs found in *DSP* exon 23 and 24, respectively. The analysis of *DSP* promoter did not reveal any sequence variants.

Table 10: SNPs identified during the screening of *DSP* exons.

Exons	Screening Technique	Variants	Comments
DSP Exon 1	Sanger	rs17133512	Common SNPs
DSP Exon 20	HRM	rs2064217	Common SNPs
DSP Exon 23	Sanger	rs28763969, rs6929069	Common SNPs
DSP Exon 24	Sanger	rs2076300, rs2744380, rs11558731	Common SNPs
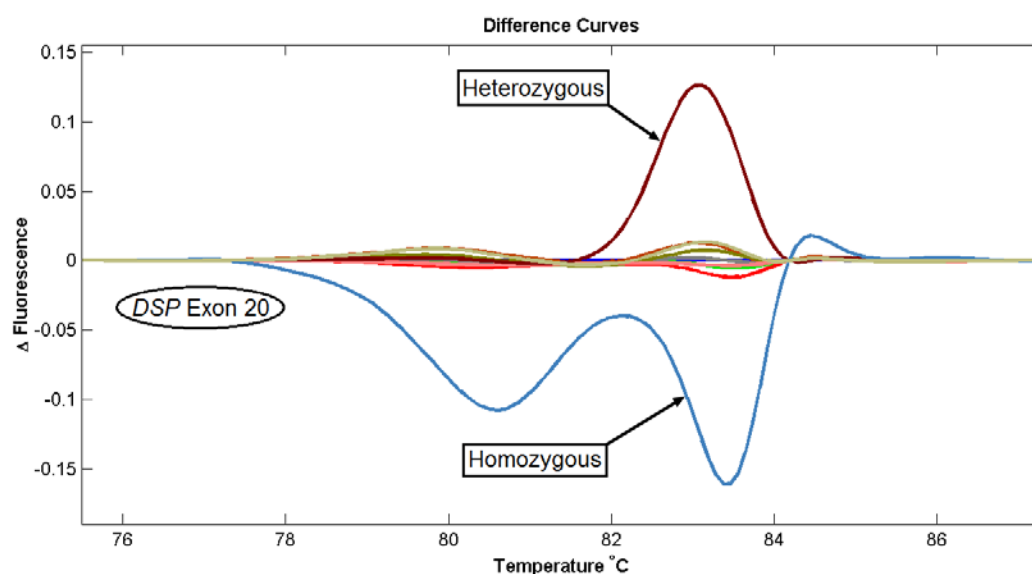
Figure 44: Difference curve showing a common synonymous variant (rs2064217) found in *DSP* exon 20. HRM curves in and/or close to Δ Fluorescence 0.0 (Normal Curves) is identical to reference sequence while away from 0.0 is sequence variant (Aberrant curves).
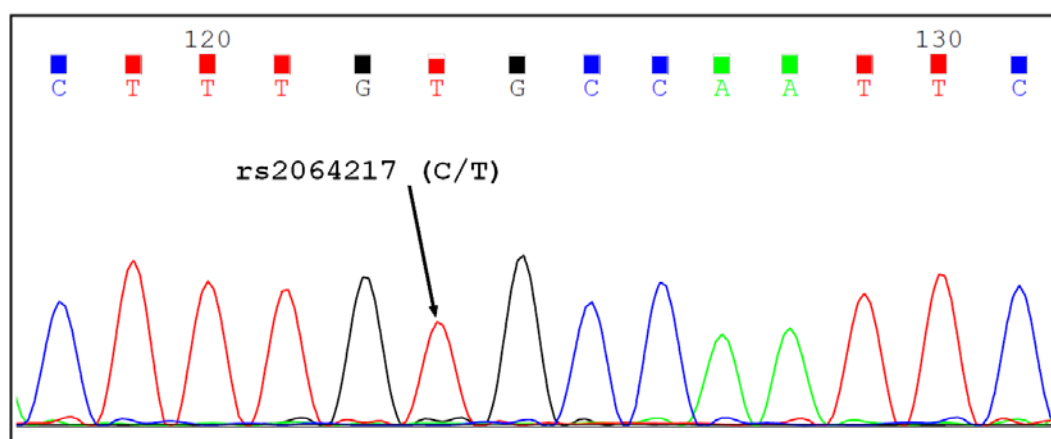
Figure 45: Sequence trace of *DSP* exon 20, with the common synonymous variant (rs2064217) indicated by an arrow. Homozygous variant where C>T substitution occurred.
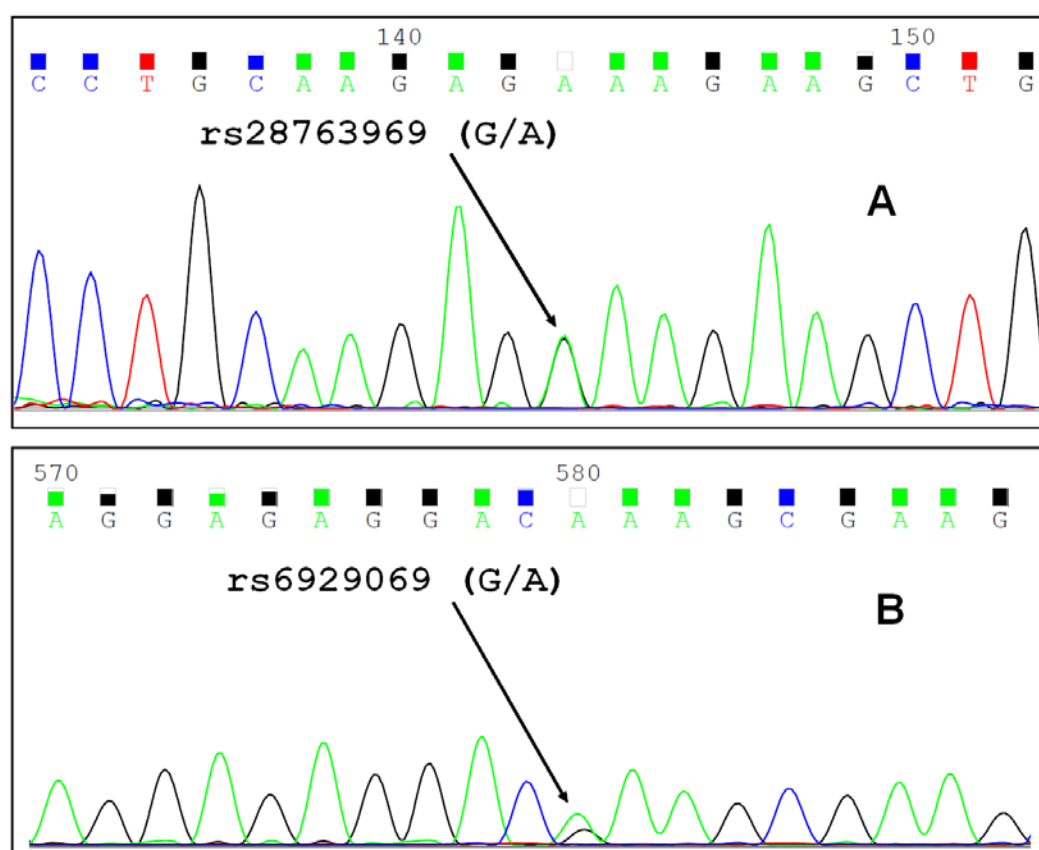
Figure 46: Sequence trace of *DSP* exon 23 showing two common synonymous variants (Heterozygous). A. Common sequence variant rs28763969 where G>A substitution occurred. B. Common sequence variant rs6929069 where G>A substitution was found. An arrow shows the position of the variant in both A and B in '+' strand.
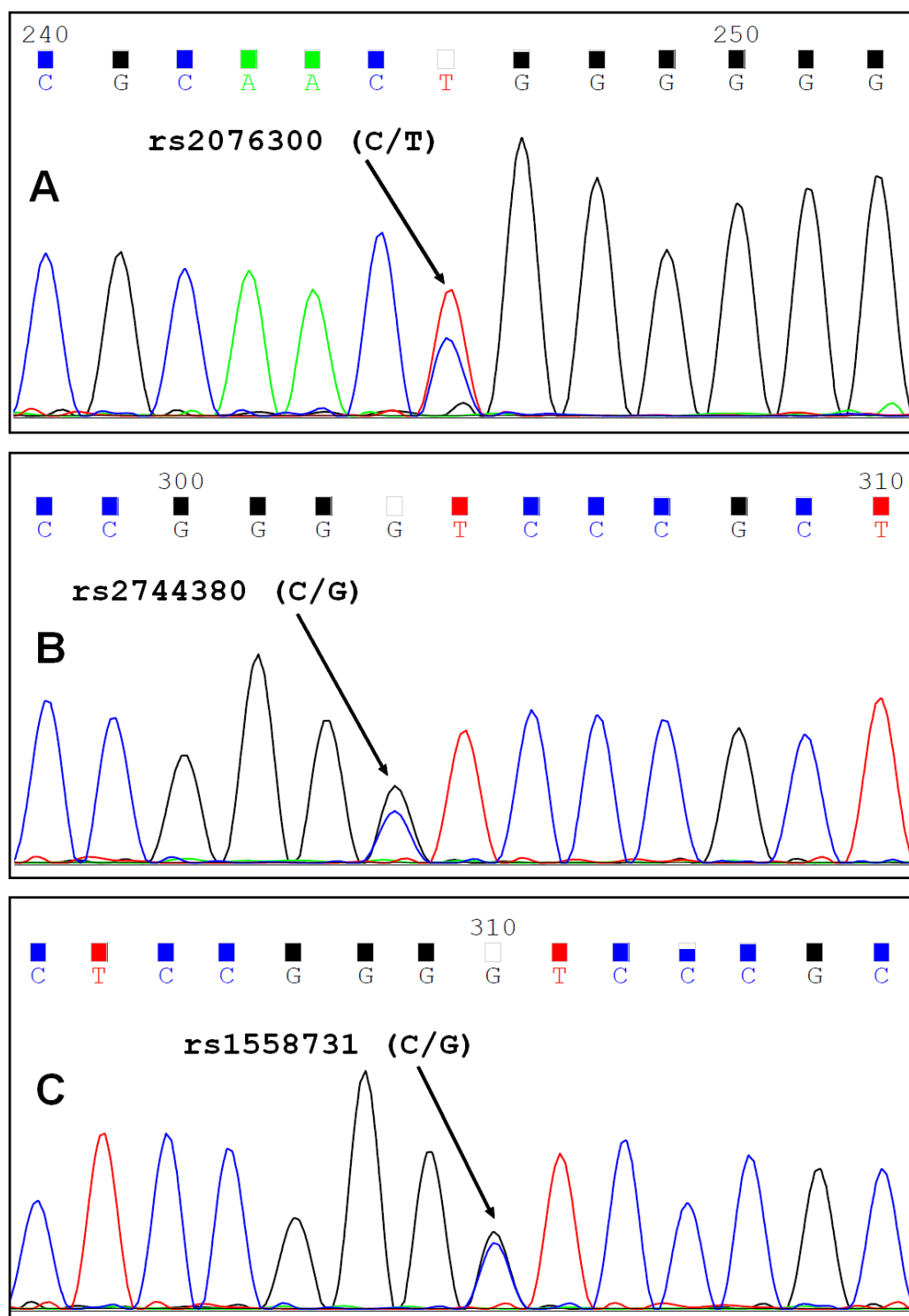
Figure 47: Sequence trace of *DSP* exon 24, showing common synonymous variants (Heterozygous). A. Common sequence variant rs2076300 where C>T substitution occurred. B. Common sequence variant rs2744380 where C>G substitution was found. C. Common sequence variant rs1558731 where C>G substitution occurred. An arrow shows the position of the variant in both A and B in '+' strand.

4.3.4 Bioinformatics Analysis of *DSP* Promoter

Bioinformatics analysis of the *DSP* promoter region identified a number of TFB sites (Figure 48). Some important transcription factors and their functions are described in Table 11. It is interesting to find that POLR2A, FOXA2, FOSL2, HNF4A, HDAC2 transcription factors are both present in *PERP* and *DSP* gene. These transcription factors are mainly involved in development, cellular adhesion and differentiation. DNase hypersensitivity site and layered chromatin marks such as H3K27ac and H3K4me3 were also found in the promoter (Figure 49). However, more research is needed to confirm the co-regulation and interaction of this TFB sites.

Table 11: Transcription factor on *DSP* promoter region regulating cell proliferation and differentiation (UCSC genome browser and GeneCard Database).

Transcription Factors	Functions	References	HGNC ID
CHD2 (Chromodomain Helicase DNA Binding Protein 2)	▪ Alter gene expression possibly by modification of chromatin structure thus altering access of the transcriptional apparatus to its chromosomal DNA template.	Woodage et al. (1997)	HGNC:1917
EP300	▪ Described in Chapter 3 in Table 7		HGNC:3373
FOSL2	▪ Same as FOSL1, describe in Table 9		HGNC:3798
FOXA1	▪ Described in Chapter 3 in Table 7		HGNC:5021
FOXA2	▪ Describe in Table 13 of this chapter		HGNC:5022
HDAC2	▪ Describe in Table 9		HGNC:4853
HNF4A	▪ Describe in Table 9		HGNC:5024
JUND	▪ Described in Chapter 3 in Table 6		HGNC:6206
MYBL2 (V-Myb Avian	▪ Transcription factor involved in the regulation of cell survival, proliferation,	Noben-Trauth et al.	HGNC:7548

Transcription Factors	Functions	References	HGNC ID
Myeloblastosis Viral Oncogene Homolog Like 2)	and differentiation.	(1996)	
MYC	▪ Described in Chapter 3 in Table 7		HGNC:7553
POLR2A	▪ Described in Chapter 3 in Table 6		HGNC:9187
TAF7 [RNA Polymerase II, TATA Box Binding Protein (TBP)-Associated Factor, 55kDa]	▪ Functions as a component of the DNA-binding general transcription factor complex TFIID ▪ Required for transcription by promoters targeted by RNA polymerase II. ▪ Plays a central role in mediating promoter responses to various activators and repressors	Chiang and Roeder (1995)	HGNC: 11541
TBP	▪ Described in Chapter 3 in Table 7		HGNC:11588
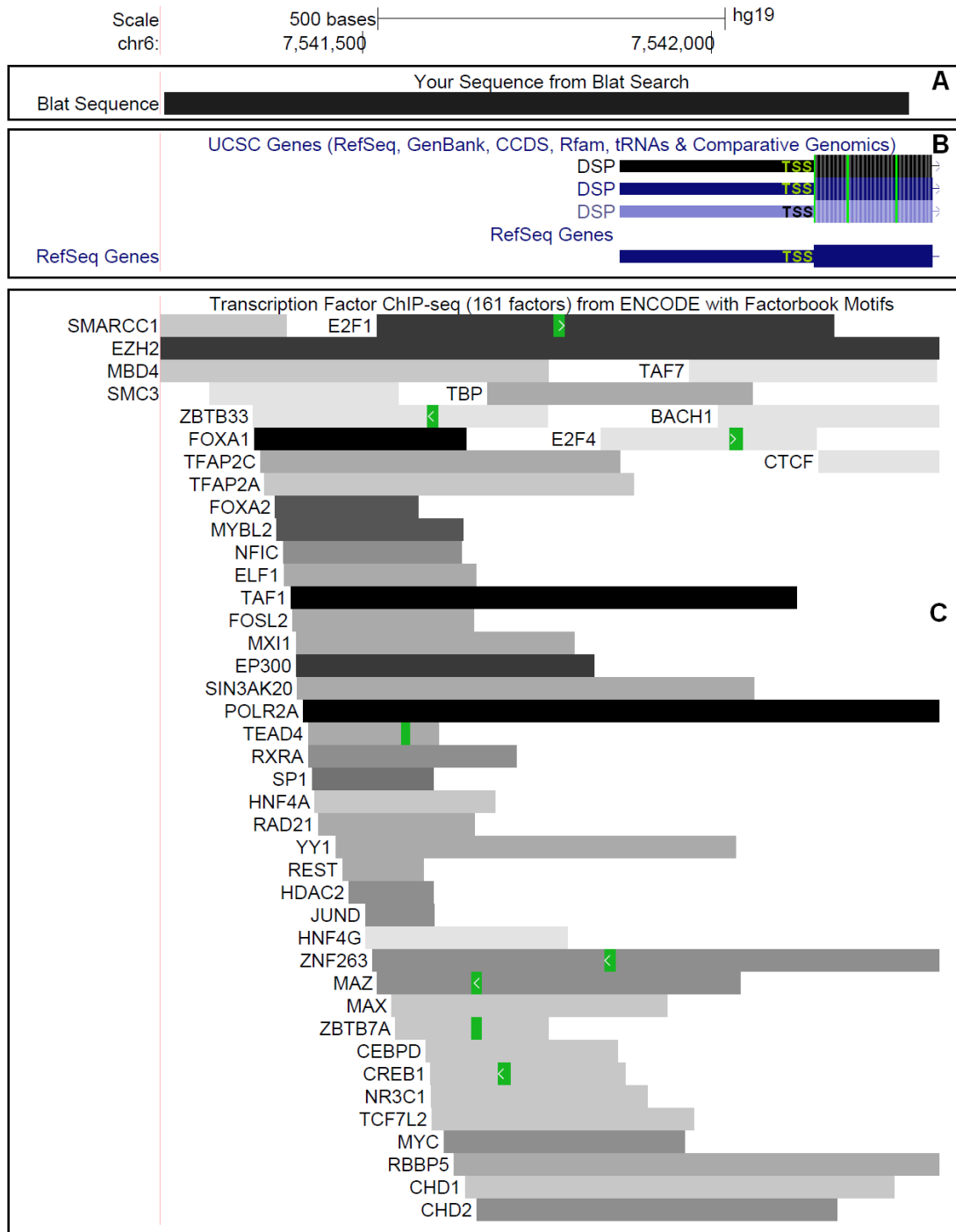
Figure 48: Bioinformatics analysis of *DSP* promoter region. A. Blat sequence showing the amplified region of *DSP* Promoter. B. Reference sequence showing 5′ UTR and exon 1 of *DSP* with transcription start site (TSS). C. Different transcription factor binding motifs as analysed by ChIP-seq of 161 factors from ENCODE project (ENCODE Project Consortium, 2012; Wang et al., 2013b). The green bar with an arrow on the transcription factor binding site denotes the consensus binding motif of the relevant transcription factor. Image Source: UCSC Genome Browser (http://genome.ucsc.edu).
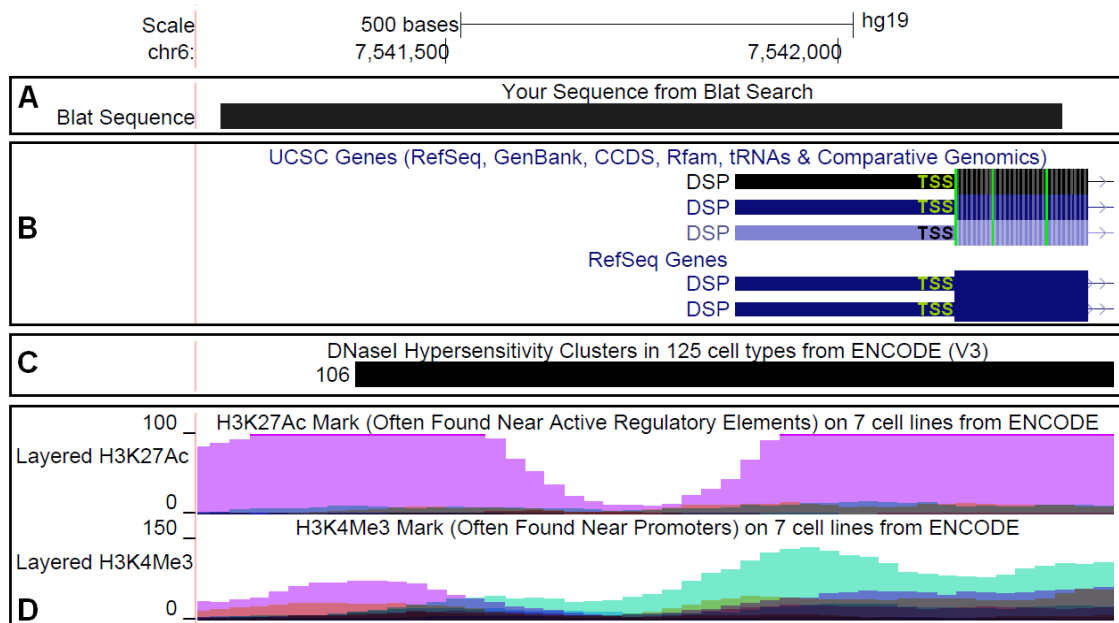
Figure 49: Bioinformatics analysis of the *DSP* promoter region, showing DNaseI cluster and histone marks. A. BLAT sequence showing the targeted promoter region analysed. B. Reference sequence showing the 5′ UTR and exon 1 of *DSP*. C. DNaseI hypersensitivity clusters from ENCODE (ENCODE Project Consortium, 2012). The number on the left side e.g. 106 means that it was found in 106 cell types out of 125 cell types. D. Enrichment of H3K27ac and H3K4me3 histones as found in the 7 cell lines from ENCODE (ENCODE Project Consortium, 2012; Gerstein et al., 2012). Each layered colour code denotes a different cell line and their enrichment e.g. the pink colour denotes enrichment of respective histones in Normal Human Lung Fibroblast (NHLF) cell line. Image Source: UCSC Genome Browser (http://genome.ucsc.edu).

4.3.5 Contamination of Buccal Derived DNA with Bacterial DNA

During *DSP* screening, a significant *Streptococcus parasanguinis* bacterial contamination in buccal DNA was found in some BEEC samples. This contamination was detected at the time of mutation screening *DSP* exon 4 for potential sequence variants using High Resolution Melting (HRM) and Sanger Sequencing. For this, we designed exon-specific forward (5' CTGTTTTCCTGCAGTGGTT 3') and reverse (5' TGGCCTGCACAGGTTTG 3') primers, predicted to generate a 254 bp product. HRM

was performed on 22 samples as previously described (de Boer et al., 2012). Five samples gave aberrant curves with HRM (Figure 50), and agarose gel analysis showed two bands (Figure 51). Sanger sequence analysis showed that the 153 bp fragment did not align with any human sequence. Alignment with other organisms using BLAST (Altschul et al., 1990) showed 98% identity (Sequence ID: gb[EU685104.0]) with *Streptococcus parasanguinis* plasmid pFW213 (Figure 52).
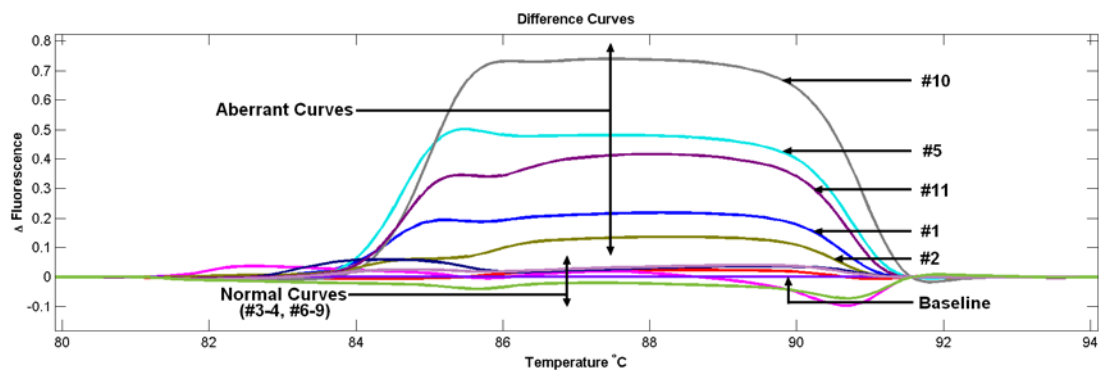
Figure 50: High Resolution Melt (HRM) analysis of PCR products from human buccal-derived DNA samples. The aberrant HRM curves (are due to contamination with *Streptococcus parasanguinis* DNA.
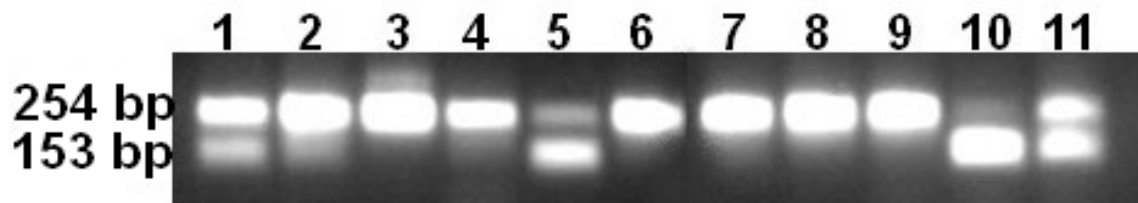
Figure 51: Gel picture showing contaminated (153 bp) and non-contaminated (254 bp) samples.

Figure 52: Alignment of the 153 bp fragment with reference sequence of *Streptococcus parasanguinis* plasmid pFW213, showing the preferential binding of human primers (Forward and Reverse) sequence (in bold).

Based on the amplified bacterial sequence with human primers, a bacterial-specific primer pair was designed (Table 12). These primers were then used to check the bacterial contamination in more buccal-derived DNA samples (n=30). Surprisingly, it has been found that almost 65% of the buccal DNA samples were contaminated with *Streptococcus parasanguinis* plasmid pFW213 (Figure 53). In this experiment (Figure 53A), 7 samples were checked for contamination that previously did not produce any HRM curve. The gel picture showed that 4 out of 7 samples were contaminated. Another 8 samples were also checked that were previously found to produce HRM curves. With a surprise, 7 out of 8 samples were found contaminated. However, it is not possible to reach any conclusion from this finding due to the smaller sample size,

whereas large sample size may reveal significance on contamination against downstream analysis.

Table 12: Primers used to amplify bacterial DNA in the buccal samples.

	Forward	Reverse
Human Primer	5' CTGTTTTCCTGCAGTGGTT 3'	5' TGGCCTGCACAGGTTTG 3'
Bacterial Primer	5' GTGGTTGGCATCCGTCTATC 3'	5' TGGCCTGCAGGTTTATTTAAG 3'
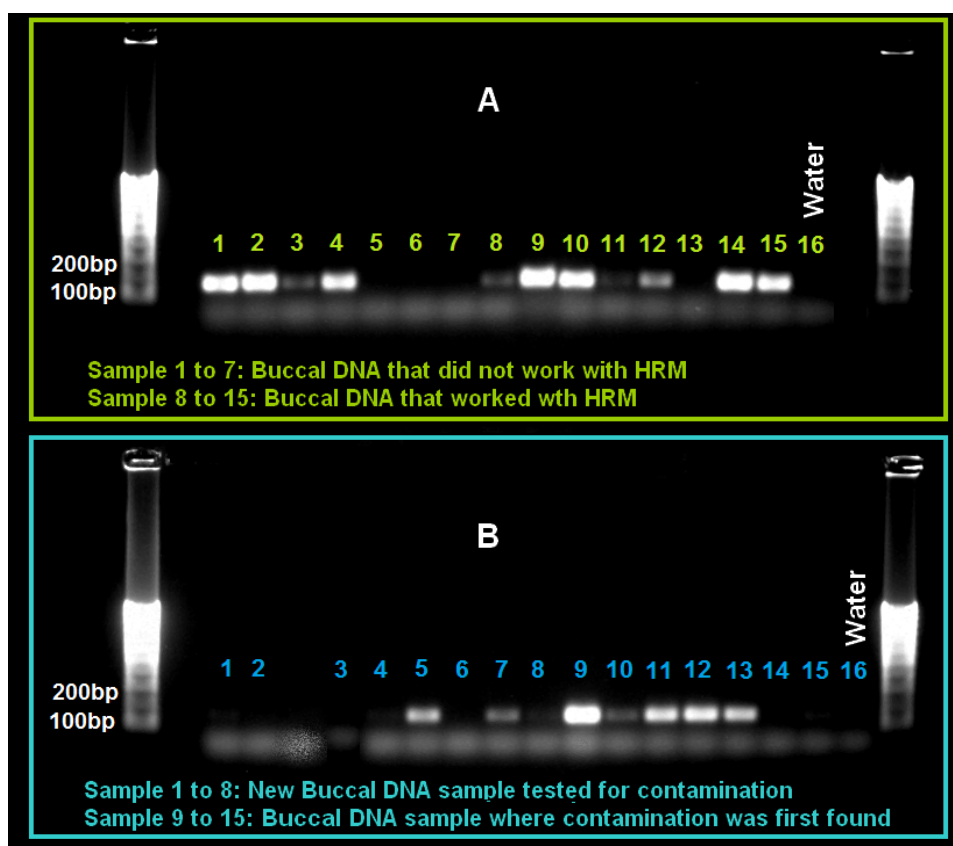
Figure 53: Bacterial contamination checked with a bacterial-specific primer pair in buccal derived DNA sample. A. Showing a comparison between DNA sample that did (Sample 1-7, n=7) and did not (Sample 8-15, n=8) produce melting curve with HRM experiment. B. Contamination of buccal DNA with bacterial DNA in new samples (Sample 1-8, n=8). Contamination confirmed with designed bacterial primer (9-15, n=7; These are the samples where contamination was first found during HRM of *DSP* exon 4 due preferential amplification of bacterial DNA with primer designed for human sequence)

4.4. Discussion

4.4.1 PERP is a Promising Candidate Gene for BEEC Pathogenesis

PERP has a critical role in epithelial stratification and also in cell-cell adhesion during development (Ihrie and Attardi, 2005; Ihrie et al., 2005). A previous report showed that *Perp* is a direct downstream target of the BEEC candidate gene *Tp63*(Ihrie and Attardi, 2005). In this study, we did not find any potential sequence variants in the coding region of *PERP*, but it is possible that intron 1 of this gene may have some valuable information on the disease phenotype. Ihrie and colleagues showed that *TAp63* and *ΔNp63* both transactivate the *Perp* reporter construct to a similar level, and this function is mediated largely through the p53/p63 consensus element in intron 1. The intron 1 is the major p53-responsive site in *Perp* (Ihrie et al., 2005) (Figure 54).
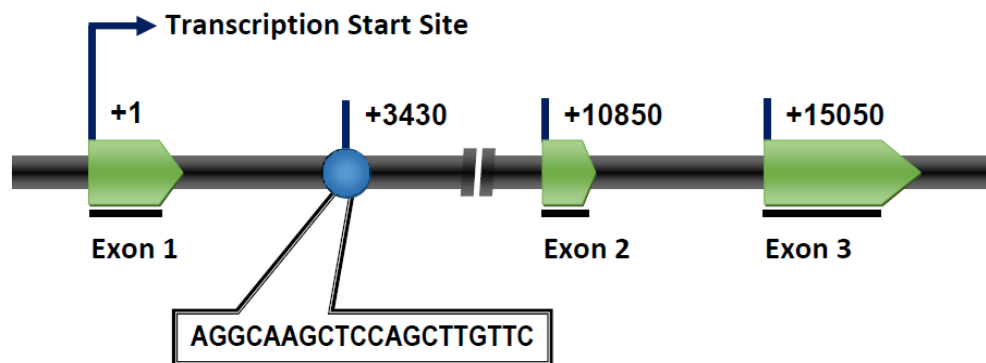
Figure 54: The human *PERP* genomic locus. Pointed green boxes are exons. Blue circle matches to the P53/P63 binding site consensus in *PERP*. Positions are labelled with respect to the transcriptional start site.

We have designed a primer pair which flanks part of the *PERP* promoter and the complete exon 1 of *PERP*, an 1102 bp product. Although we did not find any sequence variants in this part of *PERP* promoter, a complete analysis of the entire gene promoter and introns may shed some light on the pathogenesis of BEEC. Being a *TP63* target gene, it is possible that the *PERP* function may be impaired in BEEC patients. As described by Beaudry and colleagues, *PERP* induction is compromised in some patients

with Ankyloblepharon Ectodermal Dysplasia and Cleft Lip/Palate (AEC) (Beaudry et al., 2009; Ihrie et al., 2006), which is also an associated anomaly of BEEC (Celli et al., 1999; Rinne et al., 2006). Interestingly, AEC is caused by mutations in the gene encoding the *TP63* transcription factor, specifically in the Sterile Alpha Motif (SAM) domain. Previously, we have shown that there is an increased risk of BEEC associated with *ΔNP63* promoter in/del polymorphisms (Wilkins et al., 2012), therefore, it is possible that the *PERP* is dysregulated in BEEC patients. This hypothesis is corroborated by the findings of Qi and colleagues, where they found that *PERP* is over-expressed in the bladder tissue of BEEC cases (Qi et al., 2011) compared to unaffected bladder.

Bioinformatics analysis of *PERP* promoter showed a number of important transcription factor binding sites that may play a vital role in development, cellular adhesion and differentiation (Table 9). It has been reviewed in several article that in mammalian cells inter-chromosomal DNA interactions may mediate the decision which allele to activate and which to silence (de Laat and Grosveld, 2007; Simonis et al., 2006; Spilianakis and Flavell, 2006; Williams et al., 2010). Such inter-chromosomal interaction can help to regulate the transcription of a gene by remote enhancer or insulator regions through chromosome looping. Ling et al. (2006) in their study showed that 1 allele of the Igf2/H19 imprinted-control region (ICR) on mouse chromosome 7 colocalized with 1 allele of Wsb1/Nf1 on mouse chromosome 11. Omission of *Ctcf* or deletion of the maternal ICR abrogated this association and altered Wsb1/Nf1 gene expression (Ling et al., 2006). They concluded CTCF can mediate an inter-chromosomal interaction, by directing distant DNA segments to common transcription machinery. From the study of Ihrie and colleagues it is evident that transactivation of *Perp* is mediated by *p63* transcription machinery to ensure development (Ihrie and Attardi, 2005; Ihrie et al.,

2006; Ihrie et al., 2005; Ihrie et al., 2003). Therefore, the presence of CTCF in the *PERP* promoter region increased the possibility that the mediation of *P63* and *PERP* trans-chromosomal interaction is regulated by CTCF to aid the P63>PERP>Desmosome pathway in BEEC (Mahfuz et al., 2013b).

One of the major driving forces behind ribosome biogenesis is ribosomal RNA gene transcription by RNA polymerase I (PolI). The ribosomal biogenesis is the crucial to cell growth and proliferation (Nazar, 2004; Thomson et al., 2013). The key activator of PolI transcription is UBTF which is required for expression of the 18S, 5.8S, and 28S ribosomal RNAs. This transcription factor acts by facilitating recruitment of Pol I and essential basal factor SL1 (a complex of TBP and multiple TBP-associated factors or 'TAFs') (Bell et al., 1988; Cui et al., 2004; Panov et al., 2006). Study on the interaction of UBTF and ribosomal biogenesis in BEEC may reveal valuable information.

Epithelial formation is a central facet of organogenesis that relies on intercellular junction assembly to create functionally distinct apical and basal cell surfaces. It has been reported that *PERP* is involved in maintaining epithelial integrity via cellular adhesion, which also requires desmosome for its function. Studies conducted by Parviz and co-workers showed that HNF4A transcription factor is required for the epithelial transformation of the liver during development (Parviz et al., 2002; Parviz et al., 2003). Battle and colleagues showed in their study that HNF4A regulates developmental expression of a number of proteins required for cell junction assembly and adhesion (Battle et al., 2006). Very recently Bonzo et. al. (2012) identified *Perp* as a direct HNF4A target gene (Bonzo et al., 2012). HNF4A also found to bind with the most ubiquitous desmosomal gene, Desmoplakin (*DSP)* (Table 11, Figure 48).

4.4.2 Possible Involvement of other Desmosomal Proteins and Transcription Factor in BEEC

The inheritance of BEEC suggests that it is multifactorial, and there are likely to be other factors involved. It is possible that other desmosomal proteins like Desmoglein and Desmocolin (found in extracellular core of desmosome), and also the Plakoglobin and Plakophilins (found in outer dense plaque of desmosome) may have some involvement on BEEC pathogenesis. These desmosomal components facilitate contact and adherence to neighbouring cells (Delva et al., 2009). A GWEP study of Qi and colleagues showed that Desmin (*DES*) and Desmulin (*DMN*) are the two most under-expressed genes in BEEC (Qi et al., 2011). The muscle-specific intermediate filament (IF) protein is encoded by these two genes. Not surprisingly, the IFs interact with the C-terminal domain of *DSP* and form a connection with desmosomal plaques, which finally lead to epithelial stratification via cell-cell adhesion (Borrmann et al., 2000; Gallicano et al., 1998). Therefore, disruption of interactions between *DSP* and IFs during embryogenesis might contribute to BEEC progression.

Bioinformatics analysis of *DSP* promoter region has identified some very important transcription factor binding site. Transcription of gene is heavily relied on a series of highly regulated steps i.e. assembly of a pre-initiation complex (PIC) at the promoter nucleated by TFIID, followed by initiation, elongation, and termination (Hahn, 2004; Orphanides et al., 1996; Woychik and Hampsey, 2002). Gegonne and colleagues showed that TAF7, a component of the TFIID complex, bound to TAF1 controls the first steps of transcription (Gegonne et al., 2006). It interacts with and regulates the enzymatic activities of transcription factors that regulate RNA polymerase II progression. Its diverse functions in transcription initiation are consistent with its essential role in cell proliferation (Gegonne et al., 2013; Gegonne et al., 2012). It is

possible that disrupted TAF7 binding site in *DSP* may affect the transcription of the gene which eventually results in dysregulated desmosome activity.

The transcription factor MYBL2 has been suggested to play an important role in proliferation and differentiation of several cell types e.g. leukemic, neuroblastoma, glioblastoma and osteosarcomacell (Bies et al., 1996; Lin et al., 1994; Raschella et al., 1995). Some researcher reported that, it is expressed in all proliferative cells examined (Sala and Watson, 1999; Sitzmann et al., 1996). Muller et. al. (2008) showed that MYBL2 is one of 39 critical transcription factors commonly expressed in several different types of pluripotent stem cells which is implicated in the development and maintenance of the stem cell phenotype (Muller et al., 2008). Papetti and colleagues found that MYBL2 is commonly regulated in several systems of colon cell maturation both *in vitro* and *in vivo*. They also reported that MYBL2 work as a link between proliferation and differentiation in maturing intestinal epithelial cells (Papetti and Augenlicht, 2011). Detail study on interaction between MYBL2 transcription factor and *DSP* gene may uncover valuable insight.

Perp and *Dsp* both genes are required for maintaining epithelial integrity via cellular adhesion (Ihrie and Attardi, 2005). These two genes work together for their function (Ihrie et al., 2005). The transcription factor HNF4A binds to the promoter region of both of these genes. According to Battle et. al. (2006), HNF4A regulates developmental expression of a number of proteins required for cell-cell adhesion and assembly (Battle et al., 2006). Interestingly, it has been found that *Perp* is a direct HNF4A target gene (Bonzo et al., 2012). Therefore, it very tempting to speculate that HNF4A transcription factor is involved in co-regulation of *PERP* and *DSP* gene and disruption of one and/or both binding site in the promoter region of BEEC patient affect the ventral epithelial formation leading to the defect.

4.4.3 Bacterial Contamination in Buccal Swab Samples may affect DNA Analysis

Buccal swabs are a popular, inexpensive and non-invasive method of collecting DNA samples. It is a convenient procedure for collecting DNA from geographically isolated populations for larger cohort studies (Freeman et al., 1997), and has the advantage of avoiding the stressful process of venepuncture. In using buccal swab DNA, sampling or processing considerations may be important in obtaining optimal results (Walker et al., 1999). If buccal swabs are not collected and/or handled properly, potential complications may occur during subsequent analysis. Problems that can affect the interpretation include contamination, degradation, and insufficient yield (Beckett et al., 2008; Walsh et al., 1992; Zayats et al., 2009). Our inability to successfully screen all available samples in each of our selectively targeted regions like *PERP* and *DSP* promoter and also the screening of SNP (rs181052939, with minor allele frequency (AF<1%) in *PERP* promoter may be the result of sub-optimal buccal swab DNA quantity and quality contaminated with *Streptococcus parasanguinis* plasmid pFW213, hampering HRM screening process which is typically dependent on high quality sample DNA. This bacterium is the most abundant in the mouth, and a primary colonizer of human tooth surfaces (Socransky et al., 1977; Tinanoff et al., 1976) that plays an important role in dental plaque formation (Carlsson et al., 1970; Jenkinson and Lamont, 1997). A Streptococcal contamination in buccal cells is not unexpected, because of its high abundance in oral micro flora (Rudney et al., 2005; The Human Microbiome Project Consortium, 2012). Our finding demonstrates that bacterial contamination may be a significant contributor to the total amount of DNA isolated from a buccal swab, and may explain why an apparently sufficient quantity of high quality DNA does not perform as expected in human-specific PCR amplifications (Mahfuz et al., 2013a).

4.4.4 Alternative Signalling Pathways in BEEC Pathogenesis

According to Qi et al. (Qi et al., 2011), additional genes such as *SYNPO2* and *WNT5A* might be involved in the pathogenesis. *SYNPO2* is an actin-associated protein involved in actin-based motility in renal podocytes, which have a high capacity for protein synthesis and post-translational modifications (Angers and Moon, 2009; Mundel et al., 1997). It is possible that disruption in *SYNPO2* gene may increase the risk of BEEC by interrupting protein synthesis.

In BEEC patients, *WNT5A* gene was found to be highly up-regulated (14-fold over expression). It has been reported that in cloacal mesoderm over expression of mouse *Gli2* induced the expression of Wnt5A. More interestingly, *Gli2* also up-regulated p63 expression in cloacal endoderm, suggesting a common up-stream pathway is involved in the regulation of both *p63* and *Wnt5A* (Liu et al., 2007).

A recent report shows that a 32kb intergenic region between *WNT3* and *WNT9B* harbour regulatory region for classic bladder exstrophy (CBE). They reported two most significant intergenic region flanked by- i) *CYLD* and *SALL1* on chromosome 16q12.1, and ii) *EEF1E1* and *SLC35B3* on chromosome 6p24.3 (Reutter et al., 2014). Expression analyses of *Wnt3* and *Wnt9b* genes in mice revealed that both genes are expressed in the genital region during stages relevant to the development of CBE in humans (Reutter et al., 2014). It has also been reported that a homozygous nonsense mutation in *WNT3* has been associated with urorectal malformations (Niemann et al., 2004), and loss of *Wnt9b* causes urogenital defects (Carroll et al., 2005). The *WNT3* promoter region contains binding motifs for several transcription factors (Nakamura et al., 2011). Their analysis identified 185 binding sites up stream (from -759 to -993) of which following transcription factor family are most important to growth and development.

Homeodomain transcription factors	Homeobox transcription factors
cAMP-responsive element binding proteins	Octamer binding protein
Fork head domain factors	Drosophila Chorion Factor 2
HOX - PBX complexes	CLOX and CLOX homology (CDP) factors
SOX/SRY- related HMG box factors	CCAAT binding factors
GATA binding factors	LEF1/TCF
Fork head domain factors	Lim homeodomain factors

Among those transcription factor families, some are differentially expressed in human new-born bladder exstrophy tissue, e.g. HOX - PBX complexes, GATA binding factors, CCAAT binding factors, Fork head domain factors. They also have a vital role in embryonic urorectal septation process (Kang et al., 1997; Qi et al., 2011). Furthermore, it has been shown to contain PBX protein regulate the expression of *Wnt9b-Wnt3*, which in turn regulates *p63* (Ferretti et al., 2011). The study of Liu and colleagues also suggests a functional relation between *Tp63* and *Wnt5A,* where both of them seem to be regulated by a common upstream pathway (Liu et al., 2007). Thus, it is tempting to speculate that both WNT pathway and TP63 pathway (Mahfuz et al., 2013b) are somehow interlinked in BEEC pathogenesis. However, more research is needed to prove such interactions.

4.5. Conclusion

Our study did not identify any sequence variation in *PERP* and *DSP* coding regions that may contribute to BEEC pathogenesis. However, other published evidence so far supports loss or reduced *ΔNP63* expression and promoter in/del polymorphism leading to increased risk BEEC pathogenesis. To ascertain a complete picture of *PERP* and *DSP*

involvement, more research is needed on the interaction and regulation of these genes. Although, no coding mutations were found in the *PERP* and *DSP* genes, we cannot exclude the involvement of these genes in BEEC.

Chapter 5

Study of Gene Regulation

Chapter 5: Gene Regulation Study

5.1 Introduction

We previously hypothesized the involvement of *TP63>PERP>*Desmosome pathway in Bladder Exstrophy Epispadias Complex (BEEC) (Mahfuz et al., 2013b). In chapter 3 and 4, we attempted to establish a connection between the genes involved in the pathway and BEEC by screening the exons and promoter regions of relevant genes (*TP63*, *PERP* and *DSP*). Studies on the promoters and exons of *TP63* isoforms, *ΔNP63* and *TAP63,* have revealed significant information on the disease pathogenesis. Although we and others did not identify any exonic mutations in *TP63* (Ching et al., 2010; Wilkins et al., 2012), the promoter study of the predominant isoform *ΔNP63* identified three in/del polymorphisms which are significantly associated with increased risk of BEEC (Wilkins et al., 2012). We did not identify any sequence variants in the promoter region of *TAP63* in BEEC that may contribute to the disease (Darling et al., 2013).

To further test our hypothesis, we have studied the promoter and coding exons of *PERP* and *DSP*. This did not identify any sequence variants that could be linked to BEEC. However, bioinformatics analysis of these genes revealed several important regulatory elements (discussed in chapter 3 and 4) in their promoters.

In the absence of sequence variants in the promoter and exons of *PERP* and *DSP*, we next focused this study on identifying DNA elements that may regulate gene expression in bladder development. Gene regulation in eukaryotes is tightly controlled by several different factors and elements. Barrett and colleagues reviewed that the untranslated regions (UTRs) i.e. 5′ and 3′ UTRs, and introns, are a major contributor in the gene

regulation (Barrett et al., 2012). There is also evidence of transcriptional regulation of gene expression throughout the whole genome by regulatory elements present in the intergenic regions (Birney et al., 2007; Carninci et al., 2005; Cheng et al., 2005). As described in Chapter 2, gene promoters are responsible for initiation of transcription by recruiting TFIID and RNA polymerase II to the Transcription Start Site (TSS). In mammals, promoters can be divided in two separate classes *i.e.* i) simple promoter, with conserved TATA-box and a single TSS, and ii) complex promoter, containing multiple TSS with variable CpG-rich sites (Carninci et al., 2006). The complex promoters are highly enriched in vertebrates, and they recruit major regulatory elements in the gene regulation machinery, such as enhancers, including upstream and downstream promoter elements (UPE and DPEs) that contain transcription factor binding sites, and may act independently or synergistically with the core promoter to facilitate transcription initiation (Juven-Gershon et al., 2008; Smale and Kadonaga, 2003; Strachan and Read, 1999). Apart from promoter and enhancers, other elements such as insulators, activators and repressors are involved in gene regulation (Juven-Gershon et al., 2008; Sandelin et al., 2007). Evidence showed that some rare elements such as the motif ten element (MTE), downstream core element (DCE), and the X-core promoter element 1 (XCPE1) are also involved in gene regulation (Juven-Gershon et al., 2008).

Genes with complex promoters use regulatory elements, like enhancers and silencers, to selectively allow varying levels of expression as required by the gene regulation machinery. Several studies have shown that gene activation by remote enhancers is associated with long-range interactions between regulatory elements by forming enhancer-promoter chromatin loops (Cook, 2003; Fraser, 2006; Kadauke and Blobel, 2009; Nolis et al., 2009; Vilar and Saiz, 2005). It is hypothesized that proteins bound to remote enhancers interact directly with proteins bound to promoters, with the

intervening DNA being ''looped out'' (Drissen et al., 2004; Jing et al., 2008; Liu and Garrard, 2005; Tolhuis et al., 2002). Ernst and co-workers have defined chromatin signatures associated with active enhancers and promoters, by studying histone modifications using chromatin immuno-precipitation followed by high throughput sequencing (ChIP-seq) in cultured cells (Ernst and Kellis, 2010; Ernst et al., 2011). By studying chromatin state and enhancers in mouse embryonic forelimb and hind limb using whole-transcriptome analysis, coupled with genome-wide profiling of H3K27ac and H3K27me3, Cotney et al. (2012) established that histone modification profiling can be used as a tool for developmental enhancer discovery (Cotney et al., 2012). One of the crucial tools used in studying histone marks are the antibodies that recognize site-specific histone modifications (Fuchs and Strahl, 2011).

There is evidence showing that distinct combinatorial patterns of histone modifications are responsible for specifying function (Rando, 2012; Spicuglia and Vanhille, 2012; Strahl and Allis, 2000). Different classes of elements are marked by distinct patterns of histone modifications and TF binding (Heintzman et al., 2007; Wang et al., 2012). Histones can function positively and negatively in the regulation of gene expression by influencing post-translational modifications on specific amino acid residues. The International Human Epigenome Consortium identified six classes of histone H3 modifications; i) trimethylated H3K4 (H3K4me3), ii) acetylated H3K27 (H3K27ac), iii) monomethylated H3K4 (H3K4me1), iv) trimethylated H3K36 (H3K36me3), v) trimethylated H3K9 (H3K9me3), and vi) trimethylated H3K27 (H3K27me3), are particularly responsible for epigenome profiling (Bae, 2013). To date, H3K4me3 is one of the most studied chromatin modifications. The actively transcribed protein-coding promoters in eukaryotes are marked by the enrichments of both H3K4me3 and

H3K27ac. On the other hand, active enhancers can be identified by enrichments of H3K4me1 as well as H3K27ac (Kimura, 2013; Rivera and Ren, 2013).

As described earlier, enrichments of these histone modifications can be analysed by chromatin immuno-precipitation (ChIP)-based methods coupled with high throughput sequencing like NGS. During ChIP experiments, an antibody is used to target a specific modification. The antibody specificity is vital for such experiments, as this affects the quality of the data (Fuchs and Strahl, 2011). In our study, we will be using the normal human bladder urothelial cell line, SV-HUC-1. This cell line was established and immortalized by transformation of normal ureter tissue with SV40. The cell line has been repeatedly tested for production of infectious SV40 using an African Green Monkey kidney cell plaque assay, and has always tested negative (American Type Culture Collection). Several articles have been published using this cell line to study gene regulation in human bladder cancer (Hirata et al., 2012; Huang et al., 2014; Liu et al., 2013; Ueno et al., 2012). The regulation of cell proliferation, apoptosis and induction of epithelial stratification has also been studied using normal SV-HUC-1 and carcinogenic SV-HUC-1 (Huang et al., 2014; Southgate et al., 1994).

In the absence of sequence variation in coding sequence of *TP63, PERP* and *DSP*, we theorized that the transcription of these genes might be dysregulated in BEEC. Therefore, the evaluation of histone modifications at specific gene loci would be useful for deciphering the epigenomic state of BEEC. In this chapter we used ChIP assay coupled with QPCR and NGS to map H3K4me3 and H3K27ac chromatin states in a human bladder urothelial cell line, SV-HUC-1, and in human bladder tissue from BEEC patients. We believe this study will give valuable insight into identifying relevant regulatory elements in BEEC.

5.2 Materials and Methods

5.2.1 Cell line and Tissue Sample

Human normal urothelial cell line, SV-HUC-1 (ATCC$^®$ CRL9520™) from American Type Culture Collection was use to study chromatin states. Two patient tissue samples, P1 and P2, were also used in the experiments. The sample P1 is from bladder mucosa and P2 is from ureter. Tissue samples were collected during reconstructive surgery of BEEC patients and were stored at -80°C. Ethics clearance was obtained and approved from The Royal Children's Hospital (RCH) Melbourne Research Ethics Committee for the collection and use of human samples.

5.2.2 Antibodies

To study the chromatin state in the cell line and human samples, H3K4me3 and H3K27ac antibodies from Diagenode were used. Also to study the *TP63* transcription factor binding in BEEC patient, p63α (H-129) and p63 (H-137) from Santa-Cruz Biotech was used. To study the predominant *TP63* isoform, *ΔNP63*, the RR-14 antibody was kindly provided by Assoc. Prof. Satrajit Sinha, Department of Biochemistry and Centre of Excellence in Bioinformatics and Life Sciences, State University of New York at Buffalo, Buffalo, USA.

5.2.3 Cell Culture

The SV-HUC-1 cells were cultured in a 75 cm^2 tissue culture flask with Delbeco's Modified Eagle Medium (DMEM) containing 10% Foetal Bovine Serum (FBS) and 1X Penicillin-Streptomycin (PEST) antibiotic solution. Cells were then incubated at 37°C under 5% CO_2 air atmosphere. All the cell culture steps were carried out under strict aseptic conditions. Subcultures of the cells were maintained upon reaching 85% confluence. To perform cell subculture the medium was removed from the culture flask

and cell monolayer was washed twice with 1% phosphate buffered saline (PBS). Dissociation of cells monolayer was done using 1X Tryple Express solution. Upon complete dissociation fresh medium was added, aspirate and dispense into new flasks.

5.2.4 Formaldehyde Cross-linking

Cell Line: Cells were harvested on reaching 85% confluence and counted to give a final concentration of 2×10^6 cells/ml. To perform DNA-Protein cross-linking, 27 μl of 37% formaldehyde was added to the cell culture solution and incubated at room temperature for 8 minutes. Fixation was stopped by adding 114 μl of 1.25 M glycine to the tube and incubating it at room temperature for 5 minutes. Cross-linked cell were then centrifuged at 500g for 10 minutes at 4°C. The supernatant was aspirated leaving approximately 30 μl of solution behind without disturbing the cell pellet. The cell pellet was stored at -80°C for later use.

Tissue Sample: Approximately 50 mg frozen patient tissue sample was immersed in 1 ml of ice-cold PBS on a sterile 10-cm culture dish on ice. The tissue was quickly minced into smallest possible pieces using two clean razor blades. Using a sterile transfer pipette, minced tissue was transferred into two 1.5-ml tubes on ice (~500 μl/tube). The dish containing minced tissues were rinsed with another 1 ml of ice-cold PBS and distributed evenly between the two collection tubes. The samples were kept on ice.

To homogenize, the minced tissue was pipette up and down at least 20 times using an 18G (first) and then 21G needles, attached to a sterile 1 ml syringe. The homogenized tissue (~2 ml) was then passed through a 100 μm cell strainer (Falcon) into a 50 ml tube to get single cell suspension and transferred to a 1.5 ml tube. The cell suspension was then centrifuged at 4000 rpm for 6 minutes at room temperature. During this time 20

mM NaBu-PBS was freshly prepared. NaBu is a histone deacetylase inhibitor and is used when investigating acetylated epitopes. Each cell pellet was re-suspended in 500 μl of PBS or 20 mM NaBu-PBS, depending on the assay requirement. Cross- linking of cell suspension was carried out as described above. Cross-linked cell were then centrifuged at 4000×g for 6 minutes at 4°C and the cell pellet was stored at -80°C for later use.

5.2.5 Chromatin Immuno-Precipitation (ChIP)

The basic steps of the ChIP assay was described in chapter 2, figure 21. ChIP was performed using Dynabeads® magnetic beads from life technologies. First, the Dynabeads were washed twice with ice cold RIPA buffer (10 mM Tris-HCl, pH 7.5; 1 mM EDTA; 0.5 mM EGTA; 1% Triton X-100; 0.1% SDS; 0.1% Na-deoxycholate and 140 mM NaCl). Beads were separated from the buffer solution using a magnetic rack. 10 μl of prewashed Dynabeads resuspended in 90 μl RIPA buffer, i.e. a total of 100 μl to magnetic bead solution are needed for each IP. 0.2-ml PCR tubes in eight-strip format (Axygen PCR-0208-C and PCR-02CP-C) were used for immuno-precipitation. Specific antibodies [H3K4me3, H3K27ac, p63α (H-129), p63 (H-137) or RR-14] and control antibodies (IgG) were added to magnetic bead solution for binding and incubated at 40 rpm on a rotating wheel for at least 2 hours at 4°C.

Cell lysis and chromatin shearing was done using ChIP lysis buffer (50 mM Tris-HCl, pH 8.0; 10 mM EDTA; 1% SDS). In order to, shear the chromatin (200-700 bp fragment), a complete lysis buffer was prepared by adding PMSF and complete protease inhibitors to Lysis buffer. For each sample, 200 μl complete lysis buffer (165.4 μl of Lysis buffer; 2 μl of 100 mM PMSF; 28.6 μl of 7x complete protease inhibitor cocktail; 4 μl of 1M NaBu) was freshly prepare each time. Buffers were kept at room temperature until use. 130 μl of Complete Lysis buffer was added to the cross linked cells to

resuspend and incubated for 5 minutes on ice. The lysed sample (equivalent up to 3 million cells) was then transfer to a Covaris microTUBE (AFA Fiber with Snap-Cap). Shearing of the chromatin was carried out using the Covaris S220 with 2% duty cycle, 105 watts of peak incident power and 200 cycles per burst for 10 minutes. Finally, complete RIPA buffer was prepared by adding PMSF and complete protease inhibitors to RIPA buffer. For each sample, 1 ml of complete RIPA buffer (827.1 μl of RIPA buffer; 10 μl of 100 mM PMSF; 142.9 μl of 7x complete protease inhibitor cocktail and 20 μl of 1M NaBu) made freshly and 870 μl of complete RIPA buffer was added to 130 μl of sheared chromatin to reduce the SDS concentration (to approximately 0.1%) before the addition of the antibody coated beads.

The antibody coated magnetic beads were spun briefly using a bench top touch spin apparatus to bring down liquid caught in the lid. Tubes were placed in an ice-cold magnetic rack to separate the coated beads. Supernatant was discarded and 100 μl of diluted sheared chromatin was added to each IP (specific antibody-coated beads) and negative control (IgG). IP and IgG tubes were then incubated under constant rotation on a rotator at 40 rpm for overnight at 4°C. For input controls, 10 μl of diluted, sheared chromatin sample was transferred to a clean 0.2-ml PCR tubes (in eight-strip format) and store at 4°C for later analysis.

On the next day, washing of IP samples and IgG negative controls was carried out three times using 100 μl of ice-cold RIPA buffer and finally twice with TE buffer. Each time the tubes were incubated for 4 minutes at 4°C on constant rotation (40 rpm). For the first wash step with RIPA buffer, rotation was performed manually in the cold room to reduce the background. The last two wash steps were done on the rotating wheel. Wash step with TE buffer, manual rotation of tubes were done in the cold room; not on the rotating wheel, to reduce the background.

5.2.6 Reverse Cross-linking IP sample and DNA purification

Reverse cross-linking and elution of DNA was performed according to iPure Kit (Diagenode) protocol. The reaction mixture was incubated at least 4 hours (or overnight) at 65°C on a thermomixer, with continuous rotation at 1300 rpm. DNA binding was performed following the instruction of iPure Kit (Diagenode) protocol by adding 2 μl of carrier, 100 μl of 100% isopropanol and 15 μl of provided magnetic beads to each IP and Input sample. IP and input samples were incubate for 1 hour at room temperature on a rotating wheel (40 rpm). Following washing steps, two elutions in 25 μl (total volume 50 μl) was conducted to obtain the purified DNA. DNA quantity and quality was measured using Qubit® 2.0 Fluorometer (Invitrogen) following manufacturer's protocol.

5.2.7 Real Time PCR (QPCR) of Purified ChIP DNA

To determine ChIP efficiency, QPCR was performed on specific genomic locus related to our hypothesis. To determine QPCR efficiency, a standard curve was used which is generated from fragmented genomic DNA with desired ChIP primer. Primers were designed following the instruction of LowCell# Chip Kit Manual (Diagenode) and are listed in Appendix Table V.

Real-time qPCR was carried out using SYBR Green Master Mix (Applied Biosystem) on 7900HTk Fast Real-Time PCR (Applied Biosystem) machine. Most qPCR programs allow automatic calculation of the DNA quantity by comparing Ct and known quantities of DNA standards. The efficiency of chromatin immune-precipitation of particular genomic locus is calculated from qPCR data as a percentage of starting material. The % of input can be calculated by the following formula-

$$\% \text{ (ChIP/ Total input)} = 2^{[(Ct(x\%input) - \log(x\%)/\log 2) - Ct(ChIP)]} \times 100\%$$

Where, Ct(ChIP) and Ct(x%input) are threshold values obtained from exponential phase of qPCR for the IP'd DNA sample and input sample respectively, and (logx%/log2) accounts for the dilution 1:x of the input.

5.2.8 Chromatin Immuno-Precipitation Sequencing (ChIP-Seq)

The basic flow diagram of ChIP-seq was described in Chapter 2, Figure 24. At least 2 ng of ChIP DNA was used to prepare the library for sequencing. Libraries were sequenced on the Illumina HiSeq 1500 sequencer at the Monash Health Translation Precinct Medical Genomics Facility. Single-end sequencing was performed; to a length of 50bp. Sequencing files were delivered in FASTQ format. Sequence reads were filtered using the Trimmomatic Program (Lohse et al., 2012). The filtered reads were then aligned to the hg 19 version of the human reference genome, using Bowtie2 (Langmead and Salzberg, 2012) to create BAM files. Peak calling on BAM files was performed with HOMER (Heinz et al., 2010), using the histone option. HOMER was also used for generating bedgraph files for viewing in the UCSC Genome Browser (makeUCSCfile), functional annotation (annotatepeaks), and motif analysis (findMotifsGenome) (Heinz et al., 2010).

5.3 Results

5.3.1 Relative Enrichment of Targeted Loci in Urothelial Cell Line

To confirm the binding of H3K4me3 and H3K27ac on major TP63>PERP>Desmosome pathway-related genes (P63 pathway) (Mahfuz et al., 2013b), QPCR was performed using primers listed in Appendix Table V. Relative enrichment of the targeted genes of P63 pathway (*ΔNP63*, *PERP* and *DSP*) was measured. The data shows specific enrichment of H3K4me3 and H3K27ac on these targeted genes. In both cases, ChIP negative control is low *i.e.* IgG<0.05, indicating low background in the ChIP assay. In

case of H3K4me3, the targeted gene loci, *ΔNP63*, *PERP* and *DSP*, were approximately 5, 17 and 24 fold enriched respectively; compared to the negative enrichment region (FOS-) of H3K4me3 (Figure 55). On the other hand, H3K27ac was approximately 1, 2 and 3 fold enriched respectively on the same targeted loci; compared to the negative enrichment region (FOS-) of H3K27ac (Figure 56). Such, specific enrichment suggests that the DNA fragments immuno-precipitated from the genomic screen represent *bona fide* binding sites for H3K4me3 and H3K27ac on the cells. It also confirms that the ChIP assay performed on the cell line was successful.
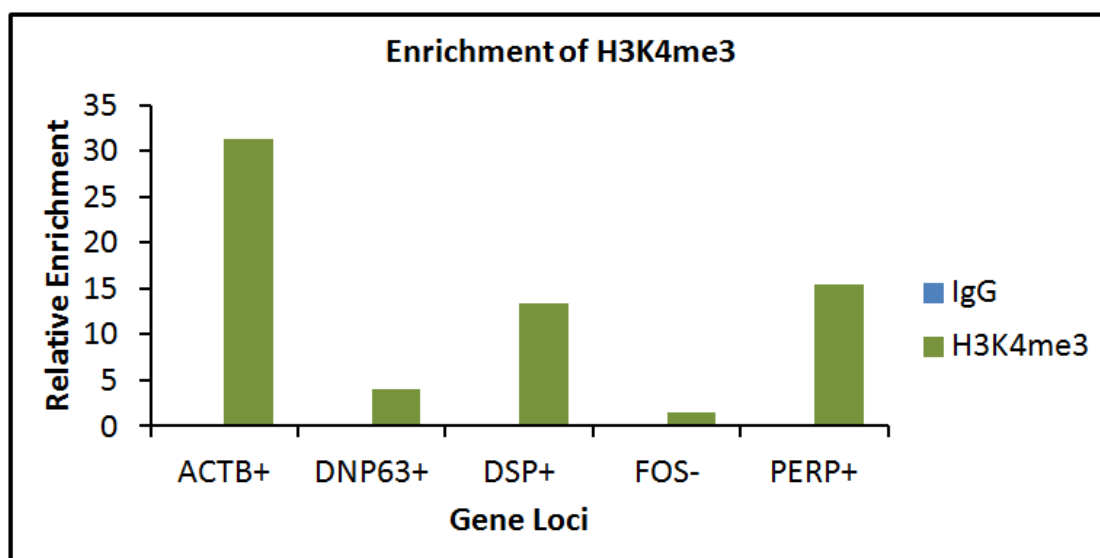
Figure 55: Relative enrichment of H3K4me3 binding sites in targeted gene loci related to P63 pathway in SV-HUC-1 cell line. ACTB+ locus was used as positive enrichment region and FOS- as negative for H3K4me3 binding sites. DNP63+ (ΔNP63+), DSP+ and PERP+ are targeted loci. The IgG was used as ChIP negative control to determine the background signal.
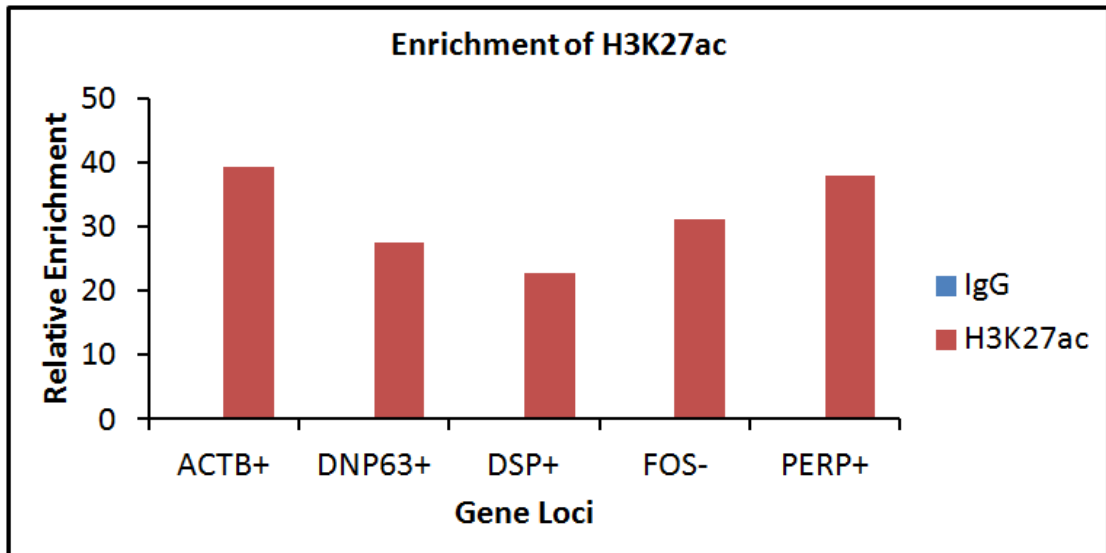
Figure 56: Relative enrichment of H3K27ac binding sites in targeted gene loci related to P63 pathway in SV-HUC-1 cell line. ACTB+ locus was used as positive enrichment region and FOS- as negative for H3K27ac binding sites. DNP63+ (ΔNP63+), DSP+ and PERP+ are targeted loci. The IgG was used as ChIP negative control to determine the background signal.

5.3.2 Genome-Wide H3K4me3 Binding Profile

ChIP-seq analysis using the H3K4me3 antibody gave a total of 36,890 peaks, of which 43% and 32% were found in the intronic and intergenic regions, respectively. Promoter, CpG-island and exon regions contain 10%, 8% and 4% of the total peaks, respectively. The remaining regions were TTS, 5'UTRs and other regions (each of them is 1% of the total) (Figure 57).
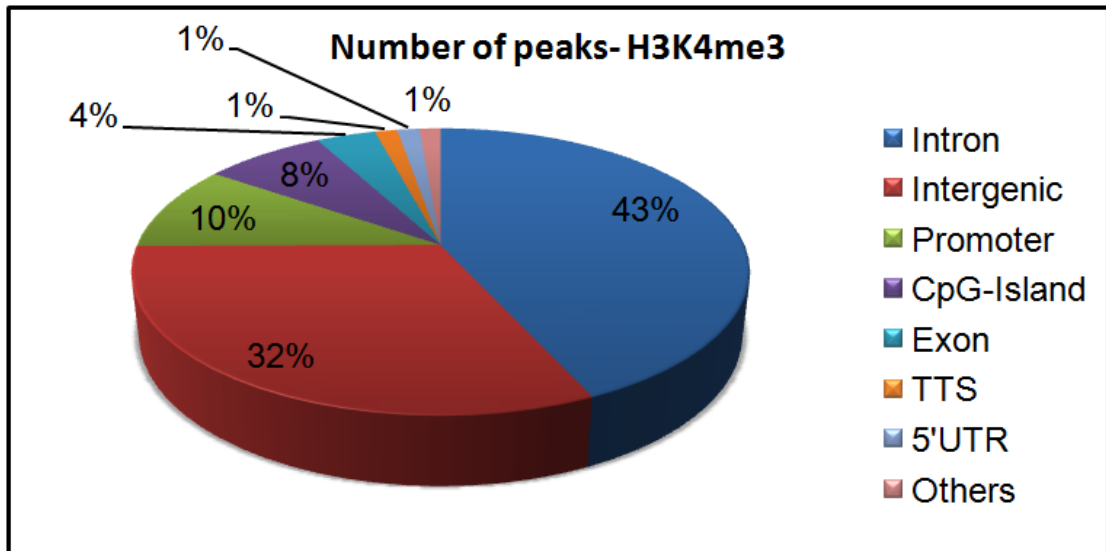
Figure 57: Distribution of the H3K4me3 binding site locations, relative to RefSeq genes. Locations of binding sites are divided into: Intron, Intergenic, Promoter, CpG-island, Exon, Transcription Termination Site (TTS), 5′ Untranslated Region (UTR) and others.

Log2 enrichment analysis of H3K4me3 binding site showed ~4.5 fold-enrichment in CpG-island and 5'UTR regions, whereas promoter and exon regions have shown ~3.5 and ~2.0 fold-enrichment, respectively. The fold-enrichment in TTS and intron region is <1, whereas no enrichment was found in the intergenic regions (Figure 58). Most of the H3K4me3 peaks were found in the introns and intergenic regions, but less or no enrichment of H3K4me3 was detected in these regions. This means that H3K4me3 is most active in the promoter and/or near the promoter (5'UTRs and CpG-islands) regions of genes.
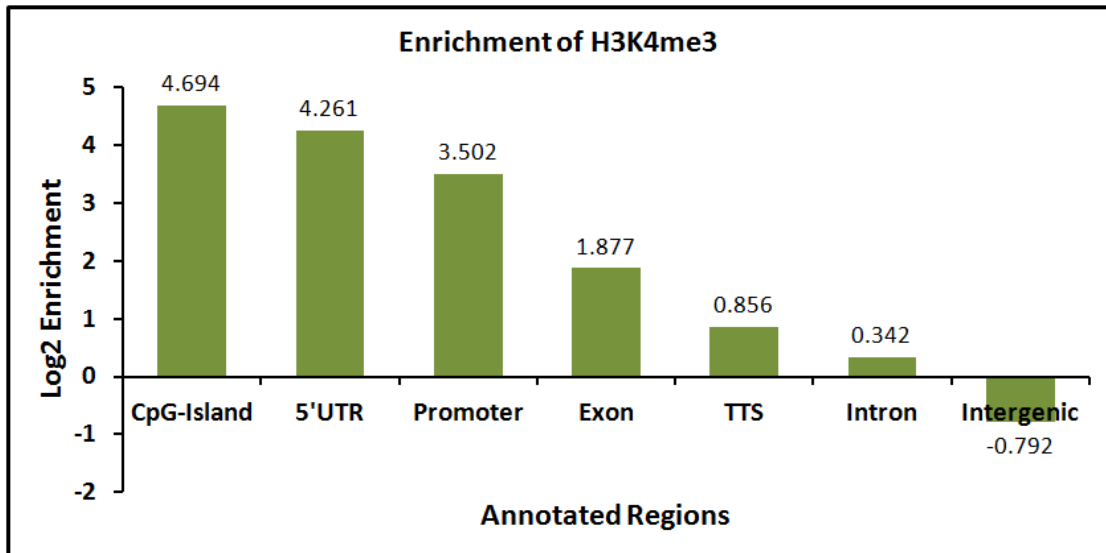
Figure 58: Enrichment of H3K4me3 annotated regions relative to RefSeq genes. Annotated regions were divided into: Intron, Intergenic, Promoter, CpG-island, Exon, Transcription Termination Site (TTS), 5′ Untranslated Region (UTR) and others.

5.3.3 Genome-Wide H3K27ac Binding Profile

ChIP-seq analysis using an H3K27ac antibody gave a total of 43,983 peaks, of which 50% and 38% were found in the intron and intergenic regions, respectively. 5% of the peaks were found in the promoter and a total of 4% peaks were found in the exon and TTS regions. The rest of the peaks were found in 3'UTRs, CpG-Islands and other regions (Figure 59).

Log2 enrichment of H3K27ac binding sited showed ~2.5 and ~2.0 fold-enrichment of the promoter and CpG-island, respectively. Approximately 1.0 fold-enrichment was found in TTS and exon regions. Fold-enrichment of 3'UTRs and intron is <1 whereas no enrichment was detected in the intergenic regions (Figure 60). Compared to the H3K4me3 enrichment in TTS, 3'UTRs and introns, the fold-enrichment of H3K27ac is higher in these sites, meaning that H3K27ac is active around promoter region as well as other regions, possibly near regulatory elements.
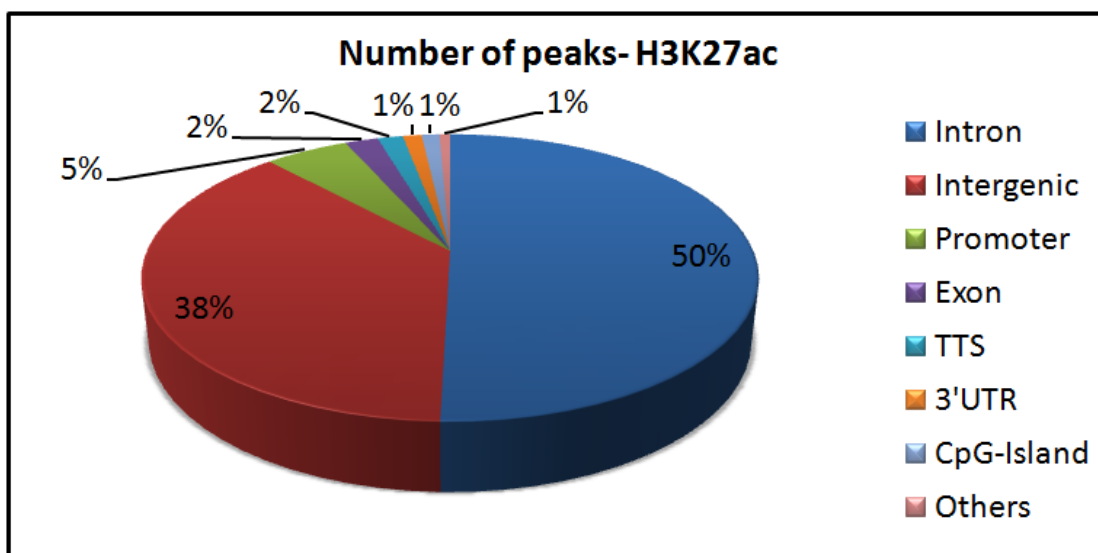
Figure 59: Distribution of the H3K27ac binding site location relative to RefSeq genes. Locations of binding sites are divided into: Intron, Intergenic, Promoter, Exon, Transcription Termination Site (TTS), 3′ Untranslated Region (UTR), CpG-island and others.
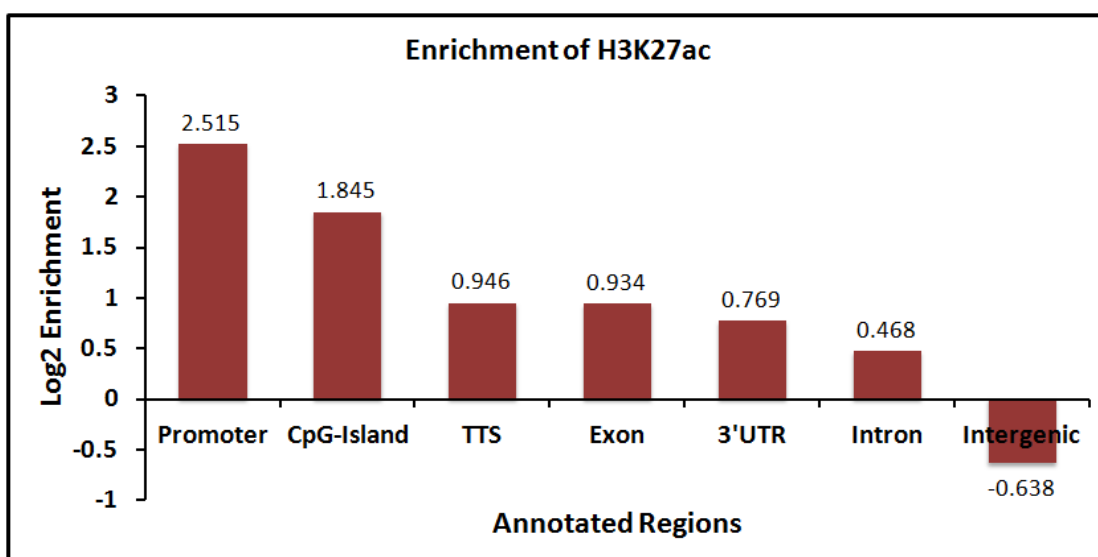
Figure 60: Enrichment of H3K27ac annotated regions relative to RefSeq genes. Annotated regions were divided into: Intron, Intergenic, Promoter, Exon, Transcription Termination Site (TTS), 3′ Untranslated Region (UTR), CpG-island and others.

5.3.4 Binding of H3K4me3 and H3K27ac near BEEC Candidate Genes

Genome Wide Expression Profiling study of Qi and colleagues have identified 162 differentially expressed BEEC candidate genes which fell into 11 biological network (BN) groups (Qi et al., 2011). In our ChIP-seq data a total of 133 of these genes

contained either H3K4me3 and/or H3K27ac binding sites. Description of these 133 BEEC candidate genes was listed in Appendix Table VI.

Binding of H3K4me3 near BEEC Candidate Genes

Our ChIP-Seq analysis on human urothelial cell line has identified H3K4me3 binding sites near 130 of these candidate genes (Table 13). H3K4me3 binding sites of these 130 genes were distributed in 7 different functional domains of the genome (Figure 61). We found that 56% of the binding sites near these genes are in the intergenic regions, following 27% and 9% in the intron and promoter regions, respectively. The exons contain 5% of the binding sites whereas 5'UTR, 3'UTR and TTS, each containing 1% of the binding sites. Comparing the H3K4me3 log2 enrichment data and distribution of binding sites of this histone among the 130 candidate gene is consistent with H3K4me3 binding sites present on and/or near the promoter and exon region of these genes acting as regulators for gene transcription.

Binding of H3K27ac near BEEC Candidate Gene

Mapping of H3K27ac has identified binding sites near 86 candidate genes of BEEC out of 162 genes (Qi et al., 2011) in SV-HUC-1 cell line (Table 13). The binding sites are distributed across six distinct functional domains. ChIP-seq analysis showed that 48% of these binding sites are located in the introns, along with 42% and 4% in the intergenic and promoter regions, respectively. A total of 6% binding sites were found in the exons, 3'UTRs and in TTS, containing 2% of the binding sites each. Combining the data from log2 enrichment (Figure 62) and distribution of the H3K27ac near 86 BEEC candidate genes, reveals that this histone mark is active in the introns possible near regulatory elements.
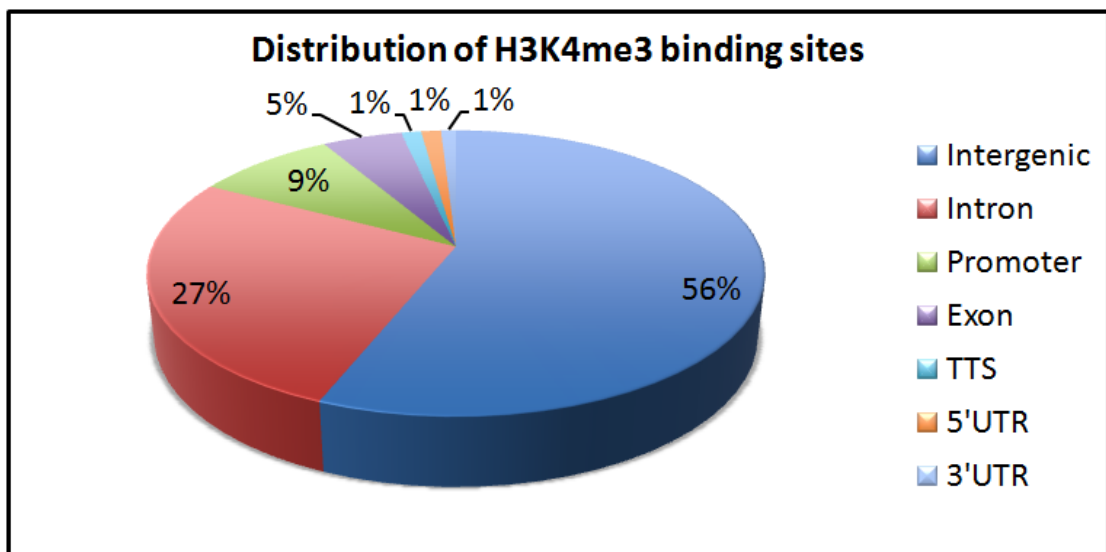
Figure 61: Distribution of H3K4me3 binding sites of 130 BEEC candidate genes across 7 different regions in human urothelial cell lines. Regions are divided into: Intergenic, Intron, Promoter-TSS, Exon, TTs, 5′UTRs, and 3′UTRs.
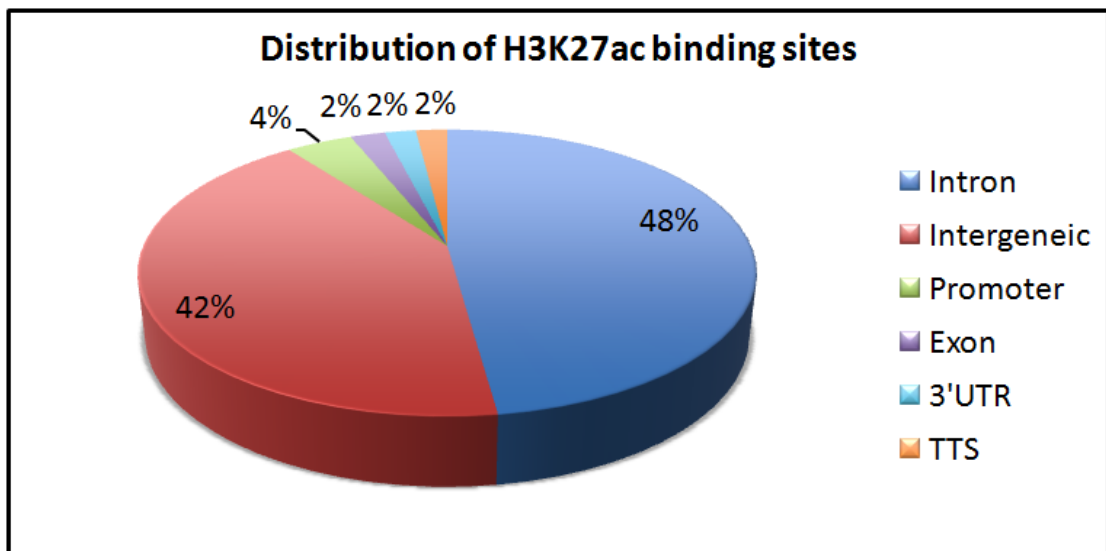
Figure 62: Distribution of H3K27ac binding sites of 86 BEEC candidate genes across 6 different regions in human urothelial cell line. Regions are divided into: Intergenic, Intron, Promoter-TSS, Exon, 5′UTRs, and 3′UTRs.

Table 13: Binding sites of H3K4me3 and H327ac found near differentially expressed BEEC candidate genes.

Biological Network	Functions	Genes Near H3K4me3 and H3K27ac Binding Sites	Number	
			H3K4me3	H3K27ac
BN1	Skeletal and Muscular Development and Function; Tissue Morphology; Cellular Assembly and Organization	*ACTN1*, CALD1*, CSRP1*, DES#, DSP*, DTNA, ERMP1*, FLNA, GEM*, KLF5*, LPP*, MYL9#, NEXN, PRKACB, PALLD*, PDLIM1*, PDLIM7*, PPIB*, PPP1R12A, S100A11*, SYNM, TNS1#, TPM1*, TPM2**	21	17
BN2	Cell Death; Cell Signalling; Embryonic Development	*AQP3*, ASS1*, BAG2, CDCA7*, CRABP2*, CTSC*, DHCR24*, FABP5, FHL1, GREM1, PHLDA2*, SC4MOL*, SFN*, SLC16A1*, SLC8A1, SOX9*, SYNPO2, TXNDC17*, WNT5A, ZAK**	20	13
BN3	Cancer; DNA Replication, Recombination and Repair; Genetic Disorders	*CA12*, CDK1*, DTL*, EFNB2, HGF, KIAA0101*, KRT18*, MCM4, MCM5*, NUSAP1, PCNA#, SACS, RRM2*, TOP2A, TSPO*, TYMS*, UHRF1*	16	9
BN4	Cell to Cell Signalling and Interactions; Cancer; Cellular Assembly and Organization	*ALDH1B1*, CD24, FERMT2*, FN1*, IGFBP2, INPP5A*, ITGA2*, ITGA5, ITGA6*, PRDX4, RHOB*, SLC25A4, TFRC*, TGM2*, TXNRD1**	15	10
BN5	Tissue Development; Organ Morphology; Amino Acid Metabolism	*AMMECR1, APRT*, DDX39A*, CA12*, CDCA7*, CSTB*, FAM129A*, GGH, GJC1*, GCOM1*, SHMT2, TIMM13, TMED3*	11	6
BN6	Developmental Disorder; Skeletal and Muscular Disorders; Tissue Development	*CAMK2G*, CDC42EP3*, CFL2, FLNC*, MLF1IP, MREG*, NUDT5*, PEX14*, PLEKHO1*, PRUNE2*, PTBP2, SGCB, SVIL**	13	9
BN7	Nervous System Development and Function; Neurological Disease; Cellular	*KLF5*, ITGA2*, MSRB3, NANS*, PERP*, RAB27B*, RNF144B*, SCHIP1*, SLMAP*, TUBA4A#*	7	7

Biological Network	Functions	Genes Near H3K4me3 and H3K27ac Binding Sites	Number	
			H3K4me3	H3K27ac
	Development			
BN8	Cell to Cell Signalling and Interactions; Cellular Movement; Tissue Development	*AFF3, AIM1*, ABRACL, **DSP***, JAM3, KIAA1217, MYO10*, SH3YL1*, ZNF185, ZEB1, ZWINT*	11	3
BN9	Connective Tissue Disorders; Inflammatory Disease; Skeletal and Muscular Disorders;	*ASS1*, CLIC4*, **CTSC***, GPR126*, TECR*, HN1*, NDFIP2-AS1, RAB23, PTBP3*, SCUBE3, SEMA3C*, SMTN**	10	8
BN10	Amino Acid Metabolism; Post Translational Modification; Small Molecule Biochemistry	*DMPK*, NUP50*, PTPLA, S100A16*, **TGM2****	5	3
BN11	Cellular Development; Gene Expression	*KCTD1**	1	1
		Total	130	86

Gene with asterisk (*) indicates both H3K4me3 and H3K27ac binding site found near and/or on these genes. Gene with hash ([#]) indicates only H3K27ac binding site found near and/or on these genes. Only H3K4me3 binding site found near and/or rest of the genes without any symbol. Genes in the '**Bold'** belongs to two different classes in the biological network.

5.3.5 H3K4me3 and H3K27ac Binding Sites in *DSP* and *PERP* Gene

We have described in Chapter 1, the possible involvement of *DSP* and *PERP* in BEEC pathogenesis. In Chapter 4 we showed that mutation screening of *DSP* and *PERP* promoters and exons did not reveal any potential sequence variants. We also hypothesized that dysregulation of gene expression may influence BEEC development. With this idea, enrichment of H3K4me3 and H3K27ac was analysed by ChIP-seq and peak score was used to interpret the enriched regions. The basic principle behind peak score is to identify regions in the genome where we find more sequencing reads than we would expect to see by chance (Heinz et al., 2010). Peak score >50 are considered as

strong enrichment and peak score<50 is considered as weak enrichment. The peak score>100 is considered as highly enriched.

Our data showed four H3K4me3 enriched loci near and/or in the *DSP* gene, and only one in *PERP* (Table 14). In *DSP,* the 1907 bp downstream locus of H3K4me3 showed a peak score of 234.1. In *PERP,* the1093 bp downstream locus showed a peak score of 101.1. The analysis also found that both these loci in *DSP* and *PERP* were in the intron regions, suggesting presence of potential regulatory elements contributing on the transcription of these genes.

Table 14: Enrichment of H3K4me3 near *PERP* and *DSP* genes.

Gene	Distance to TSS	Genomic Regions	Location of the Peaks	Peak Score
DSP	-37033	Intergenic	Chr6:7503841-7505833	32
	-9334	Intergenic	Chr6: 7531773-7533299	26.2
	-6249	Intergenic	Chr6:7535371-7535871	7.3
	1907	Intron 1	Chr6:7540843-7546712	**234.1**
PERP	1093	Intron 1	Chr6:138425600-138429534	**101.1**

Highest peak score (>50) of respective gene is shown in '**Bold**'

The mapping of H3K27ac identified a total of twenty enriched loci near and/or on the *DSP* gene, whereas only four binding sites were found in *PERP* (Table 15). The highest peak score of H3K27ac near and/or on the *DSP* was at -36962bpupstream of the promoter (peak score 134.1) and another was 1429 bp downstream of the promoter (peak score 120.3). In the case of *PERP*, the highest peak score of 10.2 was found in the intron region 1445 bp downstream of the promoter.

The presence of H3K4me3 and H3K27ac enriched with the highest peak scores in intron 1 of our target genes (*DSP* and *PERP*) encourage us to perform a bioinformatics

analysis on these regions. We then compare the analysis described in Chapter 4, Table 9; Figure 42 for *PERP*; and Table 11; Figure 48 for *DSP*; to determine the presence of any common transcription factor binding regions. Such analysis will give an idea of potential transcription factors that might be active in BBEC-related gene regulation.

Table 15: Enrichment of H3K27ac near *PERP* and *DSP* genes.

Gene	Distance to TSS	Genomic Regions	Location of the Peaks	Peak Score
	-68453	Intergenic	Chr6:7472951-7473884	14.3
	-59913	Intergenic	Chr6:7480228-7483687	40.8
	-56877	Intergenic	Chr6:7484356-7485631	18.9
	-46143	Intergenic	Chr6:7495290-7496165	10.7
	-42275	Intergenic	Chr6:7498997-7500194	17.3
	-36962	Intergenic	Chr6:7500830-7508987	**134.1**
	-31703	Intergenic	Chr6:7509729-7510605	11.7
DSP	-28000	Intergenic	Chr6:7513068-7514672	20.9
	-17632	Intergenic	Chr6:7523243-7525233	25
	-11231	Intergenic	Chr6:7525981-7535297	127
	-5465	Intergenic	Chr6:7535961-7536849	11.7
	-3566	Intergenic	Chr6:7537685-7538924	18.9
	1429	Intron 1	Chr6:7539447-7547152	**120.3**
	7924	Intron 1	Chr6:7548614-7550974	28.6
	10267	Intron 1	Chr6:7551747-7552527	11.7
PERP	1445	Intron 1	Chr6:138426873-138427558	10.2

Peak score>50 is shown in '**Bold'.** Peak score<10 are excluded from the table. All the binding sites for *DSP* and *PERP* are included in the Appendix Table VII.

Analysis of both H3K4me3 and H3K27ac high peak regions on intron 1 of *PERP*, revealed binding site of EP300, CTCF, HDAC2, MYC, FOXA1, FOXA2, FOS, GATA3, POLR2A and EGR1 transcription factors (Figure 63). We have already

described and discussed these transcriotion factors in Chapter 4 in Table 9 and Figure 42. We showed these factors may be involved in regulation of *PERP* during development. Combining the bioinformatics data from *DSP* promoter described in Chapter 4 in Table 11 and Figure 48, we found that H3K4me3 and H3K27ac high peak regions in intron 1 contain transcription factor binding sites such as MYC, JUND, EP300, FOXA1, FOXA2, HDAC2 and TBP (Figure 64). The presence of these transcription factor binding sites in the mapped H3K4me3 and H3K27ac regions both in the promoter and intron 1 of the targeted gene (*PERP* and *DSP*) supports their role in regulating transcription of the respective genes. It is very interesting to note that transcription factor binding sites for EP300, HDAC2, FOXA1 and FOXA2 are found in both the promoter and intron 1 of *PERP* and *DSP* gene, and these binding sites coincide with enrichment of H3K4me3 and H3K27ac, implicating a vital role in development.
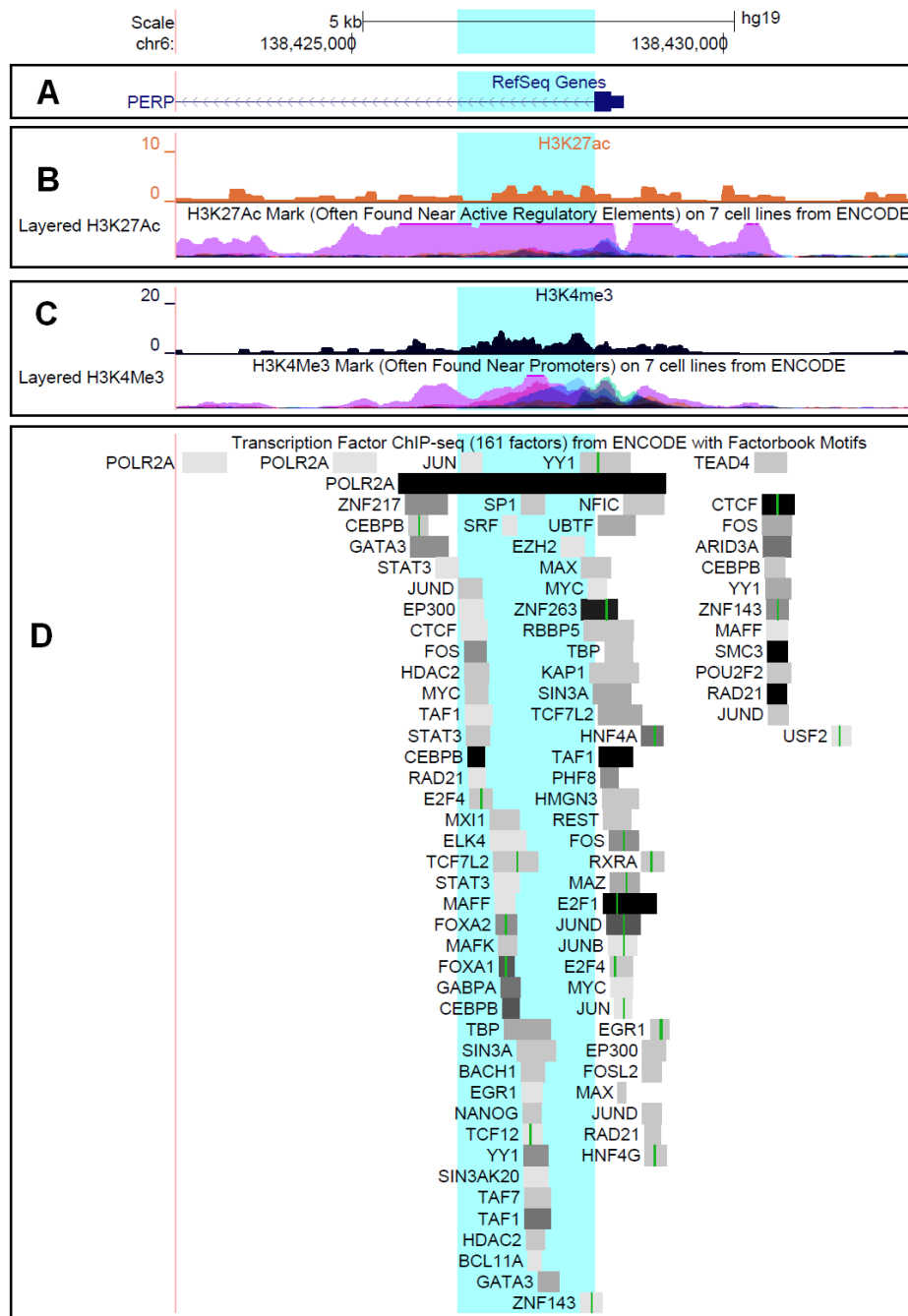
Figure 63: Bioinformatics analysis performed on the highest H3K4me3 and H3K27ac peak region in intron 1 of *PERP*. A. Reference sequence of *PERP* with TSS. B. Peak calling of our ChIP-seq data on H3K4me3 compared with the Layered H3K4me3 performed on 7 cell lines from ENCODE project (ENCODE Project Consortium, 2012). C. Peak calling of our ChIP-seq data on H3K27ac compared with the Layered H3K4me3 performed on 7 cell lines from encode project (ENCODE Project Consortium, 2012). D. Transcription factor binding sites in the high peak region in *PERP*. The highlighted portion indicates the genomic regions with high peak scores. The image is derived from the UCSC Genome Browser (http://genome.ucsc.edu).

136
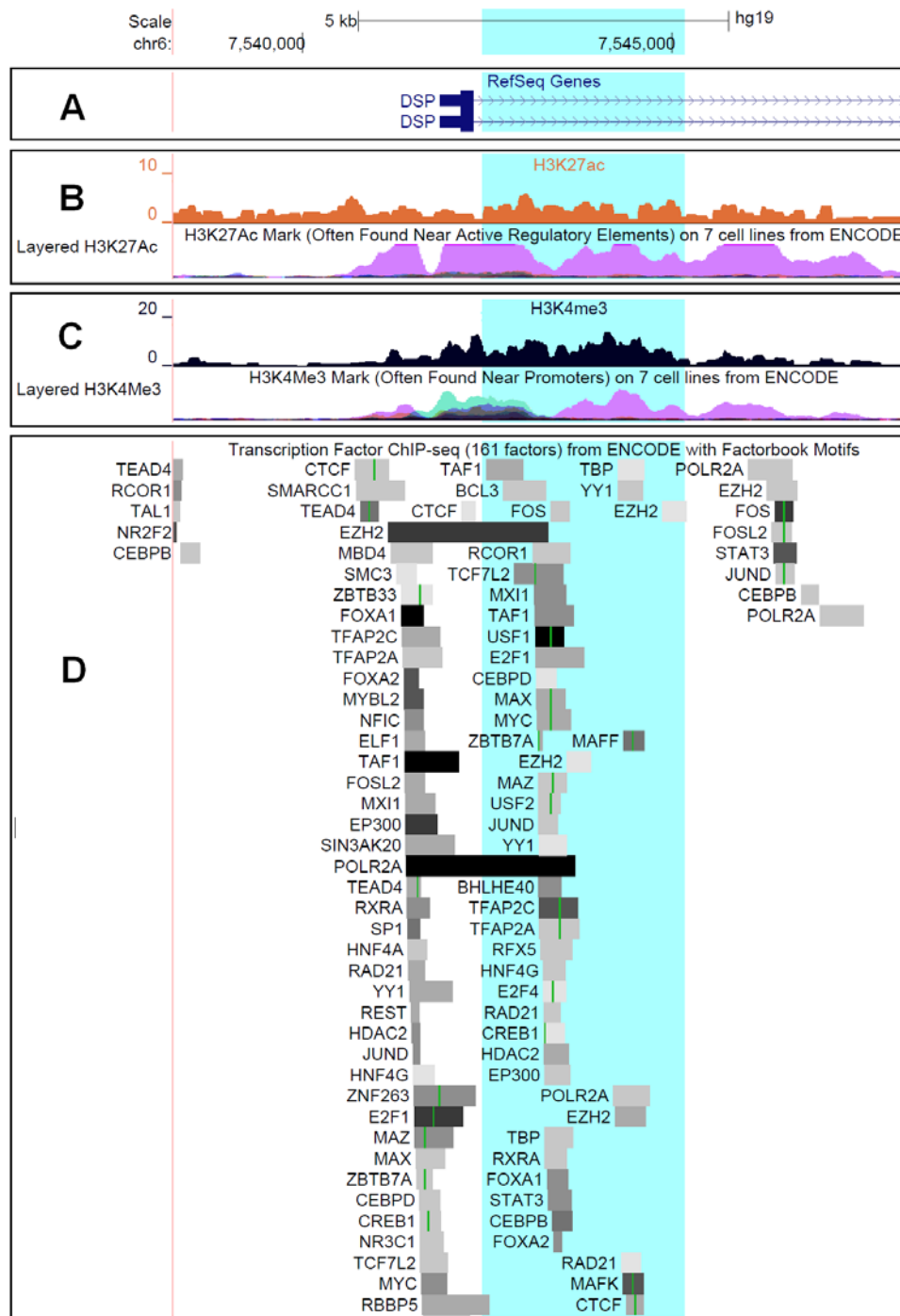
Figure 64: Bioinformatics analysis performed on the H3K4me3 and H3K27ac high peak region in intron 1 of *DSP*. A. Reference sequence of *DSP* with TSS. B. Peak calling of our ChIP-seq data on H3K4me3 compared with the Layered H3K4me3 performed on 7 cell lines from encode project (ENCODE Project Consortium, 2012). C. Peak calling of our ChIP-seq data on H3K27ac compared with the Layered H3K4me3 performed on 7 cell lines from encode project (ENCODE Project Consortium, 2012). D. Transcription factor binding sites in the high peak region in *DSP*. The highlighted portion indicates the genomic regions with high peak scores.

5.3.6 Enrichment of H3K4me3 and H3K27ac near the most Up- and Down-regulated BEEC Candidate Genes

Qi and co-workers in their research examined the BEEC candidate gene expression in mid gestation mouse embryos. According to their study, the five top ranked up-regulated gene were *CRABP2, SC4MOL, MYO10, PHLDA2* and *WNT5A* (Qi et al., 2011). They have also identified 11 top ranked down-regulated genes *i.e. ACTC1, ACTN1, CALD1, CSRP1, DES, FLNA, FLNC, FN1, LPP, TNS1* and *TPM1* in their study. This expression profile encouraged us to investigated H3K4me3 and H3K27ac enrichment at these genes in human urothelial cell lines to establish a possible link between these genes and BEEC.

The major functions of these genes are described in Table 13, and they belong to different Biological Network (BN) groups. Four out of five up-regulated genes e.g. *CRABP2, SC4MOL, MYO10, PHLDA2* belongs to the BN2 (Cell Death; Cell Signalling; Embryonic Development), and only one of the up-regulated genes *i.e.* MYO10 fall under BN8 (Cell to Cell Signalling and Interactions; Cellular Movement; Tissue Development). Among these most down-regulated genes; *ACTN1, CALD1, CSRP1, DES, FLNA, LPP, TNS1* and *TPM1,* belong to BN1 (Skeletal and Muscular Development and Function; Tissue Morphology; Cellular Assembly and Organization), whereas *FLNA* and *FN1* fall under BN4 (Cell to Cell Signalling and Interactions; Cancer; Cellular Assembly and Organization) and BN6, respectively. The gene *ACTC1* belongs to 2 biological network group *i.e.* BN3 (Cancer; DNA Replication, Recombination and Repair; Genetic Disorders) and BN5 (Tissue Development; Organ Morphology; Amino Acid Metabolism). However, we did not find any H3K4me3 and H3K27ac enrichment near *ACTC1.*

Mapping of H3K4me3 in the most Up- and Down-regulated Candidate Genes

Our data identified some enrichment of H3K4me3 near these up-regulated genes (Table 16). We found high peak scores, 85.8 and 107.6, in the intron 1 of *SC4MOL* and *MYO10* gene respectively. The highest peak score (74.2) for *PHLDA2* was found in exon 1 of the gene. Among these five top ranked up-regulated genes, *WNT5A* showed the highest peak score (162.9) in the Promoter-TSS region at 85 bp downstream (Figure 65). The peak score data on H3K4me3 of these up regulated genes suggests that regulations of most of these genes are linked with the intron 1, near the TSS.

Table 16: H3K4me3 peak locations near the most up-regulated BEEC candidate genes.

Gene	Distance to TSS	Genomic Regions	Location of the Peaks	Peak Score
CRABP2	3790	Intron 1	Chr1:156671419-156671919	8
	419	Intron 1	Chr1:156674099-156675981	**34.9**
SC4MOL	863	Intron 1	Chr4:166247406-166251956	**85.8**
MYO10	1684	Intron 1	Chr5: 16932102-16937300	**107.6**
	46628	Intron 1	Chr5:16889419-16890096	10.9
PHLDA2	104	Exon 1	Chr11:2948845-2952248	**74.2**
WNT5A	85	Promoter-TSS	Chr3:55518720-55524450	**162.9**

Highest peak score of respective gene is shown in '**Bold**'
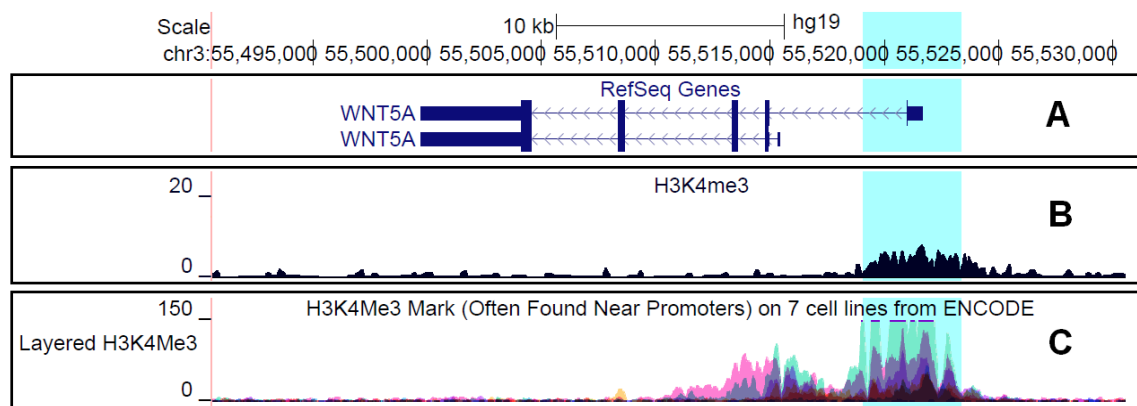
Figure 65: H3K4me3 enrichment in the promoter of the *WNT5A* gene. A. Reference sequence of *WNT5A* highlighting the region of interest. B. Peak calling of our ChIP-seq data with H3K4me3 near *WNT5A* gene, performed on human urothelial cell line. The highlighted portion indicates the genomic regions with high peak scores. C. Layered H3K4me3 peaks from ENCODE project (ENCODE Project Consortium, 2012), showing the same high peak regions performed on 7 cell lines. Each colour denotes peak calling of different cell lines.

For the most down-regulated genes, the highest peak scores (>150) were found in *ACTN1* (205.0), *FLNC* (207.9), *FN1* (196.3) and *TPM1* (169.4). The peak score >50 was found in *CALD1* (106.2 and 99.6) and *CSRP1* (85.5) (Table 17). Among these down-regulated genes, the highest peak score was found in *FLNC* at 524 bp down-stream of the TSS (Figure 66). All the highest peaks of these genes were found in the introns or in exons, suggesting that gene regulation is linked with these high peak regions.

Table 17: H3K4me3 peak locations near the most down-regulated BEEC candidate gene.

Gene	Distance to TSS	Genomic Regions	Location of the Peaks	Peak Score
ACTN1	1365	Intron 1	Chr14:69440795-69448642	**205**
	1116	Intron 1	Chr7:134463364-134467197	**106.2**
CALD1	1328	Intron 1	Chr7:134575808-134579151	**99.6**
	8327	Intron 1	Chr7:134584068-134584888	11.6

Gene	Distance to TSS	Genomic Regions	Location of the Peaks	Peak Score
CSRP1	122	Promoter-TSS	Chr1:201474060-201478470	**86.5**
FLNA	502	Intron 1	ChrX:153602050-153602958	13.8
	2044	Intron 1	ChrX:153600590-153601334	12.4
	4187	Intron 2	ChrX:153598234-153599404	21.8
FLNC	524	Exon 1	Chr7:128467896-128474118	**207.9**
FN1	-183927	Intron 8	Chr2:216484391-216485045	10.9
	396	Exon 1	Chr2:216296060-216304731	**196.3**
LPP	41242	Intron 2	Chr3: 187984106-187984765	11.6
	70053	Intron 2	Chr3:188012832-188013661	**13.1**
TPM1	714	Intron 1	Chr15:63340718-63341983	17.4
	1238	Intron 2	Chr15:63332747-63339406	**169.4**

Highest peak score of respective gene is shown in '**Bold'.** Peak score <10 are excluded from the table. All peaks are listed in Appendix Table VIII.
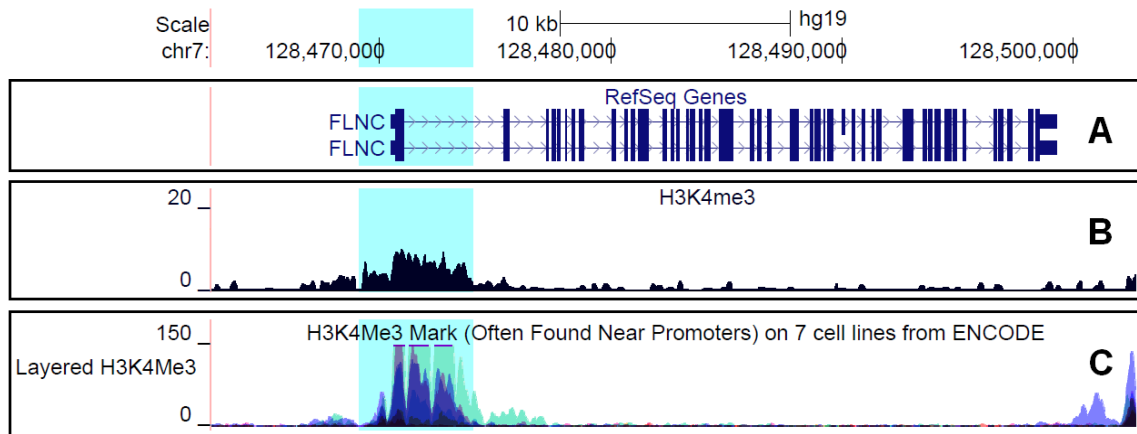
Figure 66: H3K4me3 enrichment in the exon 1 of *FLNC* gene. A. Reference sequence of *FLNC* highlighting the region of interest. B. Peak calling of our ChIP-seq data with H3K4me3 on *FLNC* gene, performed on human urothelial cell line. The highlighted portion indicates the genomic regions with high peak scores. C. Layered H3K4me3 peaks from ENCODE project (ENCODE Project Consortium, 2012), showing the same high peak regions performed on 7 cell lines. Each colour denotes peak calling of different cell lines. The image is derived from the UCSC Genome Browser (http://genome.ucsc.edu).

Mapping of H3K27ac in the most Up- and Down-regulated Candidate Genes

As for H3K4me3, a similar study was performed on H3K27ac enrichment to locate possible regulatory regions. Compared to the H3K4me3 binding sites, the peak scores of H3K27ac were low (<50) (Table 18). The highest peak scores for *CRABP2* (7.1) and *PHLDA2* (17.3) were found in the intergenic regions. For *SC4MOL* and *MYO10* highest peak score was 8.2 and 7.1 found in the intron 1 and intron 2, respectively. Among these most up-regulated genes, the highest peak score was found near *PHLDA2* in the intergenic region (Figure 67) at -3087 bp upstream of TSS.

Table 18: H3K27ac peak locations near the most up-regulated BEEC candidate gene.

Gene	Distance to TSS	Genomic Regions	Location of the Peaks	Peak Score
CRABP2	-5681	Intergenic	Chr1:156681039-156681539	**7.1**
	1767	Intron 1	Chr1:156673442-156673942	6.6
SC4MOL	-1208	Intergenic	Chr4:166247360-166247860	7.1
	3727	Intron 1	Chr4:166252295-166252795	**8.2**
MYO10	-65669	Intergenic	Chr5:17001804-17002304	6.6
	65868	Intron 2	Chr5:16870267-16870767	**7.1**
	125874	Intron 3	Chr5:16810261-16810761	6.6
PHLDA2	-5288	Intergenic	Chr11:2955573-2956303	10.2
	-3078	Intergenic	Chr11:2953010-2954446	**17.3**
	826	Exon 2	Chr11:2949574-2950074	6.6
	1859	TTS	Chr11:2948541-2949041	7.1

Highest peak score of respective gene is shown in '**Bold'**
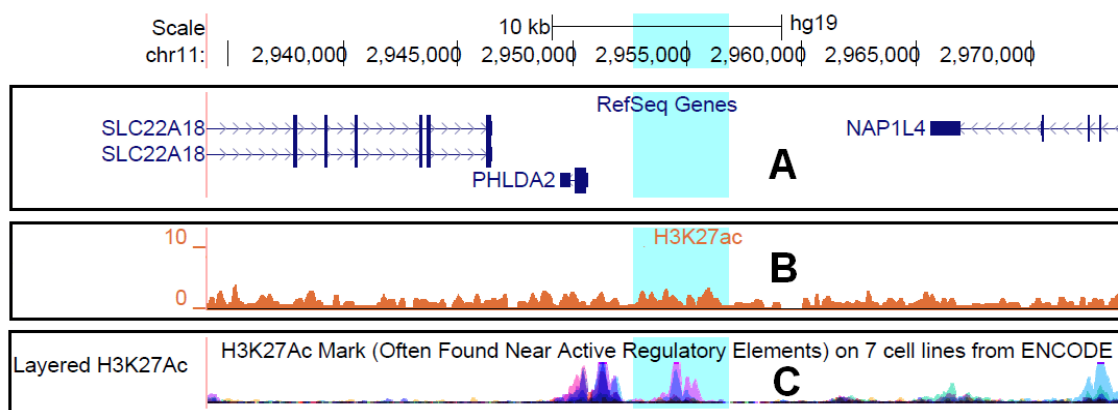
Figure 67: H3K27ac enrichment in the intergenic region of *PHLDA2* gene. A. Reference sequence of *PHLDA2* highlighting the region of interest. B. Peak calling of our ChIP-seq data with H3K27ac near *PHLDA2* gene, performed on human urothelial cell line. The highlighted portion indicates the genomic regions with high peak scores. C. Layered H3K27ac peaks from ENCODE project (ENCODE Project Consortium, 2012), showing the same high peak regions performed on 7 cell lines. Each colour denotes peak calling of different cell lines. The image is derived from the UCSC Genome Browser (http://genome.ucsc.edu).

Several H3K27ac enriched regions were found in the most down-regulated BEEC candidate genes. We found 14 binding sites for *ACTN1*, 6 for *CALD1,* 10 for *CSRP1*, 1 for *DES*, 4 for *FN1*, and 7 for LPP. In the case of *FLNC*, *TNS1* and *TPM1*, we found 5 H3K27ac enriched regions for each of these genes (Appendix Table IX). Peak score ≥10 of these genes are shown in the Table 19. Most of the peak scores for *ACTN1* and *CALD1* were found in intron 1, near the TSS. The highest peak scores for these two genes were 18.9 and 28.0 at 4325 bp and 1431 bp downstream, respectively. A peak score of 11.2 was found in intron 1 of *FLNC* gene at 586 bp and 4490 bp downstream to TSS. The most down-regulated BEEC candidate gene, *CSRP1,* has the highest peak score (45.4) found in the intron 3 of the gene at 7333 bp down-stream of TSS (Figure 68).

Table 19: H3K27ac peak locations near the most down-regulated BEEC candidate gene.

Gene	Distance to TSS	Genomic Regions	Location of the Peaks	Peak Score
ACTN1	4325	Intron 1	Chr14:69441123-69442393	**18.9**
	12059	Intron 1	Chr14:69433625-69434424	11.2
	15976	Intron 1	Chr14:69429734-69430480	10.2
	38177	Intron 1	Chr14:69407269-69408543	17.3
CALD1	1431	Intron 1	Chr7:134464452-134466738	**28**
	1600	Intron 1	Chr7:134576758-134578745	25
	11760	Intron 1	Chr7:134587338-134588484	15.8
CSRP1	1457	Intron 1	Chr1:201474045-201474975	15.8
	7332	Intron 3	Chr1:201456750-201459988	**45.4**
FLNC	586	Intron 1	Chr7:128470692-128471447	**11.2**
	4490	Intron 1	Chr7:128474556-128475391	**11.2**
FN1	-30124	Intergenic	Chr2:216330561-216331269	10.7
	-26560	Intergenic	Chr2:216326948-216327754	**11.7**
	1013	Intron 1	Chr2:216299462-216300095	10.2
LPP	-1156	Exon 1	Chr3:187869855-187871159	**15.3**
TNS1	-11190	Intergenic	Chr2-218819279-218820694	**19.9**
	19316	Intron 1	Chr2-218789067-218789893	11.2

Highest peak score of respective gene is shown in '**Bold'**
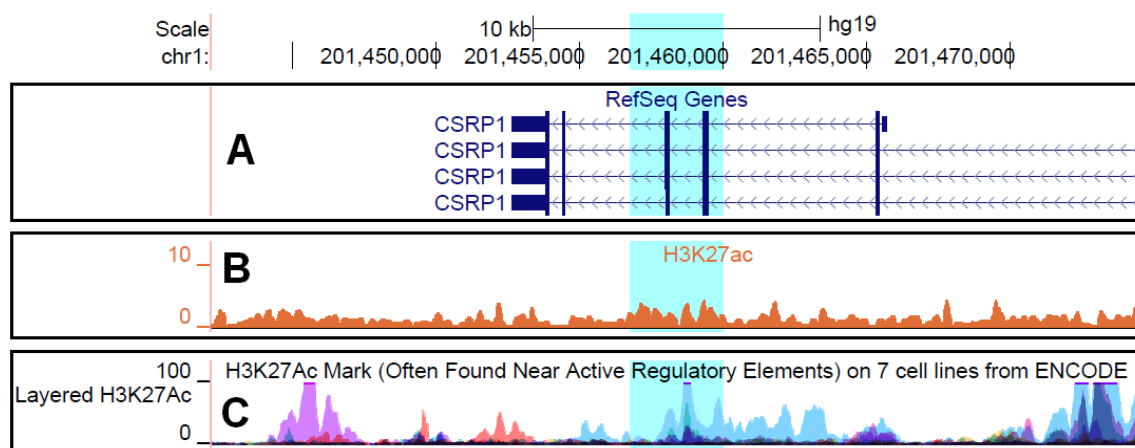
Figure 68: H3K27ac enriched regions in the intron 3 of *CSRP1* gene. A. Reference sequence of *CSRP1* highlighting the region of interest. B. Peak calling of our ChIP-seq data with H3K27ac on *CSRP1* gene, performed in human urothelial cell line. The highlighted portion indicates the genomic regions with high peak scores. C. Layered H3K27ac peaks from ENCODE project (ENCODE Project Consortium, 2012), showing the same high peak regions, performed on 7 cell lines. Each colour code denotes peak calling of different cell lines. The image is derived from the UCSC Genome Browser (http://genome.ucsc.edu).

5.3.7 Overlapping Enrichment of H3K4me3 and H3K27ac with Published *TP63* Binding Data

We previously showed that TP63 is a candidate gene for BEEC (Cheng et al., 2006), and in/del polymorphisms in the promoter of the *TP63* predominant isoform, *ΔNP63,* are associated with increased risk of this disease (Wilkins et al., 2012). In chapter 3 we showed that there is no sequence variation detected in the promoter of the other *TP63* isoform, *TAP63*(Darling et al., 2013), and also no sequence variation was detected in the *TP63* exons (Ching et al., 2010; Wilkins et al., 2012) that could contribute to the disease pathogenesis. Such findings encouraged us to investigate whether there is any overlap between the H3K4me3 and H3K27ac with published TP63 binding sites (Kouwenhoven et al., 2010; Sethi et al., 2014).

The H3K4me3 and TP63 intersect data gave a total of 762 peaks, of which 47% was found in the intron regions, followed by 32% in the intergenic regions. 14% of the overlapping peaks were located in the promoter regions (Figure 69). Exon, TTS and 5'UTR each contains 2% of the peaks, whereas the 3'UTR contains 1% of the peaks. Introns and promoters together contain about 61% of the peaks, meaning any of these overlapping sites may act as a regulator for gene transcription where TP63 may be a major contributor.
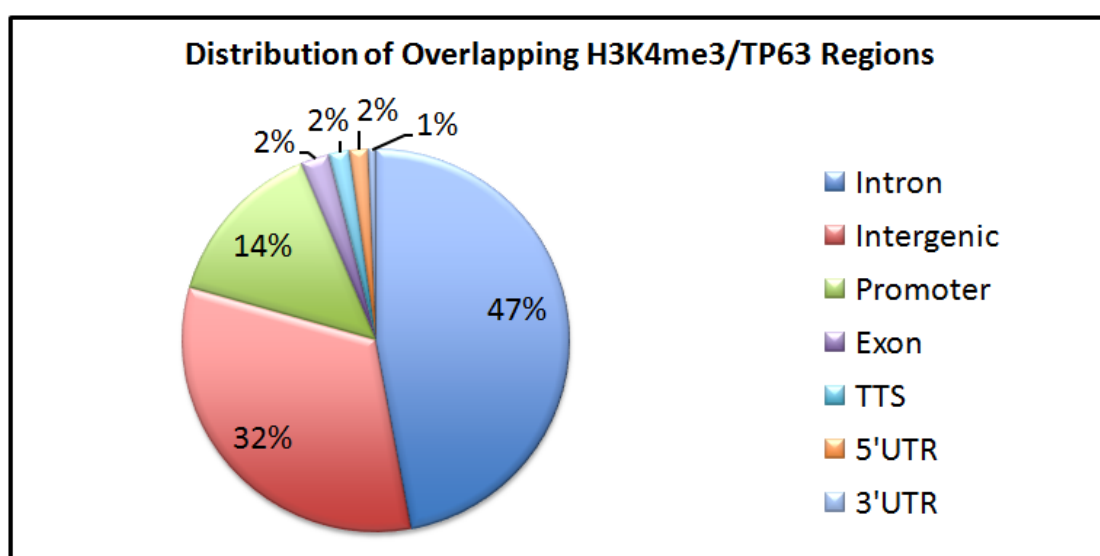
Figure 69: Distribution of H3K4me3 enrichment overlapped with the published *TP63* binding site in the genome. Overlapping H3K4me3/TP63 loci are distributed over eight distinct genomic regions *i.e.* Intron, Intergenic, Promoter, Exon, TTS, 5'UTR, 3'UTR and Non-coding.

In the case of H3K27ac and TP63 intersect data, a total of 808 peaks were found. Among these peaks 46% were located in the intergenic regions, whereas 44% of them were found in the introns. The Exons, 5'UTRs and 3'UTRs regions each contains 1% of the peaks, whereas 2% was found in TTS (Figure 70). The promoter region contains 5% of the peaks which is lower than the H3K4me3/TP63 overlapping peaks, suggesting that the TP63 binding site are less active in these regions. Therefore, it is possible that overlapping TP63 peaks in the intron regions may act as regulators for gene transcription.
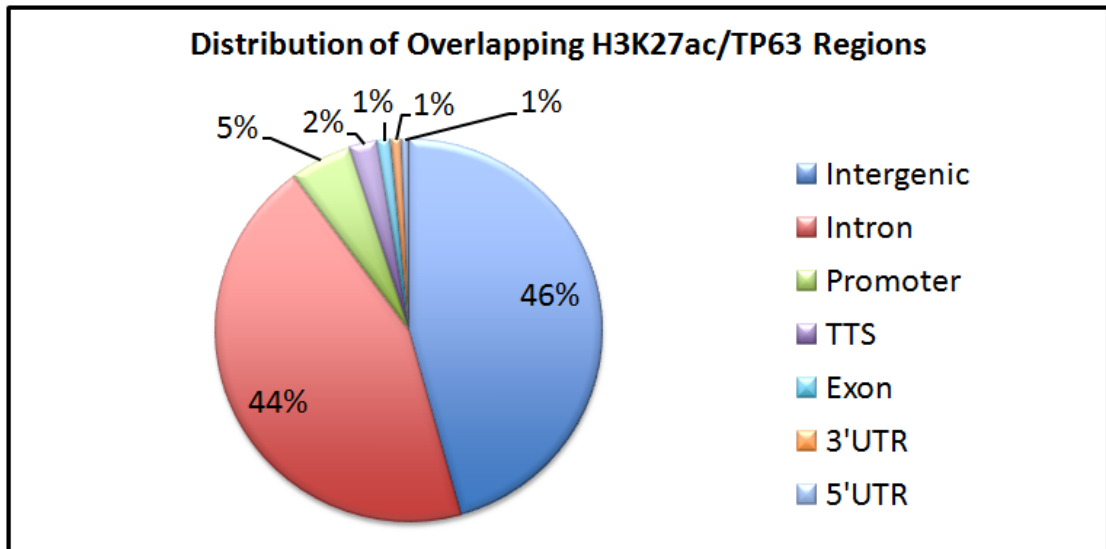
Figure 70: Distribution of H3K27ac enrichment overlapped with the published *TP63* binding site in the genome. Overlapping H3K4me3/TP63 loci are scattered over 8 distinct genomic regions *i.e.* Intron, Intergenic, Promoter, Exon, TTS, 5'UTR, 3'UTR and Non-coding.

5.3.8 Overlapping Binding sites of H3K4me3/TP63 and H3K27ac/TP63 near Published BEEC Candidate Genes.

We also investigated overlapping binding sites of H3K4me3/TP63 and H3K27ac/TP63 near published BEEC candidate genes (Qi et al., 2011). The H3K4me3/TP63 intersect data identified 15 candidate genes with overlapping binding regions (Table 20). Peak distribution of these genes across the genomic regions showed equal percentage of peaks (33% each) found in the intergenic and intron regions. Promoter regions contain 29% and TTS contains 5% of the peaks (Figure 71). Bioinformatics analysis of the individual overlapping peak region revealed binding sites of several important transcription factors (Table 20), that are previously discussed in Chapter 3 and 4. Considering the log2 enrichment of the H3K4me3 (Figure 58), it can be concluded that genes containing overlapping sites in the promoter region could be promising candidate genes for BEEC pathogenesis. However, intron regions also show positive enrichment of H3K4me3 (Figure 58), so they cannot be excluded as being involved in BEEC.

Table 20: Overlapping H3K4me3/TP63 loci near published BEEC candidate gene.

Gene	Distance to TSS	Genomic Regions	Location of the Peaks	Transcription Factors
CDCA7	1445	Intron 1	Chr2:174220852-174221160	N/A
*DSP**	-36074	Intergenic	Chr6:7505760-7505833	JUN, FOS, EP300, HDAC2
ITGA2	18206	Intron 1	Chr5:52303193-52303531	FOS
MREG	7085	Intron 1	Chr2:216871056-216871467	TBP, MYC
MYO10	659	Intron 1	Chr5:16935590-16935862	MYC, FOS
*PCNA**	-102	Promoter-TSS	Chr20:5100449-5101049	FOS, HDAC2, HNF4A
PERP	2066	Intron 1	Chr6:138426370-138426818	JUN, FOS, EP300, HDAC2
*PTBP3**	71	Promoter-TSS	Chr9:115095075-115095395	CTCF, FOXA2
*RNF144B**	-61103	Intergenic	Chr6:18326349-18326607	CTCF
	-46135	Intergenic	Chr6:18341199-18341693	MYC, FOS
	-34190	Intergenic	Chr6:18353138-18353645	MYC, FOS
	-155	Promoter-TSS	Chr6:18387120-18387733	FOXA1, FOXA2, MYC, JUND
S100A11	12847	Intergenic	Chr1:151996630-151996698	N/A
SCHIP1	-107516	Intron 1	Chr3:159450119-159450149	FOXA1, EP300
	-55577	Intron 4	Chr3:159501779-159502367	FOXA1, EP300, FOS, MYC
*SEMA3C**	-22130	Intergenic	Chr7:80570578-80571017	JUN, MYC, FOS, FOSL2, EP300
	-862	Promoter-TSS	Chr7:80549312-80549746	MYC, EP300
*SFN**	-3846	Intergenic	Chr1:27185541-27186033	FOS
	2095	TTS	Chr1:27191265-27192191	RELA, FOSL1, FOS, JUN, MYC
SMTN	769	Promoter-TSS	Chr22:31477830-31478273	FOS, POLR2A
*TGM2**	-730	Promoter-TSS	Chr20:36794198-36794662	N/A

Genes with the asterisk (*) also found in the H3K27ac/TP63 intersect data.
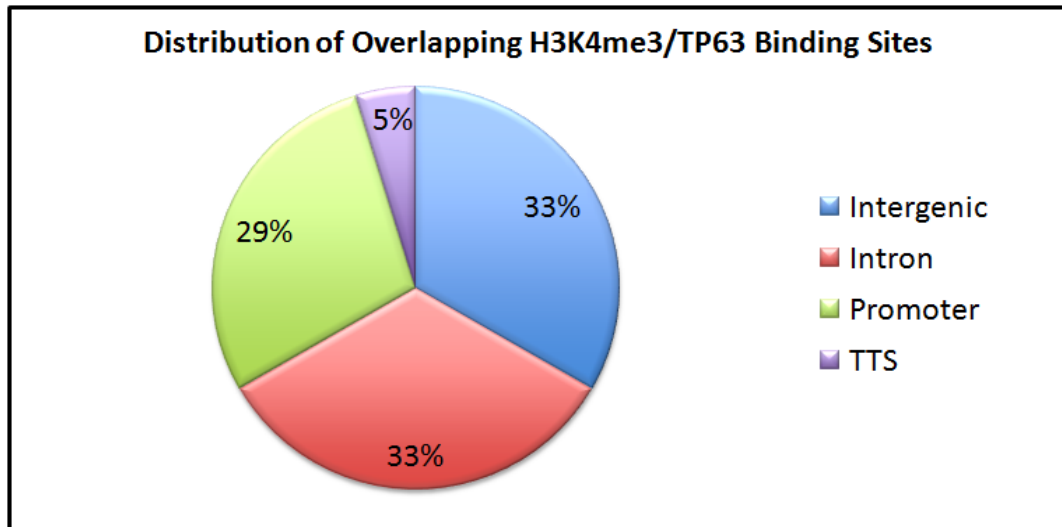
Figure 71: Distribution of H3K4me3 enrichment overlapped with the published TP63 binding site found near and/or on the BEEC candidate genes. Overlapping H3K4me3/TP63 loci are scattered over 4 distinct genomic regions *i.e.* Intron, Intergenic, Promoter, and TTS.

Our H3K27ac/TP63 intersect data located 11 BEEC candidate genes with overlapping loci (Table 21). Peak distribution for these 11 genes showed that 58% of the peaks are located in the intergenic regions, followed by 21% and 16% in the intron and promoter regions, respectively. Only 5% of the peaks are located in the non-coding regions (Figure 72). Bioinformatics analysis of the individual overlapping peak region showed binding sites of several important transcription factors (Table 21), that are previously discussed in Chapter 3 and 4. Considering the log2 enrichment data of H3K27ac, it can be said that overlapping sites in the promoter and intron may harbour potential regulatory elements.

From our overlapping data near published BEEC candidate gene (Qi et al., 2011), we found five of the gene are present in both H3K4me3/TP63 and H3K27ac/TP63 intersect regions *i.e. DSP*, *PCNA*, *PTBP3*, *RNF144B*, *SEMA3C* and *TGM2* (Table 20 and 21). Among these five genes, *DSP* is a vital member of TP63>PERP>Desmosome pathway (Mahfuz et al., 2013b) and also crucial BEEC pathogenesis as described in Chapter 1.

Table 21: Overlapping H3K27ac/TP63 loci near published BEEC candidate gene.

Gene	Distance to TSS	Genomic Regions	Location of the Peaks	Transcription Factors
ASS1	-17605	Intergenic	Chr9:133302260-133302719	FOS, MYC. JUN, CTCF
DSP*	-35904	Intergenic	Chr6:7505760-7506173	FOSL2, JUND, FOS, EP300
	-61465	Intergenic	Chr6:7480229-7480581	FOS, MYC. JUN, EP300, POLR2A
	12150	Intron 1	Chr6:7553822-7554219	MYC, EP300, FOXA1, POLR2A
ITGA6	9721	Intron 1	Chr2:173301933-173302138	POLR2A
LPP	448741	Intron 6	Chr3:188391901-188391967	FOS, MYC, JUN, EP300,
PCNA*	147	Non-coding	Chr20:5100449-5100552	N/A
PTBP3*	71	Promoter-TSS	Chr9:115095075-115095395	CTCF, EP300, TBP, MYC
RHOB	-45495	Intergenic	Chr2:20601132-20601548	CTCF, MYC
RNF144B*	-61103	Intergenic	Chr6:18326349-18326607	CTCF
	-46135	Intergenic	Chr6:18341199-18341693	FOS, POLR2A, MYC
	-34190	Intergenic	Chr6:18353138-18353645	FOS, POLR2A, MYC
	23324	Intron 2	Chr6:18410685-18411126	MYC
	-155	Promoter-TSS	Chr6:18387120-18387733	FOS, MYC, JUN, EP300, HDAC2, FOXA1, FOXA2
SEMA3C*	-154537	Intergenic	Chr7:80703093-80703316	FOS, MYC,
	-22312	Intergenic	Chr7:80570942-80571017	FOS, MYC, JUN, FOSL2
SFN	-3957	Intergenic	Chr1:27185319-27186033	FOS, POLR2A,
TGM2*	-18896	Intergenic	Chr20:36812396-36812796	FOS, FOXA1, FOS, EP300
	-720	Promoter-TSS	Chr20:36794198-36794643	N/A

Genes with the asterisk (*) also found in the H3K4me3/TP63 intersect data.
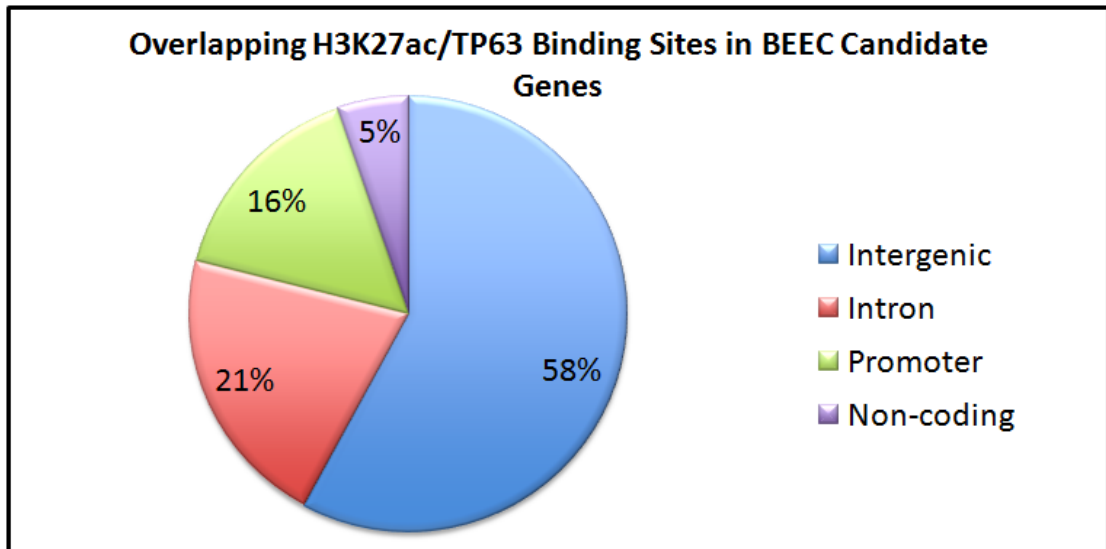
Figure 72: Distribution of H3K27ac enrichment overlapped with the published TP63 binding site found near and/or on the BEEC candidate genes. Overlapping H3K27c/TP63 loci are scattered over 8 distinct genomic regions *i.e.* Intron, Intergenic, Promoter, Exon, TTS, 5'UTR, 3'UTR and Non-coding.

5.3.9 Known Motif Enrichment in H3K4me3 and H3K27ac in Human Urothelial Cell Line

A *de novo* motif search was performed on the 36867 regions with H3K4me3 enrichment. The top three ranked motifs for H3K4me3 are IRF2, ISRE and SCL (Figure 73). For H3K27ac, motif search was performed on 43942 regions. The top three ranked motifs for H3K27ac are SCL, EBF1 and ZFX (Figure 74). It is interesting to find that the SCL motif is present in bothH3K4me3 and H3K27ac target sequence among the top three ranked binding motifs.
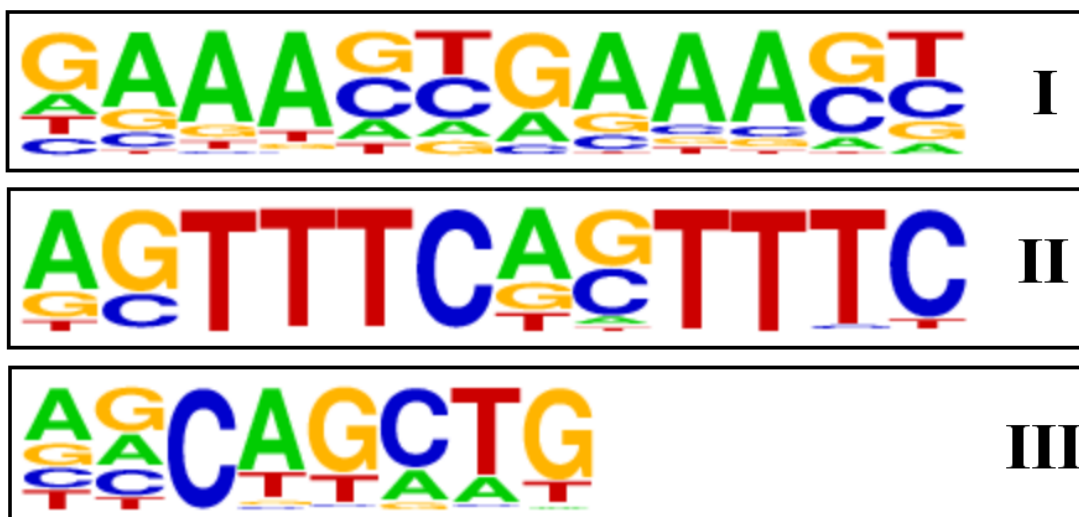
Figure 73: Three top ranked binding motifs in H3K4me3 target sequences. I. IRF2 (IRF), II. ISRE (IRF) and III. SCL. Enriched motifs were identified by HOMER *de novo* motif analysis.
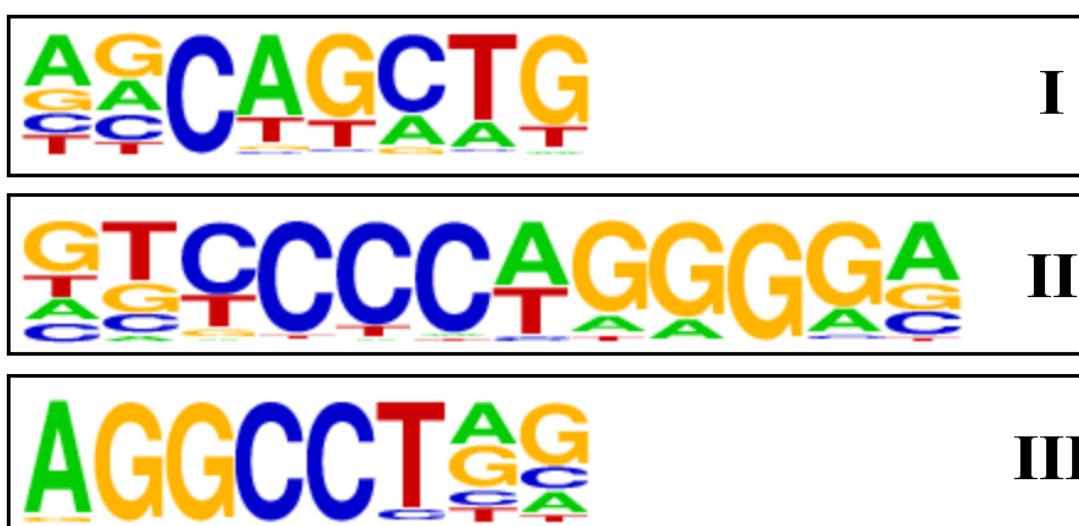
Figure 74: Three top ranked binding motifs in H3K27ac target sequences. I. SCL, II. EBF1 (EBF) and III. ZFX (Zf). Enriched motifs were identified by HOMER *de novo* motif analysis.

5.3.10 Relative Enrichment of H3K4me3 and H3K27ac in Targeted Loci of BEEC Tissue Sample

We also performed ChIP assay using H3K4me3 antibodies on tissue from two BEEC patients. The sample from patient 1 (P1) is bladder mucosa tissue, and in patient 2 (P2) the sample is from ureter. Relative enrichment of H3K4me3 in P1 was measured by QPCR using designed primer shown in Appendix Table V. Fold enrichment on targeted

loci (*ΔNP63*, *PERP* and *DSP*) were very low compared to the negative enrichment region (FOS-) of H3K4me3 (Figure 75). The value of ChIP negative control is low *i.e.* IgG<0.05, indicating low background noise in ChIP assay. In case of P2, enrichment of H3K4me3 on the targeted gene loci, (*ΔNP63*, *PERP* and *DSP*) is higher than the P1 patient. Fold-enrichment data on P2 showed ~1.5 fold in *DSP* and *PERP* loci, whereas in *ΔNP63* loci, it close to ~0.5 fold. In the P2 patient IgG is also <0.05, meaning low background noise in ChIP assay (Figure 76).
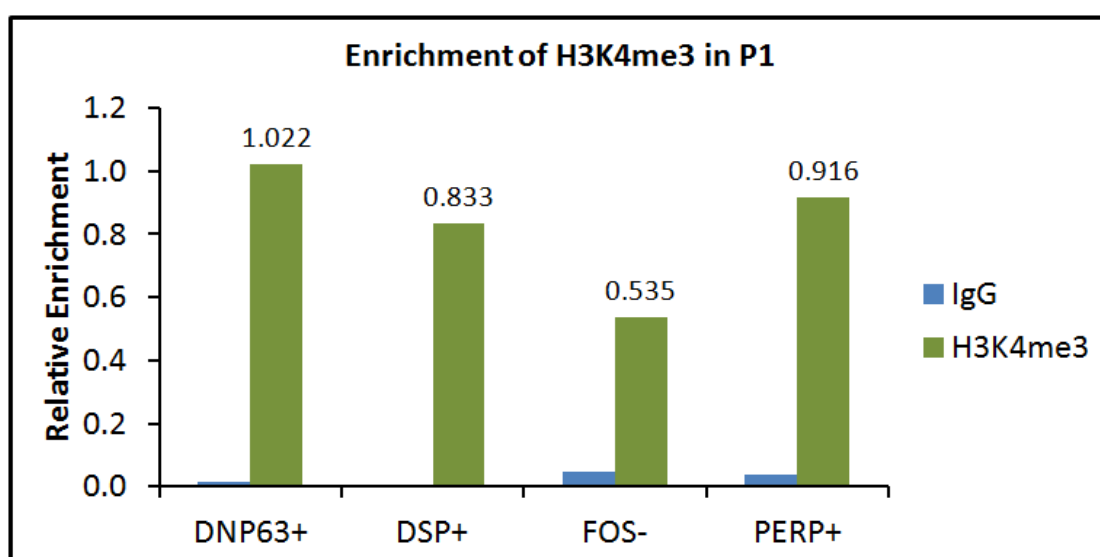
Figure 75: Relative enrichment of H3K4me3 in targeted gene loci related to P63 pathway in patient (P1) bladder mucosa tissue. FOS- locus was used as negative enrichment region H3K4me3 binding sites. DNP63+ (ΔNP63+), DSP+ and PERP+ are targeted loci. The IgG was used as ChIP negative control to determine the background signal.
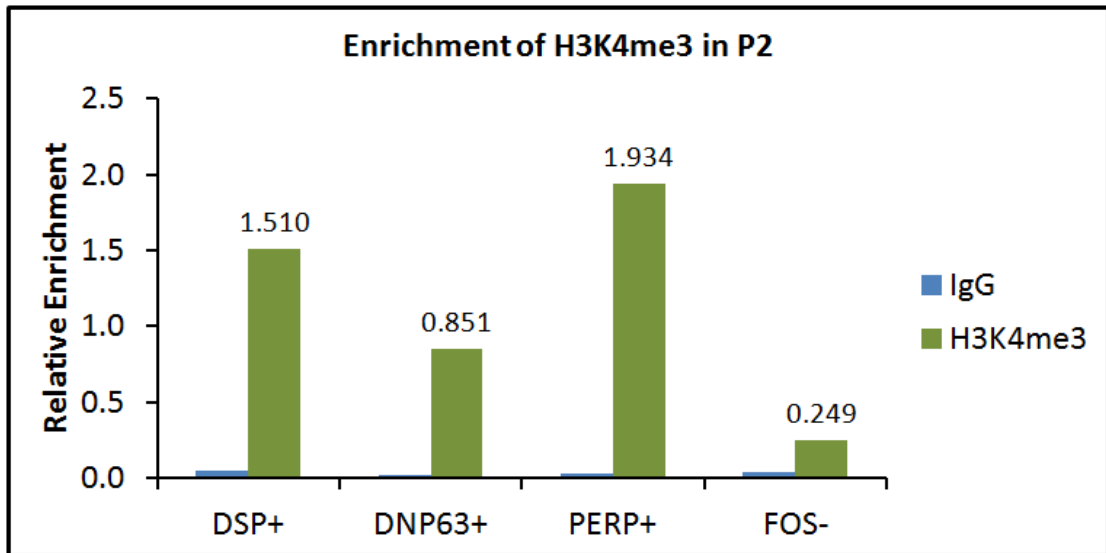
Figure 76: Relative enrichment of H3K4me3 in targeted gene loci related to P63 pathway in patient (P2) ureter tissue. FOS- locus was used as negative enrichment region H3K4me3 binding sites. DNP63+ (ΔNP63+), DSP+ and PERP+ are targeted loci. The IgG was used as ChIP negative control to determine the background signal.

5.3.11 Binding of TP63 Transcription Factor in Cell Line and BEEC Tissue Sample

To confirm our H3K4me3/TP63 and H3K27ac/TP63 intersect data we have performed ChIP analysis on the normal human urothelial cell line as well BEEC patient tissue sample. To do this we have used p63α (H-129) and p63 (H-137) antibodies from Santa-Cruz Biotech. To study the predominant *TP63* isoform, *ΔNP63*, we also used a RR-14 antibody which is specific to *ΔNP63* isoform (Birkaya et al., 2007; Ortt et al., 2008; Romano et al., 2006). Unfortunately, we were not able to purify enough immuno-precipitated DNA to do further analysis on relative enrichment and binding site of TP63 in the genome.

5.4 Discussion

PERP and *DSP* are major components of the P63 Pathway. After *TP63*, these genes are also considered amongst the most prominent candidate genes in BEEC aetiology (Qi et al., 2011). Little information is available on how these genes are involved in BEEC progression, except that Qi and colleagues found over-expression of *PERP* and *DSP* in BEEC patient tissue samples (Qi et al., 2011). In Chapter 4, we screened the exons and promoters of these genes. However, we did not find any sequence variants in our samples that could be linked to BEEC.

In the absence of sequence variants, we next focused our investigation on finding regulatory elements influencing the transcription of these and other possible BEEC candidate genes. Several research groups showed that different kinds of regulatory elements can be marked by particular histone modifications and TF binding (Heintzman et al., 2007; Hon et al., 2009; Wang et al., 2012). With this in mind, we focused our research on genome wide mapping of H3K4me3 and H3K27ac in normal human urothelial cell line. These two histone marks are often found near promoter-TSS, and H3K27ac in particular is found near active regulatory (enhancer) elements (Bae, 2013; ENCODE Project Consortium, 2012; Kimura, 2013; Rivera and Ren, 2013).

ChIP assay using normal human urothelial cell line (SV-HUC-1) with H3K4me3 and H3K27ac showed significant enrichment around our candidate genes *i.e. ΔNP63*, *PERP* and *DSP*. The location of these regions is consistent with 7 cell lines used in ENCODE project (ENCODE Project Consortium, 2012). We also checked the enrichment of H3K4me3 in 2 BEEC patient samples, 1 in bladder mucosa tissue and another in ureter tissue. The result of this investigation showed very low level of fold-enrichment on our targeted loci. The low level may be due to the antibody specificity towards the tissues (Fuchs and Strahl, 2011).

Our genome wide profiling of H3K4me3 in human urothelial cell line showed that this histone mark is highly enriched around promoter and in the exons. Previous studies have shown that H3K4me3 enrichment was found to be associated with the promoter and 5′-coding regions of active genes in yeast (Bernstein et al., 2002; Santos-Rosa et al., 2002) and also in higher eukaryotes (Bernstein et al., 2005; Schneider et al., 2004). The mapping of H3K27ac is also consistent with the finding of Koch and colleagues where they found enrichment of H3ac just downstream of TSS (Koch et al., 2007).

Qi and colleagues identified 162 BEEC candidates gene with differential expression in patient tissue samples. The gene were categorized into 11 different biological network mostly related to the developmental functions, organ morphology, cellular assembly and adhesion (Qi et al., 2011). Our data on H3K4me3 and H3K27ac enrichment near these genes showed a wide distribution across the different functional groups. However, a combined study between enrichment and binding site distribution of these two histone marks near BEEC candidate genes suggests the presence of the possible regulatory elements around the promoter and/or in the exons for H3K4me3, whereas for H3K27ac the regulatory elements are mostly located in the intron regions. Although enrichment of these histone marks is highly concentrated in intergenic region, we found an overall negative enrichment of the modification sites. This finding is in agreement with Koch et. al. (2007) who reported that away from TSSs, the modification sites are enriched with H3K4me1 but relatively depleted in H3K4me3 and H3ac (Koch et al., 2007).

PERP and *DSP* genes, being a major contributor in the p63 pathway, were studied for enrichment of H3K4me3 and H3K27ac to locate possible regulatory elements for these genes. Significant enrichment for both histone marks were found between 1000 to 2000 bp of the TSS in the intron 1 of the both *PERP* and *DSP*. Bioinformatics analysis of these loci revealed the presence of EP300, HDAC2, FOXA1 and FOXA2 transcription

factor binding sites. We also located binding sites for the same transcription factors in the promoter region of both genes. It has been reported in several articles that members of the FOXA subfamily of transcription factors are involved in normal urogenital development and differentiation (Mauney et al., 2010; Oottamasathien et al., 2007; Thomas et al., 2008; Varley et al., 2009). Transcription factor EP300 is important in the processes of cell proliferation and differentiation that regulates transcription via chromatin remodelling (Gayther et al., 2000). Further study of these regions in *PERP* and *DSP* may improve our understanding of the transcription of these genes is regulated.

Qi and colleagues reported the up-regulation of five BEEC candidate genes, namely *CRABP2*, *SC4MOL*, *MYO10*, *PHLDA2* and *WNT5A* (Qi et al., 2011). They have also identified 11 down-regulated genes *i.e. ACTC1, ACTN1, CALD1, CSRP1, DES, FLNA, FLNC, FN1, LPP, TNS1* and *TPM1*. These genes are reported to be expressed in the ventrocuadal mesoderm, the allantoic/umbilical region and/or the hindgut endoderm including cloacal membrane, suggesting the involvement of these genes in the formation of the infra-umbilical region (Qi et al., 2011). The binding profile of H3K4me3 near the most up-regulated gene showed high peak score (>50) in *SC4MOL*, *MYO10, PHLDA2* and *WNT5A* gene loci, listed in Table 16 of this chapter. Mapping of H3K27ac to these up-regulated gene showed peak score<50 as listed in Table 18. This finding is consistent with the finding of Koch et al. (2007).

Methylsterol monooxygenase 1, encoded by *SC4MOL,* is localized to the endoplasmic reticulum membrane and is believed to function in cholesterol biosynthesis. Evidence suggests that methylsterols play an important role in epidermal biology by their influence on cell proliferation, intracellular signalling, vesicular trafficking and immune response (He et al., 2014). Like the bladder, the epidermis consists of stratified

epithelium, so detailed study of *SC4MOL* may reveal valuable information on bladder development. Both H3K4me3 and H3K27ac enrichment with high peak scores was detected in the intron 1 (between 800 to 4000 bp downstream of promoter) of this gene, suggesting that this region is involved in active transcriptional regulation.

The *MYO10* gene encodes a member of the myosin super family and represents an unconventional myosin. Evidence suggests that it involved in regulating cell shape, cell spreading and cell adhesion (Almagro et al., 2010; Hasson et al., 1996; Liu and Cheney, 2012). Epithelial cells have critical functions in normal physiology and disease, and epithelial morphogenesis is required for the proper development of tissues and organs. Liu and co-worker reported the involvement of *MYO10* in junction formation, regulation of paracellular permeability, and epithelial morphogenesis (Liu et al., 2012). A high peak score of H3K4me3 enrichment in intron 1 of this gene suggest active involvement in gene regulation. Therefore, the study of this gene in BEEC may help to determine the molecular mechanism of urothelial development.

PHLDA2, previously known as *TSSC3*, has been shown to be imprinted, with preferential expression from the maternal allele in placenta and liver (Hu et al., 1997; Qian et al., 1997). Genomic imprinting is defined as gamete or zygote specific epigenetic modification involved in parental origin-specific gene expression in somatic cells of the offspring. Lee and Feinberg showed that *PHLDA2* is the first apoptosis-related gene in any species found to be imprinted (Lee and Feinberg, 1998). Muller et al. (2000) stated that this gene is expressed from the maternal allele during normal human development and found to be imprinted in placenta, liver, and foetal tissues. They also showed by inducing a polymorphism in exon 1 that this gene loses its imprinting function (Muller et al., 2000). Our binding profile of H3K4me3 identified high peak score in the exon 1 of this gene, suggesting that this exon is active during

gene transcription. It is possible that disruption of this gene may be responsible for spontaneous BEEC cases, without any familial history.

The promoter region of *WNT5A* showed the highest peaks score among the entire top ranked up-regulated BEEC candidate genes. *WNT5A* is a member of WNT protein family involved in signal transduction (Clark et al., 1993). Abnormal WNT signalling and/or mutations in these genes can cause a range of birth defects and other diseases, including cancer (Andre et al., 2012; Angers and Moon, 2009; Logan and Nusse, 2004). Liu and co-worker showed that both *Wnt5a* and *p63* expression is induced by Gli2 in the cloacal mesoderm, meaning that these genes are regulated by a common up-stream pathway (Liu et al., 2007). There is evidence that WNT protein family is involved in several developmental processes, including regulation of cell fate and patterning during embryogenesis (Gao et al., 2011; Heisenberg et al., 2000; Qian et al., 2007; Rauch et al., 1997). As *WNT5A* mapped into the Chromosome 3 (Clark et al., 1993), it is possible that intra-chromosomal interaction with *TP63* may influence the developmental processes.

Among the entire top ranked down-regulated genes (Qi et al., 2011), we found high H3K4me3 peak score (>50) in *ACTN1*, *CALD1*, *CSRP1*, *FLNC*, *FN1* and *TPM1*. In case of H3K27ac, the peak score is <50 in the entire top ranked down-regulated genes; and we found peak score>25 in the *ACTN1*, *CALD1* and *CSRP1* genes.

In our study we found high H3K4me3 and H3K27ac peak scores between the 1000-4500 bp regions of alpha actinin gene, *ACTN1*, meaning these regions are most active during gene regulation. *ACTN1* is a member of the spectrin gene super family which represents a diverse group of cytoskeletal proteins. It is an actin-binding protein with multiple roles in different cell types (Youssoufian et al., 1990). *ACTN1* is thought to

anchor actin to a variety of intracellular structures which acts as a bundling protein (Kanchanawong et al., 2010). Coding mutations in this gene are associated with a rare congenital disorder characterized by abnormal blood platelets count and morphology (Gueguen et al., 2013; Kunishima et al., 2013). Yang and Glass reported the involvement of *ACTN1* in human glomerular disease by investigating the expression pattern of human glomerular mesangial cells *in vivo* and *in vitro* (Yang and Glass, 2008). Glomerular mesangial cells have been characterized as modified smooth muscle cell. Like smooth muscle in urinary bladder, it contracts or relaxes in response to a number of vasoactive agents (Schlondorff, 1987). Therefore, mutation screening in BEEC patients and analysing the expression pattern of *ACTN1* in bladder development might help determining new insight in the disease aetiology.

We have identified high peak scores in intron 1 and intron 3 of *CSRP1* for H3K4me3 and H3K27ac enrichment, respectively. This gene is a member of the cysteine-rich protein (CSRP) family that also includes a group of LIM domain proteins, also characterized as evolutionarily conserved proteins. This protein family gene was implicated in the regulatory processes important for cytoskeletal remodelling, development and cellular differentiation mediated by protein-protein interactions (Erdel and Weiskirchen, 1998; Liebhaber et al., 1990; Weiskirchen and Gunther, 2003). Qi and colleagues found a strong and distinct expression of *CSRP1, TNS1* and *TPM1* in mid-gestation mouse embryo in the ventrolateral area surrounding the cloacal membrane and also in the basement of allantois including bladder primordia (Qi et al., 2011). This gene is also known to interact with *ACTN1* which is involved in smooth muscle function (Harper et al., 2000). With high peak score of the regulatory histones in the urothelial cell line on the intron of *ACTN1* and *CSRP1,* and their known interaction in the smooth muscle, it is tempting to hypothesize that these two most under-expressed genes in

BEEC patient are involved the pathogenesis of the disease. However, more detailed study is required to confirm this speculation.

We also found high H3K4me3 peak scores in exon 1 of the both *FLNC* and *FN1* genes. *FLNC* is a muscle-specific filamin, which plays a central role in muscle cells, probably by functioning as a large actin-cross-linking protein (Gariboldi et al., 1994; Maestrini et al., 1993). Fibronectins, *FN1*, are involved in cell adhesion and migration processes including embryogenesis, cell motility, wound healing, and maintenance of cell shape (Kornblihtt et al., 1985; Oldberg and Ruoslahti, 1986). Study of these genes in BEEC patient may uncover valuable information related to smooth muscle development and cellular adhesion in maintaining epithelial integrity.

De novo motif scan in H3K4me3 regions identified significant enrichment for IRF2 transcription factor binding sites. IRF2 regulates NFƙB activity via the modulation of NFƙB subcellular localization (Chae et al., 2008). NFƙB is a ubiquitous transcription factor which is found in numerous cell types (DiDonato et al., 2012). It expresses cytokines, chemokines, growth factors, cell adhesion molecules, and some acute phase protein (Baldwin, 1996). Involvement of NFƙB has also been reported in maintaining epithelial integrity (Nenci et al., 2007). NFƙB works as a transcriptional regulator which requires RELA for activation. Chae et al. (2008) showed that IRF2 recruits transcription factors into the nucleus via physical interaction which than facilitate the activation complex. In BEEC murine model, bladder epithelia is developmentally compromised *i.e.* urothelial layer is not stratified compare to wild-type (Cheng et al., 2006). Therefore, it is possible that in BEEC patient the IRF2>RELA/P65>NFƙB signalling is disrupted leading to abnormal bladder development.

De novo motif scan in H3K4me3 and H3K27ac regions identified significant enrichment of the SCL transcription factor, officially known as STIL (SCL/TAL1 interrupting locus). STIL is highly conserved in vertebrate species (Aplan et al., 1991; Collazo-Garcia et al., 1995; Gottgens et al., 2000), and is a ubiquitous-expressed protein in many cell types. It plays an important role in embryonic development as well as in cellular growth and proliferation (Collazo-Garcia et al., 1995; Izraeli et al., 1997; Sun et al., 2014). A number of studies show that SCL/TAL1is required for the Sonic Hedgehog response pathway (Carr et al., 2014; Izraeli et al., 2001; Sun et al., 2014). Interestingly, multiple studies have confirmed that *Shh* is involved in development of the bladder, through the *Shh-Ptc1-Gli2-Bmp4* pathway (Cheng et al., 2008; Haraguchi et al., 2007; Shiroyanagi et al., 2007). It has also been established that SCL acts as a transcriptional regulator, controlling multiple developmental pathways in a variety of organisms, from worms to man (Barton et al., 1999; Garrell and Campuzano, 1991; Harfe et al., 1998; Krause et al., 1990; Murre et al., 1994; Murre et al., 1989; Srivastava et al., 1997). As development of the bladder requires epithelial-mesenchymal interaction (Baskin et al., 1996a; Baskin et al., 1996b) and *Shh* is responsible for inducing mesenchyme development (Haraguchi et al., 2007), it is possible that the SCL transcription factor control the development of bladder by regulating the *SHH* pathway. Detailed study of this transcription in BEEC patient may reveal important insight in to the disease aetiology.

From our overlapping binding data of TP63 with H3K4me3 and H3K27ac near published BEEC candidate genes (Qi et al., 2011), we found six genes present in both H3K4me3/TP63 and H3K27ac/TP63 intersect regions *i.e. DSP*, *PCNA*, *PTBP3*, *RNF144B*, *SEMA3C* and *TGM2*. We have already discussed the involvement of *DSP,* as it is our one of the primary targets in analysing the contribution of p63 pathway in

BEEC pathogenesis. On the other hand, *PTBP3* and *SEMA3C* are particularly interesting because they are known for their regulatory function in cell differentiation and proliferation during developmental processes (Chen et al., 2014; Reidy and Tufro, 2011; Sadvakassova et al., 2009). The presence of TP63 binding sites in these genes made us think that *TP63* might influence the transcriptional regulation of these genes during development.

5.5 Conclusion

In this chapter we identified regulatory elements in some known BEEC candidate genes, with specific attention to our p63 pathway genes *i.e.TP63*, *PERP* and *DSP*. In the previous chapter we have screened the *PERP* and *DSP* promoter and exon in search for potential sequence variant in these genes. However, in the absence of any contributing variants to BEEC aetiology, we concluded that there might be some other loci that are responsible for the regulation of these genes. Our genome wide binding profile of H3K4me3 and H3K27ac, and also the overlapping regions of these histone modifications with TP63 binding showed that intron 1 of *PERP* and *DSP* might be involved in BEEC pathogenesis. These regions have not previously been screened for genetic variants in relation to BEEC. We have also found and discussed some interesting genomic regions near and/or on previously published BEEC candidate genes. These data will be a valuable resource for future BEEC studies, to better determine the role of genetic and epigenetic variation in BEEC pathogenesis.

Chapter 6

Conclusions and Future Directions

Chapter 6: Conclusions and Future Directions

In this study, we have investigated possible genetic and epigenetic factors in the pathogenesis of BEEC. Based on the *TP63* knockout mouse model (Yang et al., 1998; Yang et al., 1999); the only genetic model for BEEC, we previously showed that *p63* is a candidate gene for BEEC pathogenesis (Cheng et al., 2006). To better understand BEEC in humans, sequence variation studies of the promoter region of *TP63* isoforms, *TAP63* and *ΔNP63*, were carried out. Although, *TAP63* promoter region did not reveal any novel variants contributing to the disease (Darling et al., 2013), we found three insertion/deletion polymorphisms in the promoter of *ΔNP63* which are associated with significant risk of BEEC (Wilkins et al., 2012). *ΔNP63* mRNA expression analysis in BEEC patients >1 year of age showed variable expression patterns. However, expression analysis in BEEC infants/neonates showed a significant reduction in *ΔNP63* expression. Such reduction is consistent with the finding of Yang et al. (1998) and also our study on the BEEC murine knock out model (Cheng et al., 2006). Our study of *TP63* coding exons did not reveal any sequence variation in BEEC patients, which also is in agreement with Ching et al. (2010).

In the absence of *TP63* exonic mutations in BEEC, we next focused our research direction on identifying other potential candidate genes for BEEC pathogenesis. Genome wide expression profiling study carried by Qi and co-workers (Qi et al., 2011) showed that *PERP*, *SYNPO2* and genes of WNT pathway may contribute to BEEC aetiology. As with *TP63*, *PERP* is required for maintaining the integrity of the stratified epithelia and is also a direct downstream target of *P63* (Ihrie and Attardi, 2005; Ihrie et al., 2006; Ihrie et al., 2005; Ihrie et al., 2003; Marques et al., 2005; Marques et al.,

2006). Interestingly, *PERP* is a critical component of desmosome, and they function together maintain epithelial integrity (Ihrie and Attardi, 2005; Ihrie et al., 2006; Ihrie et al., 2005). These findings, led us to hypothesize that the TP63>PERP>Desmosome pathway (P63 pathway) might be involved in BEEC pathogenesis (Mahfuz et al., 2013b).

Expression analysis of *PERP* and *DSP* showed significant over-expression of these genes in BEEC patient tissue samples, a similar finding to that reported by Qi et al. (2011). Over-expression of *PERP* and *DSP* may be a compensatory response to the absence, or reduced expression, of *ΔNP63* in BEEC patients (possibly due to a negative feedback loop). We screened *PERP* and *DSP* for sequence variation in their promoters and exons, but did not identify any potential sequence variant that may be contributed to the disease. Mutation screening on the promoter regions of *PERP* also did not reveal any novel variants that may contribute to BEEC. Unfortunately, the high GC ratio in the promoter region of *DSP* rendered it unsuitable for screening with PCR and sequencing. As a part of this study, we found significant *Streptococcus parasanguinis* bacterial contamination in buccal samples. A Streptococcal contamination in buccal cells is not unexpected, because of its high abundance in oral micro flora (Rudney et al., 2005; The Human Microbiome Project Consortium, 2012). Our finding demonstrates that bacterial contamination may be a significant contributor to the total amount of DNA isolated from a buccal swab. This may explain why an apparently sufficient quantity of DNA does not perform as expected in human-specific PCR amplifications (Mahfuz et al., 2013a).

In the absence of contributing sequence variants of *PERP* and *DSP* in BEEC patients, we focused our research in finding regulatory elements that might be involved in the regulation of these genes. To test our hypothesis, we have looked at H3K4me3 and H3K27ac enrichment in a human urothelial cell line, as these histone modifications are associated with regulatory regions (ENCODE Project Consortium, 2012; Kimura, 2013; Rivera and Ren, 2013). The majority of the enriched regions are found in the introns, intergenic regions and promoter-TSS. The highest levels of enrichment are found in promoter-TSS, less enrichment found in introns, and no or negative enrichment in the intergenic region. This matched with the finding of Koch and co-workers who reported that away from TSSs, the modification sites are relatively depleted in H3K4me3 and H3ac histones (Koch et al., 2007). The analysis of our ChIP-seq confirms enrichment of these histone modifications in the intron 1 of the *PERP* and *DSP*. Several articles have shown that *PERP* intron 1 contains a major P53/P63 responsive element and *P63* binds this element to exert its function (Ihrie et al., 2005; Ince et al., 2002; Jheon et al., 2011). The study of Yang and co-workers demonstrated that *DSP* was methylated in the promoter region and/or in intron 1 in cell lines and primary lung tumors, indicating that epigenetic regulation is responsible for *DSP* gene silencing (Yang et al., 2012). Our histone modification enrichment study and published research suggests that intron 1 of both *PERP* and *DSP* are essential for epigenetic regulation in epithelial cells.

De novo motif scan have identified IRF2 and SCL transcription factor in the enriched H3K4me3 and H3K27ac regions. IRF2 recruits RELA in the nucleus to activate NFƙB complex, which in turn expresses cell adhesion molecule and maintain epithelial integrity (Baldwin, 1996; Chae et al., 2008; Nenci et al., 2007). It has been shown in several researches that SCL is required for the Sonic Hedgehog (*SHH*) response

pathway (Carr et al., 2014; Izraeli et al., 2001; Sun et al., 2014). Evidence confirmed that *Shh* induce mesenchymal development (Haraguchi et al., 2007) and development of bladder needs epithelial-mesenchymal interaction (Baskin et al., 1996a; Baskin et al., 1996b). So, it is very tempting to speculate that both IRF2-RELA- NFƙB and SCL-SHH signalling pathway are involved in bladder development and detailed study on this factor on BEEC patient may reveal new insight in disease aetiology.

The identification of H3K4me3 and H3K27ac enrichment that overlapped with TP63 binding sites identified six genes (*DSP*, *PCNA*, *PTBP3*, *RNF144B*, *SEMA3C* and *TGM2)*. Interestingly, these genes are mainly involved in DNA repair and replication, cell death or regulation of developmental processes (Baple et al., 2014; Chen et al., 2014; Conforti et al., 2013; Reidy and Tufro, 2011; Sadvakassova et al., 2009; Sayan et al., 2010), suggesting that these gene may be transactivated by TP63 for their function.

In conclusion, we believe that perturbation of *P63-PERP-DSP* signalling or regulation may render individuals more susceptible to developing BEEC.

Future Directions

We did not find any potential sequence variant in *TAP63* promoter of BEEC patients, but our bioinformatics analysis have identified several potential transcription factor binding sites in the promoter of *TAP63*. Amongst those, FOXA1 and RELA are particularly interesting for their involvement in development and differentiation. FOXA1 is known to be involved in normal urogenital development and differentiation (Mauney et al., 2010; Oottamasathien et al., 2007; Thomas et al., 2008; Varley et al., 2009). Detailed study on this transcription factors with ChIP and ChIP-seq may uncover more precise information on the transcriptional control of *TP63* gene in BEEC.

RELA is responsible for activation of the NFƙB complex (Chen et al., 1999), an ubiquitous transcription factor involved in the control of epithelial integrity (Nenci et al., 2007). Interestingly, our *de novo* motif scan found that IRF2 motif is highly enriched in normal urothelial cell line and this transcription factor is responsible for recruiting RELA to the nucleus during activation of NFƙB complex (Chae et al., 2008). Baldwin and co-worker showed that NFƙB expresses cytokines and cell adhesion molecules (Baldwin, 1996) which is important during epithelial stratification. BEEC murine models shows little or no stratification in bladder urothelia (Cheng et al., 2006) which may also be caused by the disruption of IRF2>RELA>NFƙB signalling pathway. Therefore, expression and mutation analysis of this signalling pathway in BEEC patients could uncover possible mechanisms of disease pathogenesis. Also, a RELA and IRF2 transcription factor binding study on BEEC epigenomes via ChIP and ChIP-seq may reveal valuable information on the disease progression.

In our study we also tried to study the TP63 transcription factor binding. Unfortunately, due to rarity of BEEC, we have restricted access to patient tissue samples. With additional samples, it may be possible to map TP63 binding in the genome, to identify other genes regulated by this transcription factor.

Our study did not find any potential sequence variation in the *PERP* and *DSP* exons contributing to BEEC progression. However, the genome wide binding of H3K4me3 and H3K27ac histones in normal urothelial cell line showed enrichment in the intron 1 of both genes. It was known that *PERP* intron 1 contains a major P63 consensus element (Ihrie et al., 2005; Ihrie et al., 2003) and methylation of *DSP* intron 1 is associated with several cancers (Yang et al., 2012). There has been no sequence

variation analysis of this locus performed in BEEC patients. Mutation screening of this intron 1, and examining the effect of any variants on transcriptional regulation in an urothelial cell line might explain the regulatory mechanism of these genes. Such experiment may also uncover the transactivation efficiency in *PERP* by TP63.

Bioinformatics analysis of the *PERP* and *DSP* intron 1 also identified binding sites for some common transcription factors that were also found in the promoter region of the both genes. Among these factors, members of FOXA subfamily transcription factor (FOXA1 and FOXA2) and EP300 are particularly interesting, because these transcription factors were also found in the promoter region of both *TP63* isoforms. By studying these transcription factors, it should be possible to compare the precise regulation of P63 signalling pathway between BEEC patients and healthy individuals.

With more resources and time it would be interesting to carry out a comparative genome wide mapping of H3K4me3 and H3K27ac, as well as other H3 histone modifications, in normal urothelial cell line, healthy individuals and BEEC patient tissue samples. Such experiments will allow us to locate specific regulatory loci relevant to bladder development and BEEC. To validate the involvement of these loci, expression analysis, mutation screening and reporter assays could be performed to find a potential link with BEEC pathogenesis.

References

References

A.D.A.M. Medical Encyclopedia (2014). Myelomeningocele (PubMed Health).

Aguirre-Lamban, J., Riveiro-Alvarez, R., Garcia-Hoyos, M., Cantalapiedra, D., Avila-Fernandez, A., Villaverde-Montero, C., Trujillo-Tiebas, M.J., Ramos, C., Ayuso, C., Aguirre-Lamban, J., *et al.* (2010). Comparison of High-Resolution Melting Analysis with Denaturing High-Performance Liquid Chromatography for Mutation Scanning in the ABCA4 Gene. Investigative Ophthalmology & Visual Science *51*, 2615-2619.

AlleLogic Biosciences (2004). A Breakthrough in PCR Technology: AllGloTM Fluorogenic Reagents for RT-PCR. In AllGlo Technologu Overview (AlleLogic Biosciences Corp).

Almagro, S., Durmort, C., Chervin-Petinot, A., Heyraud, S., Dubois, M., Lambert, O., Maillefaud, C., Hewat, E., Schaal, J.P., Huber, P., *et al.* (2010). The motor protein myosin-X transports VE-cadherin along filopodia to allow the formation of early endothelial cell-cell contacts. Molecular and cellular biology *30*, 1703-1717.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J Mol Biol *215*, 403-410.

Ambrose, S.S., and O'Brien, D.P., 3rd (1974). Surgical embryology of the exstrophy-epispadias complex. The Surgical clinics of North America *54*, 1379-1390.

American Type Culture Collection (2014). SV-HUC-1 (https://www.atcc.org/: ATCC).

Amiel, J., Bougeard, G., Francannet, C., Raclin, V., Munnich, A., Lyonnet, S., and Frebourg, T. (2001). TP63 gene mutation in ADULT syndrome. European journal of human genetics : EJHG *9*, 642-645.

Andre, P., Wang, Q., Wang, N., Gao, B., Schilit, A., Halford, M.M., Stacker, S.A., Zhang, X., and Yang, Y. (2012). The Wnt coreceptor Ryk regulates Wnt/planar cell polarity by modulating the degradation of the core planar cell polarity component Vangl2. Journal of Biological Chemistry *287*.

Angel, P., and Karin, M. (1991). The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer *1072*, 129-157.

Angers, S., and Moon, R.T. (2009). Proximal events in Wnt signal transduction. Nat Rev Mol Cell Biol *10*, 468-477.

Anonymus (1987). Epidemiology of bladder exstrophy and epispadias: a communication from the International Clearinghouse for Birth Defects Monitoring Systems. Teratology *36*, 221-227.

Aplan, P.D., Lombardi, D.P., and Kirsch, I.R. (1991). Structural characterization of SIL, a gene frequently disrupted in T-cell acute lymphoblastic leukemia. Molecular and cellular biology *11*, 5462-5469.

Arey, L.B. (1965). Developmental Anatomy: A Textbook and Laboratory Manual of Embryology (Saunders).

Arte, S. (2001). Phenotypic and genotypic features of familial hypodontia (University of Helsinki).

Bae, J.B. (2013). Perspectives of international human epigenome consortium. Genomics & informatics *11*, 7-14.

Baldwin, A.S., Jr. (1996). The NF-kappa B and I kappa B proteins: new discoveries and insights. Annual review of immunology *14*, 649-683.

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299-308.

Bani, M., Tezkirecioglu, A.M., Akal, N., and Tuzuner, T. (2010). Ectodermal dysplasia with anodontia: a report of two cases. European journal of dentistry *4*, 215.

Baple, E.L., Chambers, H., Cross, H.E., Fawcett, H., Nakazawa, Y., Chioza, B.A., Harlalka, G.V., Mansour, S., Sreekantan-Nair, A., Patton, M.A.*, et al.* (2014). Hypomorphic PCNA mutation underlies a human DNA repair disorder. J Clin Invest *124*, 3137-3146.

Barbieri, C.E., and Pietenpol, J.A. (2006). p63 and epithelial biology. In Experimental cell research (United States), pp. 695-706.

Barrett, L.W., Fletcher, S., and Wilton, S.D. (2012). Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. Cell Mol Life Sci *69*, 3613-3634.

Barrow, L.L., van Bokhoven, H., Daack-Hirsch, S., Andersen, T., van Beersum, S.E., Gorlin, R., and Murray, J.C. (2002). Analysis of the p63 gene in classical EEC syndrome, related syndromes, and non-syndromic orofacial clefts. J Med Genet *39*, 559-566.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. In Cell (United States), pp. 823-837.

Barton, L.M., Göttgens, B., and Green, A.R. (1999). The stem cell leukaemia (SCL) gene: A critical regulator of haemopoietic and vascular development. International Journal of Biochemistry and Cell Biology *31*, 1193-1207.

Baskin, L.S., Hayward, S.W., Sutherland, R.A., DiSandro, M.J., Thomson, A.A., Goodman, J., and Cunha, G.R. (1996a). Mesenchymal-epithelial interactions in the bladder. World journal of urology *14*, 301-309.

Baskin, L.S., Hayward, S.W., Young, P., and Cunha, G.R. (1996b). Role of mesenchymal-epithelial interactions in normal bladder development. J Urol *156*, 1820-1827.

Battle, M.A., Konopka, G., Parviz, F., Gaggl, A.L., Yang, C., Sladek, F.M., and Duncan, S.A. (2006). Hepatocyte nuclear factor 4alpha orchestrates expression of cell adhesion proteins during the epithelial transformation of the developing liver. Proc Natl Acad Sci U S A *103*, 8419-8424.

Beaudoin, S., Barbet, P., and Bargy, F. (2004). Pelvic development in the rabbit embryo: implications in the organogenesis of bladder exstrophy. Anatomy and embryology *208*, 425-430.

Beaudry, V.G., Ihrie, R.A., Jacobs, S.B., Nguyen, B., Pathak, N., Park, E., and Attardi, L.D. (2010). Loss of the desmosomal component perp impairs wound healing in vivo. Dermatol Res Pract *2010*, 759731.

Beaudry, V.G., Pathak, N., Koster, M.I., and Attardi, L.D. (2009). Differential PERP regulation by TP63 mutants provides insight into AEC pathogenesis. Am J Med Genet A *149A*, 1952-1957.

Beckett, S.M., Laughton, S.J., Pozza, L.D., McCowage, G.B., Marshall, G., Cohn, R.J., Milne, E., and Ashton, L.J. (2008). Buccal swabs and treated cards: methodological considerations for molecular epidemiologic studies examining pediatric populations. American journal of epidemiology *167*, 1260-1267.

Behrens, J., Von Kries, J.P., Kühl, M., Bruhn, L., Wedlich, D., Grosschedl, R., and Birchmeier, W. (1996). Functional interaction of β-catenin with the transcription factor LEF- 1. Nature *382*, 638-642.

Bektas, M., and Rubenstein, D.S. (2009). Perp and pemphigus: a disease of desmosome destabilization. J Invest Dermatol *129*, 1606-1608.

Bell, S.P., Learned, R.M., Jantzen, H.M., and Tjian, R. (1988). Functional cooperativity between transcription factors UBF1 and SL1 mediates human ribosomal RNA synthesis. Science (New York, NY) *241*, 1192-1197.

Bengtsson, M., Karlsson, H.J., Westman, G., and Kubista, M. (2003). A new minor groove binding asymmetric cyanine reporter dye for real-time PCR. Nucleic Acids Res *31*, e45.

Berg, J.M., Tymoczko, J.L., and Stryer, L. (2002). The Greater Complexity of Eukaryotic Genomes Requires Elaborate Mechanisms for Gene Regulation. In Biochemistry (W.H. Freeman), p. 1050.

Berger, I., and Shaul, Y. (1991). Structure and function of human jun-D. Oncogene *6*, 561-566.

Bernstein, B.E., Humphrey, E.L., Erlich, R.L., Schneider, R., Bouman, P., Liu, J.S., Kouzarides, T., and Schreiber, S.L. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes. Proc Natl Acad Sci U S A *99*, 8695-8700.

Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R.*, et al.* (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell *120*, 169-181.

Berrocal, T., Lopez-Pereira, P., Arjonilla, A., and Gutierrez, J. (2002). Anomalies of the distal ureter, bladder, and urethra in children: embryologic, radiologic, and pathologic features. Radiographics : a review publication of the Radiological Society of North America, Inc *22*, 1139-1164.

Bertola, D.R., Kim, C.A., Albano, L.M., Scheffer, H., Meijer, R., and van Bokhoven, H. (2004). Molecular evidence that AEC syndrome and Rapp-Hodgkin syndrome are variable expression of a single genetic disorder. Clinical genetics *66*, 79-80.

Betz, R., Gray, S.G., Ekstrom, C., Larsson, C., and Ekstrom, T.J. (1998). Human histone deacetylase 2, HDAC2 (Human RPD3), is localized to 6q21 by radiation hybrid mapping. Genomics *52*, 245-246.

Beyersmann, D. (2000). Regulation of mammalian gene expression. Exs *89*, 11-28.

Bies, J., Hoffman, B., Amanullah, A., Giese, T., and Wolff, L. (1996). B-Myb prevents growth arrest associated with terminal differentiation of monocytic cells. Oncogene *12*, 355-363.

Bingle, C.D., and Gowan, S. (1996). Molecular cloning of the forkhead transcription factor HNF-3 alpha from a human pulmonary adenocarcinoma cell line. Biochimica et biophysica acta *1307*, 17-20.

Birkaya, B., Ortt, K., and Sinha, S. (2007). Novel in vivo targets of ?Np63 in keratinocytes identified by a modified chromatin immunoprecipitation approach. BMC Mol Biol *8*, 43.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E.*, et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799-816.

Blakstad, O. (2008). Experimental Research (Explorable.com: https://explorable.com/experimental-research).

Blanpain, C., and Fuchs, E. (2007). p63: revving up epithelial stem-cell potential. Nat Cell Biol *9*, 731-733.

Boemers, T.M.L., de Jong, T.P.V.M., Rövekamp, M.H., Bax, N.M.A., and van Gool, J.D. (1994). Covered exstrophy associated with an anorectal malformation: A rare variant of classical bladder exstrophy. Pediatric Surgery International *9*, 438-440.

Bondurand, N., Kuhlbrodt, K., Pingault, V., Enderich, J., Sajus, M., Tommerup, N., Warburg, M., Hennekam, R.C., Read, A.P., Wegner, M., *et al.* (1999). A molecular analysis of the yemenite deaf-blind hypopigmentation syndrome: SOX10 dysfunction causes different neurocristopathies. Human molecular genetics *8*, 1785-1789.

Bonne, S., Gilbert, B., Hatzfeld, M., Chen, X., Green, K.J., and van Roy, F. (2003). Defining desmosomal plakophilin-3 interactions. The Journal of cell biology *161*, 403-416.

Bonzo, J.A., Ferry, C.H., Matsubara, T., Kim, J.H., and Gonzalez, F.J. (2012). Suppression of hepatocyte proliferation by hepatocyte nuclear factor 4alpha in adult mice. The Journal of biological chemistry *287*, 7345-7356.

Bornslaeger, E.A., Corcoran, C.M., Stappenbeck, T.S., and Green, K.J. (1996). Breaking the connection: displacement of the desmosomal plaque protein desmoplakin from cell-cell interfaces disrupts anchorage of intermediate filament bundles and alters intercellular junction assembly. The Journal of cell biology *134*, 985-1001.

Borrmann, C.M., Mertens, C., Schmidt, A., Langbein, L., Kuhn, C., and Franke, W.W. (2000). Molecular diversity of plaques of epithelial-adhering junctions. Annals of the New York Academy of Sciences *915*, 144-150.

Bossis, G., Malnou, C.E., Farras, R., Andermarcher, E., Hipskind, R., Rodriguez, M., Schmidt, D., Muller, S., Jariel-Encontre, I., and Piechaczyk, M. (2005). Down-regulation of c-Fos/c-Jun AP-1 dimer activity by sumoylation. Molecular and cellular biology *25*, 6964-6979.

Bougeard, G., Hadj-Rabia, S., Faivre, L., Sarafan-Vasseur, N., and Frebourg, T. (2003). The Rapp-Hodgkin syndrome results from mutations of the TP63 gene. European journal of human genetics : EJHG *11*, 700-704.

Boyadjiev, S.A., Dodson, J.L., Radford, C.L., Ashrafi, G.H., Beaty, T.H., Mathews, R.I., Broman, K.W., and Gearhart, J.P. (2004). Clinical and molecular characterization of the bladder exstrophy-epispadias complex: analysis of 232 families. BJU Int *94*, 1337-1343.

Boyadjiev, S.A., South, S.T., Radford, C.L., Patel, A., Zhang, G., Hur, D.J., Thomas, G.H., Gearhart, J.P., and Stetten, G. (2005). A reciprocal translocation 46,XY,t(8;9)(p11.2;q13) in a bladder exstrophy patient disrupts CNTNAP3 and presents evidence of a pericentromeric duplication on chromosome 9. Genomics *85*, 622-629.

Brivanlou, A.H., and Darnell, J.E., Jr. (2002). Signal transduction and the control of gene expression. Science (New York, NY) *295*, 813-818.

Bruch, S.W., Adzick, N.S., Goldstein, R.B., and Harrison, M.R. (1996). Challenging the embryogenesis of cloacal exstrophy. J Pediatr Surg *31*, 768-770.

Buermans, H.P., and den Dunnen, J.T. (2014). Next generation sequencing technology: Advances and applications. Biochimica et biophysica acta *1842*, 1932-1941.

Buratowski, S. (1994). The basics of basal transcription by RNA polymerase II. Cell *77*, 1-3.

Burchard, E.G., Silverman, E.K., Rosenwasser, L.J., Borish, L., Yandava, C., Pillari, A., Weiss, S.T., Hasday, J., Lilly, C.M., Ford, J.G., *et al.* (1999). Association between a sequence variant in the IL-4 gene promoter and FEV(1) in asthma. American journal of respiratory and critical care medicine *160*, 919-922.

Cadeddu, J.A., Benson, J.E., Silver, R.I., Lakshmanan, Y., Jeffs, R.D., and Gearhart, J.P. (1997). Spinal abnormalities in classic bladder exstrophy. British journal of urology *79*, 975-978.

Caplin, B.E., Rasmussen, R.P., Bernard, P.S., and Wittwer, C.T. LightCycler hybridization probes.

Carlsson, J., Grahnen, H., Jonsson, G., and Wikner, S. (1970). Establishment of Streptococcus sanguis in the mouths of infants. Archives of oral biology *15*, 1143-1148.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., *et al.* (2005). The transcriptional landscape of the mammalian genome. Science (New York, NY) *309*, 1559-1563.

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C.A., Taylor, M.S., Engstrom, P.G., Frith, M.C., *et al.* (2006). Genome-wide analysis of mammalian promoter architecture and evolution. Nature genetics *38*, 626-635.

Carr, A.L., Sun, L., Lee, E., Li, P., Antonacci, C., Gorbea, E., Finlay, C., and Li, L. (2014). The human oncogene SCL/TAL1 interrupting locus is required for mammalian dopaminergic cell proliferation through the Sonic hedgehog pathway. Cellular Signalling *26*, 306-312.

Carroll, D.K., Carroll, J.S., Leong, C.O., Cheng, F., Brown, M., Mills, A.A., Brugge, J.S., and Ellisen, L.W. (2006). p63 regulates an adhesion programme and cell survival in epithelial cells. Nat Cell Biol *8*, 551-561.

Carroll, T.J., Park, J.S., Hayashi, S., Majumdar, A., and McMahon, A.P. (2005). Wnt9b plays a central role in the regulation of mesenchymal to epithelial transitions underlying organogenesis of the mammalian urogenital system. Dev Cell *9*, 283-292.

Castillo-Martin, M., Domingo-Domenech, J., Karni-Schmidt, O., Matos, T., and Cordon-Cardo, C. (2010). Molecular pathways of urothelial development and bladder tumorigenesis. Urologic Oncology: Seminars and Original Investigations *28*, 401-408.

Caton, A.R., Bloom, A., Druschel, C.M., and Kirby, R.S. (2007). Epidemiology of bladder and cloacal exstrophies in New York State, 1983-1999. Birth Defects Res A Clin Mol Teratol *79*, 781-787.

Celli, J., Duijf, P., Hamel, B.C., Bamshad, M., Kramer, B., Smits, A.P., Newbury-Ecob, R., Hennekam, R.C., Van Buggenhout, G., van Haeringen, A.*, et al.* (1999). Heterozygous germline mutations in the p53 homolog p63 are the cause of EEC syndrome. Cell *99*, 143-153.

Chae, M., Kim, K., Park, S.M., Jang, I.S., Seo, T., Kim, D.M., Kim, I.C., Lee, J.H., and Park, J. (2008). IRF-2 regulates NF-kappaB activity by modulating the subcellular localization of NF-kappaB. Biochem Biophys Res Commun *370*, 519-524.

Chan, I., McGrath, J.A., and Kivirikko, S. (2005). Rapp-Hodgkin syndrome and the tail of p63. Clinical and experimental dermatology *30*, 183-186.

Chandra, V., Huang, P., Potluri, N., Wu, D., Kim, Y., and Rastinejad, F. (2013). Multidomain integration in the structure of the HNF-4alpha nuclear receptor complex. Nature *495*, 394-398.

Chartier, F.L., Bossu, J.P., Laudet, V., Fruchart, J.C., and Laine, B. (1994). Cloning and sequencing of cDNAs encoding the human hepatocyte nuclear factor 4 indicate the presence of two isoforms in human liver. Gene *147*, 269-272.

Chen, B., Zhao, A.G., Shao, J., Mu, X.Y., Jiang, L., and Liu, J.W. (2014). The effects of PTBP3 silencing on the proliferation and differentiation of MKN45 human gastric cancer cells. Life sciences *114*, 29-35.

Chen, F., Castranova, V., Shi, X., and Demers, L.M. (1999). New insights into the role of nuclear factor-kappaB, a ubiquitous transcription factor in the initiation of diseases. Clinical chemistry *45*, 7-17.

Chen, X., Bonne, S., Hatzfeld, M., van Roy, F., and Green, K.J. (2002). Protein binding and functional characterization of plakophilin 2. Evidence for its diverse roles in desmosomes and beta -catenin signaling. The Journal of biological chemistry *277*, 10512-10522.

Cheng, J., Kapranov, P., Drenkow, J., Dike, S., Brubaker, S., Patel, S., Long, J., Stern, D., Tammana, H., Helt, G.*, et al.* (2005). Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. Science (New York, NY) *308*, 1149-1154.

Cheng, W., Jacobs, W.B., Zhang, J.J., Moro, A., Park, J.H., Kushida, M., Qiu, W., Mills, A.A., and Kim, P.C. (2006). DeltaNp63 plays an anti-apoptotic role in ventral bladder development. Development *133*, 4783-4792.

Cheng, W., Yeung, C.K., Ng, Y.K., Zhang, J.R., Hui, C.C., and Kim, P.C. (2008). Sonic Hedgehog mediator Gli2 regulates bladder mesenchymal patterning. J Urol *180*, 1543-1550.

Chiang, C.M., and Roeder, R.G. (1995). Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators. Science (New York, NY) *267*, 531-536.

Chilosi, M., Zamò, A., Brighenti, A., Malpeli, G., Montagna, L., Piccoli, P., Pedron, S., Lestani, M., Inghirami, G., Scarpa, A.*, et al.* (2003). Constitutive expression of DeltaN-p63alpha isoform in human thymus and thymic epithelial tumours. Virchows Arch *443*, 175-183.

Ching, B.J., Wittler, L., Proske, J., Yagnik, G., Qi, L., Draaken, M., Reutter, H., Gearhart, J.P., Ludwig, M., and Boyadjiev, S.A. (2010). p63 (TP73L) a key player in

embryonic urogenital development with significant dysregulation in human bladder exstrophy tissue. Int J Mol Med *26*, 861-867.

Chipail, A., Constantinescu, V., Covic, M., and Angheloni, T. (1976). [Phenotypic and cytogenetic analysis of an unusual malformative syndrome (trisomy 9 p+)]. Revista de pediatrie, obstetrica si ginecologie Pediatria *25*, 201-210.

Chranowska, K.H., Krajewska-Walasek, M., Rump, Z., Wisniewski, L., and Fryns, J.P. (1990). Anodontia as the sole clinical sign of the ectrodactyly-ectodermal dysplasia-cleft lip (EEC) syndrome. Genetic counseling (Geneva, Switzerland) *1*, 67-73.

Clark, C.C., Cohen, I., Eichstetter, I., Cannizzaro, L.A., McPherson, J.D., Wasmuth, J.J., and Iozzo, R.V. (1993). Molecular cloning of the human proto-oncogene Wnt-5A and mapping of the gene (WNT5A) to chromosome 3p14-p21. Genomics *18*, 249-260.

Collas, P. (2010). The current state of chromatin immunoprecipitation. Molecular biotechnology *45*, 87-100.

Collazo-Garcia, N., Scherer, P., and Aplan, P.D. (1995). Cloning and characterization of a murine SIL gene. Genomics *30*, 506-513.

Conforti, F., Yang, A.L., Piro, M.C., Mellone, M., Terrinoni, A., Candi, E., Tucci, P., Thomas, G.J., Knight, R.A., Melino, G.*, et al.* (2013). PIR2/Rnf144B regulates epithelial homeostasis by mediating degradation of p21WAF1 and p63. Oncogene *32*, 4758-4765.

Cook, P.R. (2003). Nongenic transcription, gene regulation and action at a distance. J Cell Sci *116*, 4483-4491.

Cotney, J., Leng, J., Oh, S., DeMare, L.E., Reilly, S.K., Gerstein, M.B., and Noonan, J.P. (2012). Chromatin state signatures associated with tissue-specific gene expression and enhancer activity in the embryonic limb. Genome research *22*, 1069-1080.

Cui, C.P., Zhang, Y., Yu, Y.T., Li, Y.Z., Geng, Y., and Zhao, S.F. (2004). [Expression profiling of UBF, a novel member of UBC family]. Zhongguo ying yong sheng li xue za zhi = Zhongguo yingyong shenglixue zazhi = Chinese journal of applied physiology *20*, 66-69.

Darling, T., Mahfuz, I., White, S.J., and Cheng, W. (2013). No TAP63 promoter mutation is detected in bladder exstrophy-epispadias complex patients. J Pediatr Surg *48*, 2393-2400.

Das, P.M., Ramachandran, K., vanWert, J., and Singal, R. (2004). Chromatin immunoprecipitation assay. BioTechniques *37*, 961-969.

de Boer, C.M., Eini, R., Gillis, A.M., Stoop, H., Looijenga, L.H., and White, S.J. (2012). DICER1 RNase IIIb domain mutations are infrequent in testicular germ cell tumours. BMC Res Notes *5*, 569.

de Laat, W., and Grosveld, F. (2007). Inter-chromosomal gene regulation in the mammalian cell nucleus. Current opinion in genetics & development *17*, 456-464.

de Riese, W., and Warmbold, H. (1986). Adenocarcinoma in extrophy of the bladder. A case report and review of the literature. International urology and nephrology *18*, 159-162.

DeGraff, D.J., Clark, P.E., Cates, J.M., Yamashita, H., Robinson, V.L., Yu, X., Smolkin, M.E., Chang, S.S., Cookson, M.S., Herrick, M.K., *et al.* (2012). Loss of the urothelial differentiation marker FOXA1 is associated with high grade, late stage bladder cancer and increased tumor proliferation. PloS one *7*, e36669.

DeGregori, J., Leone, G., Miron, A., Jakoi, L., and Nevins, J.R. (1997). Distinct roles for E2F proteins in cell growth control and apoptosis. Proc Natl Acad Sci U S A *94*, 7245-7250.

Delva, E., Tucker, D.K., and Kowalczyk, A.P. (2009). The desmosome. Cold Spring Harb Perspect Biol *1*, a002543.

Di Como, C.J., Urist, M.J., Babayan, I., Drobnjak, M., Hedvat, C.V., Teruya-Feldstein, J., Pohar, K., Hoos, A., and Cordon-Cardo, C. (2002). p63 expression profiles in human normal and tumor tissues. Clinical cancer research : an official journal of the American Association for Cancer Research *8*, 494-501.

di Fiore, M.S.H., and Schmidt, I.G. (1981). Atlas of human histology (Lea & Febiger).

Dianzani, I., Garelli, E., Gustavsson, P., Carando, A., Gustafsson, B., Dahl, N., and Anneren, G. (2003). Rapp-Hodgkin and AEC syndromes due to a new frameshift mutation in the TP63 gene. J Med Genet *40*, e133.

DiDonato, J.A., Mercurio, F., and Karin, M. (2012). NF-kappaB and the link between inflammation and cancer. Immunological reviews *246*, 379-400.

Dominguez-Sola, D., Ying, C.Y., Grandori, C., Ruggiero, L., Chen, B., Li, M., Galloway, D.A., Gu, W., Gautier, J., and Dalla-Favera, R. (2007). Non-transcriptional control of DNA replication by c-Myc. Nature *448*, 445-451.

Draaken, M., Mughal, S.S., Pennimpede, T., Wolter, S., Wittler, L., Ebert, A.K., Rosch, W., Stein, R., Bartels, E., Schmidt, D., *et al.* (2013). Isolated bladder exstrophy associated with a de novo 0.9 Mb microduplication on chromosome 19p13.12. Birth Defects Res A Clin Mol Teratol.

Draaken, M., Reutter, H., Schramm, C., Bartels, E., Boemers, T.M., Ebert, A.K., Rösch, W., Schröder, A., Stein, R., Moebus, S., *et al.* (2010). Microduplications at 22q11.21 are associated with non-syndromic classic bladder exstrophy. Eur J Med Genet *53*, 55-60.

Drissen, R., Palstra, R.J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S., and de Laat, W. (2004). The active spatial organization of the beta-globin locus requires the transcription factor EKLF. Genes Dev *18*, 2485-2490.

Dudek, R.W., and Fix, J.D. (1998). Embryology, 2nd edn (Williams & Wilkins).

Dwight, Z., Palais, R., and Wittwer, C.T. (2011). uMELT: prediction of high-resolution melting curves and dynamic melting profiles of PCR products in a rich web application. Bioinformatics (Oxford, England) *27*, 1019-1020.

Ebert, A.K., Reutter, H., Ludwig, M., and Rösch, W.H. (2009). The exstrophy-epispadias complex. Orphanet J Rare Dis *4*, 23.

Eckner, R., Ewen, M.E., Newsome, D., Gerdes, M., DeCaprio, J.A., Lawrence, J.B., and Livingston, D.M. (1994). Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. Genes Dev *8*, 869-884.

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Erali, M., and Wittwer, C.T. (2010). High resolution melting analysis for gene scanning. In Methods (United States: 2010 Elsevier Inc), pp. 250-261.

Erdel, M., and Weiskirchen, R. (1998). Assignment1 of CSRP1 encoding the LIM domain protein CRP1, to human chromosome 1q32 by fluorescence in situ hybridization. Cytogenetics and cell genetics *83*, 10-11.

Ernst, J., and Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. Nature biotechnology *28*, 817-825.

Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., *et al.* (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. Nature *473*, 43-49.

Facchini, V., Gadducci, A., Colombi, L., Ceccarelli, P., Lamacchia, M., Simi, U., and Fioretti, P. (1987). Carcinoma developing in bladder exstrophy. Case report. British journal of obstetrics and gynaecology *94*, 795-797.

Ferretti, E., Li, B., Zewdu, R., Wells, V., Hebert, J.M., Karner, C., Anderson, M.J., Williams, T., Dixon, J., Dixon, M.J., *et al.* (2011). A conserved Pbx-Wnt-p63-Irf6 regulatory module controls face morphogenesis by promoting epithelial apoptosis. Dev Cell *21*, 627-641.

Filippova, G.N., Fagerlie, S., Klenova, E.M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P.E., Collins, S.J., and Lobanenkov, V.V. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. Molecular and cellular biology *16*, 2802-2813.

Flores, E.R., Tsai, K.Y., Crowley, D., Sengupta, S., Yang, A., McKeon, F., and Jacks, T. (2002). p63 and p73 are required for p53-dependent apoptosis in response to DNA damage. Nature *416*, 560-564.

Frank, B., Rigas, S.H., Bermejo, J.L., Wiestler, M., Wagner, K., Hemminki, K., Reed, M.W., Sutter, C., Wappenschmidt, B., Balasubramanian, S.P., *et al.* (2008). The CASP8 -652 6N del promoter polymorphism and breast cancer risk: a multicenter study. Breast cancer research and treatment *111*, 139-144.

Fraser, P. (2006). Transcriptional control thrown for a loop. Current opinion in genetics & development *16*, 490-495.

Freeman, B., Powell, J., Ball, D., Hill, L., Craig, I., and Plomin, R. (1997). DNA by mail: an inexpensive and noninvasive method for collecting DNA samples from widely dispersed populations. Behavior genetics *27*, 251-257.

Fuchs, E., and Raghavan, S. (2002). Getting under the skin of epidermal morphogenesis. In Nature reviews Genetics (England), pp. 199-209.

Fuchs, S.M., and Strahl, B.D. (2011). Antibody recognition of histone post-translational modifications: emerging issues and future prospects. Epigenomics *3*, 247-249.

Gallicano, G.I., Kouklis, P., Bauer, C., Yin, M., Vasioukhin, V., Degenstein, L., and Fuchs, E. (1998). Desmoplakin is required early in development for assembly of desmosomes and cytoskeletal linkage. The Journal of cell biology *143*, 2009-2022.

Gao, B., Song, H., Bishop, K., Elliot, G., Garrett, L., English, M.A., Andre, P., Robinson, J., Sood, R., Minami, Y.*, et al.* (2011). Wnt signaling gradients establish planar cell polarity by inducing Vangl2 phosphorylation through Ror2. Dev Cell *20*, 163-176.

Gariboldi, M., Maestrini, E., Canzian, F., Manenti, G., De Gregorio, L., Rivella, S., Chatterjee, A., Herman, G.E., Archidiacono, N., Antonacci, R.*, et al.* (1994). Comparative mapping of the actin-binding protein 280 genes in human and mouse. Genomics *21*, 428-430.

Garrell, J., and Campuzano, S. (1991). The helix-loop-helix domain: A common motif for bristles, muscles and sex. BioEssays : news and reviews in molecular, cellular and developmental biology *13*, 493-498.

Garrod, D., and Chidgey, M. (2008). Desmosome structure, composition and function. In Biochimica et biophysica acta (Netherlands), pp. 572-587.

Garrod, D.R., Merritt, A.J., and Nie, Z. (2002). Desmosomal cadherins. Curr Opin Cell Biol *14*, 537-545.

Gayther, S.A., Batley, S.J., Linger, L., Bannister, A., Thorpe, K., Chin, S.F., Daigo, Y., Russell, P., Wilson, A., Sowter, H.M.*, et al.* (2000). Mutations truncating the EP300 acetylase in human cancers. Nature genetics *24*, 300-303.

Gearhart, J.P. (2001). The bladder exstrophy-epispadias-cloacal exstrophy complex. In Pediatric Urology, J.P. Gearhart, and R.C. Rink, eds. (Mouriquand PDE. Philadelphia: W. B. Saunders Co), pp. 511-546.

Gearhart, J.P., and Jeffs, R.D. (1989). State-of-the-art reconstructive surgery for bladder exstrophy at the Johns Hopkins Hospital. Am J Dis Child *143*, 1475-1478.

Gegonne, A., Devaiah, B.N., and Singer, D.S. (2013). TAF7: traffic controller in transcription initiation. Transcription *4*, 29-33.

Gegonne, A., Tai, X., Zhang, J., Wu, G., Zhu, J., Yoshimoto, A., Hanson, J., Cultraro, C., Chen, Q.R., Guinter, T.*, et al.* (2012). The general transcription factor TAF7 is essential for embryonic development but not essential for the survival or differentiation of mature T cells. Molecular and cellular biology *32*, 1984-1997.

Gegonne, A., Weissman, J.D., Zhou, M., Brady, J.N., and Singer, D.S. (2006). TAF7: a possible transcription initiation check-point regulator. Proc Natl Acad Sci U S A *103*, 602-607.

Gerstein, M.B., Kundaje, A., Hariharan, M., Landt, S.G., Yan, K.K., Cheng, C., Mu, X.J., Khurana, E., Rozowsky, J., Alexander, R.*, et al.* (2012). Architecture of the human regulatory network derived from ENCODE data. Nature *489*, 91-100.

Getsios, S., Huen, A.C., and Green, K.J. (2004). Working out the strength and flexibility of desmosomes. Nat Rev Mol Cell Biol *5*, 271-281.

Geyer, P.K., Green, M.M., and Corces, V.G. (1990). Tissue-specific transcriptional enhancers may act in trans on the gene located in the homologous chromosome: the molecular basis of transvection in Drosophila. The EMBO journal *9*, 2247-2256.

Ghassibe, M., Bayet, B., Revencu, N., Verellen-Dumoulin, C., Gillerot, Y., Vanwijck, R., and Vikkula, M. (2005). Interferon regulatory factor-6: a gene predisposing to isolated cleft lip with or without cleft palate in the Belgian population. European journal of human genetics : EJHG *13*, 1239-1242.

Ghioni, P., D'Alessandra, Y., Mansueto, G., Jaffray, E., Hay, R.T., La Mantia, G., and Guerrini, L. (2005). The protein stability and transcriptional activity of p63alpha are regulated by SUMO-1 conjugation. Cell Cycle *4*, 183-190.

Ginsberg, D., Vairo, G., Chittenden, T., Xiao, Z.X., Xu, G., Wydner, K.L., DeCaprio, J.A., Lawrence, J.B., and Livingston, D.M. (1994). E2F-4, a new member of the E2F transcription factor family, interacts with p107. Genes Dev *8*, 2665-2679.

Gokhman, D., Livyatan, I., Sailaja, B.S., Melcer, S., and Meshorer, E. (2013). Multilayered chromatin analysis reveals E2f, Smad and Zfx as transcriptional regulators of histones. Nat Struct Mol Biol *20*, 119-126.

Gottgens, B., Barton, L.M., Gilbert, J.G., Bench, A.J., Sanchez, M.J., Bahn, S., Mistry, S., Grafham, D., McMurray, A., Vaudin, M.*, et al.* (2000). Analysis of vertebrate SCL loci identifies conserved enhancers. Nature biotechnology *18*, 181-186.

Green, K.J., and Gaudry, C.A. (2000). Are desmosomes more than tethers for intermediate filaments? Nat Rev Mol Cell Biol *1*, 208-216.

Griffiths, A.J.F., Miller, J.H., Suzuki, D.T., Lewontin, R.C., and Gelbart, W.M. (2003). Transcription: an overview of gene regulation in eukaryotes. In An Introduction to Genetic Analysis (Macmillan Higher Education).

Gueguen, P., Rouault, K., Chen, J.M., Raguenes, O., Fichou, Y., Hardy, E., Gobin, E., Pan-Petesch, B., Kerbiriou, M., Trouve, P., *et al.* (2013). A missense mutation in the alpha-actinin 1 gene (ACTN1) is the cause of autosomal dominant macrothrombocytopenia in a large French family. PloS one *8*, e74728.

Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. Nat Struct Mol Biol *11*, 394-403.

Haiman, C.A., Garcia, R.R., Kolonel, L.N., Henderson, B.E., Wu, A.H., and Le Marchand, L. (2008). A promoter polymorphism in the CASP8 gene is not associated with cancer risk. Nature genetics *40*, 259-260; author reply 260-251.

Haraguchi, R., Motoyama, J., Sasaki, H., Satoh, Y., Miyagawa, S., Nakagata, N., Moon, A., and Yamada, G. (2007). Molecular analysis of coordinated bladder and urogenital organ formation by Hedgehog signaling. In Development (England), pp. 525-533.

Harfe, B.D., Gomes, A.V., Kenyon, C., Liu, J., Krause, M., and Fire, A. (1998). Analysis of a Caenorhabditis elegans twist homolog identifies conserved and divergent aspects of mesodermal patterning. Genes and Development *12*, 2623-2635.

Harper, B.D., Beckerle, M.C., and Pomies, P. (2000). Fine mapping of the alpha-actinin binding site within cysteine-rich protein. The Biochemical journal *350 Pt 1*, 269-274.

Hasson, T., Skowron, J.F., Gilbert, D.J., Avraham, K.B., Perry, W.L., Bement, W.M., Anderson, B.L., Sherr, E.H., Chen, Z.Y., Greene, L.A., *et al.* (1996). Mapping of unconventional myosins in mouse and human. Genomics *36*, 431-439.

Hatzfeld, M. (1999). The armadillo family of structural proteins. International review of cytology *186*, 179-224.

Hatzfeld, M., Green, K.J., and Sauter, H. (2003). Targeting of p0071 to desmosomes and adherens junctions is mediated by different protein domains. J Cell Sci *116*, 1219-1233.

Hay, R.J., and Wells, R.S. (1976). The syndrome of ankyloblepharon, ectodermal defects and cleft lip and palate: an autosomal dominant condition. The British journal of dermatology *94*, 277-289.

He, M., Smith, L.D., Chang, R., Li, X., and Vockley, J. (2014). The role of sterol-C4-methyl oxidase in epidermal biology. Biochimica et biophysica acta *1841*, 331-335.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A.*, et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nature genetics *39*, 311-318.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Heisenberg, C.P., Tada, M., Rauch, G.J., Saude, L., Concha, M.L., Geisler, R., Stemple, D.L., Smith, J.C., and Wilson, S.W. (2000). Silberblick/Wnt11 mediates convergent extension movements during zebrafish gastrulation. Nature *405*, 76-81.

Higuchi, R., Dollinger, G., Walsh, P.S., and Griffith, R. (1992). Simultaneous amplification and detection of specific DNA sequences. Bio/technology (Nature Publishing Company) *10*, 413-417.

Hirata, H., Hinoda, Y., Ueno, K., Shahryari, V., Tabatabai, Z.L., and Dahiya, R. (2012). MicroRNA-1826 targets VEGFC, beta-catenin (CTNNB1) and MEK1 (MAP2K1) in human bladder cancer. Carcinogenesis *33*, 41-48.

Holland, P.M., Abramson, R.D., Watson, R., and Gelfand, D.H. (1991). Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase. Proc Natl Acad Sci U S A *88*, 7276-7280.

Holthofer, B., Windoffer, R., Troyanovsky, S., and Leube, R.E. (2007). Structure and function of desmosomes. International review of cytology *264*, 65-163.

Hon, G.C., Hawkins, R.D., and Ren, B. (2009). Predictive chromatin signatures in the mammalian genome. Human molecular genetics *18*, R195-201.

Hood, L.E., Hunkapiller, M.W., Hunkapiller, T.J., and Smith, L.M. (1992). Automated DNA sequencing technique (Google Patents).

Hu, R.J., Lee, M.P., Connors, T.D., Johnson, L.A., Burn, T.C., Su, K., Landes, G.M., and Feinberg, A.P. (1997). A 2.5-Mb transcript map of a tumor-suppressing subchromosomal transferable fragment from 11p15.5, and isolation and sequence analysis of three novel genes. Genomics *46*, 9-17.

Huang, L., Tang, Z., Li, D., Wang, G., Huang, K., Zhou, J., Li, Y., Gao, B., and Chen, J. (2014). [Ketamine induces apoptosis of human uroepithelial SV-HUC-1 cells]. Zhong nan da xue xue bao Yi xue ban = Journal of Central South University Medical sciences *39*, 703-707.

Huang, Y.P., Wu, G., Guo, Z., Osada, M., Fomenkov, T., Park, H.L., Trink, B., Sidransky, D., Fomenkov, A., and Ratovitski, E.A. (2004). Altered sumoylation of p63alpha contributes to the split-hand/foot malformation phenotype. Cell Cycle *3*, 1587-1596.

Huber, W.E., Price, E.R., Widlund, H.R., Du, J., Davis, I.J., Wegner, M., and Fisher, D.E. (2003). A tissue-restricted cAMP transcriptional response: SOX10 modulates alpha-melanocyte-stimulating hormone-triggered expression of microphthalmia-associated transcription factor in melanocytes. The Journal of biological chemistry *278*, 45224-45230.

Huen, A.C., Park, J.K., Godsel, L.M., Chen, X., Bannon, L.J., Amargo, E.V., Hudson, T.Y., Mongiu, A.K., Leigh, I.M., Kelsell, D.P., *et al.* (2002). Intermediate filament–membrane attachments function synergistically with actin-dependent contacts to regulate intercellular adhesive strength. The Journal of cell biology *159*, 1005-1017.

Husmann, D.A., and Vandersteen, D.R. (1999). Anatomy of the cloacal esxtrophy. In The epispadias-exstrophy complex, J.P. Gearhard, and R. Matthews, eds. (New York: Kluwer Academic/Plenum), pp. 199-206.

Ianakiev, P., Kilpatrick, M.W., Toudjarska, I., Basel, D., Beighton, P., and Tsipouras, P. (2000). Split-hand/split-foot malformation is caused by mutations in the p63 gene on 3q27. Am J Hum Genet *67*, 59-66.

Ihrie, R.A., and Attardi, L.D. (2005). A new Perp in the lineup: linking p63 and desmosomal adhesion. Cell Cycle *4*, 873-876.

Ihrie, R.A., Bronson, R.T., and Attardi, L.D. (2006). Adult mice lacking the p53/p63 target gene Perp are not predisposed to spontaneous tumorigenesis but display features of ectodermal dysplasia syndromes. Cell Death Differ *13*, 1614-1618.

Ihrie, R.A., Marques, M.R., Nguyen, B.T., Horner, J.S., Papazoglu, C., Bronson, R.T., Mills, A.A., and Attardi, L.D. (2005). Perp is a p63-regulated gene essential for epithelial integrity. Cell *120*, 843-856.

Ihrie, R.A., Reczek, E., Horner, J.S., Khachatrian, L., Sage, J., Jacks, T., and Attardi, L.D. (2003). Perp is a mediator of p53-dependent apoptosis in diverse cell types. Curr Biol *13*, 1985-1990.

Ince, T.A., Cviko, A.P., Quade, B.J., Yang, A., McKeon, F.D., Mutter, G.L., and Crum, C.P. (2002). p63 Coordinates anogenital modeling and epithelial cell differentiation in the developing female urogenital tract. Am J Pathol *161*, 1111-1117.

Izraeli, S., Colaizzo-Anas, T., Bertness, V.L., Mani, K., Aplan, P.D., and Kirsch, I.R. (1997). Expression of the SIL gene is correlated with growth induction and cellular proliferation. Cell Growth and Differentiation *8*, 1171-1179.

Izraeli, S., Lowe, L.A., Bertness, V.L., Campaner, S., Hahn, H., Kirsch, I.R., and Kuehn, M.R. (2001). Genetic evidence that Sil is required for the Sonic Hedgehog response pathway. Genesis (New York, NY : 2000) *31*, 72-77.

Jamora, C., DasGupta, R., Kocieniewski, P., and Fuchs, E. (2003). Links between signal transduction, transcription and adhesion in epithelial bud development. Nature *422*, 317-322.

Jenkinson, H.F., and Lamont, R.J. (1997). Streptococcal adhesion and colonization. Critical reviews in oral biology and medicine : an official publication of the American Association of Oral Biologists *8*, 175-200.

Jheon, A.H., Mostowfi, P., Snead, M.L., Ihrie, R.A., Sone, E., Pramparo, T., Attardi, L.D., and Klein, O.D. (2011). PERP regulates enamel formation via effects on cell-cell adhesion and gene expression. J Cell Sci *124*, 745-754.

Jin, B., Li, Y., and Robertson, K.D. (2011). DNA methylation: superior or subordinate in the epigenetic hierarchy? Genes & cancer *2*, 607-617.

Jing, H., Vakoc, C.R., Ying, L., Mandat, S., Wang, H., Zheng, X., and Blobel, G.A. (2008). Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. Mol Cell *29*, 232-242.

Johnson, D.S., Mortazavi, A., Myers, R.M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. In Science (New York, NY) (United States), pp. 1497-1502.

Johnson, P.F., and McKnight, S.L. (1989). Eukaryotic transcriptional regulatory proteins. Annual review of biochemistry *58*, 799-839.

Jonkman, M.F., Pasmooij, A.M., Pasmans, S.G., van den Berg, M.P., Ter Horst, H.J., Timmer, A., and Pas, H.H. (2005). Loss of desmoplakin tail causes lethal acantholytic epidermolysis bullosa. Am J Hum Genet *77*, 653-660.

Joulin, V., Bories, D., Eleouet, J.F., Labastie, M.C., Chretien, S., Mattei, M.G., and Romeo, P.H. (1991). A T-cell specific TCR delta DNA binding protein is a member of the human GATA family. The EMBO journal *10*, 1809-1816.

Jumlongras, D., Bei, M., Stimson, J.M., Wang, W.F., DePalma, S.R., Seidman, C.E., Felbor, U., Maas, R., Seidman, J.G., and Olsen, B.R. (2001). A nonsense mutation in MSX1 causes Witkop syndrome. Am J Hum Genet *69*, 67-74.

Juven-Gershon, T., Hsu, J.Y., Theisen, J.W., and Kadonaga, J.T. (2008). The RNA polymerase II core promoter - the gateway to transcription. Curr Opin Cell Biol *20*, 253-259.

Kadauke, S., and Blobel, G.A. (2009). Chromatin loops in gene regulation. Biochimica et biophysica acta *1789*, 17-25.

Kadonaga, J.T. (2012). Perspectives on the RNA polymerase II core promoter. Wiley interdisciplinary reviews Developmental biology *1*, 40-51.

Kaestner, K.H., Hiemisch, H., Luckow, B., and Schutz, G. (1994). The HNF-3 gene family of transcription factors in mice: gene structure, cDNA sequence, and mRNA distribution. Genomics *20*, 377-385.

Kanchanawong, P., Shtengel, G., Pasapera, A.M., Ramko, E.B., Davidson, M.W., Hess, H.F., and Waterman, C.M. (2010). Nanoscale architecture of integrin-based cell adhesions. Nature *468*, 580-584.

Kang, S., Graham, J.M., Jr., Olney, A.H., and Biesecker, L.G. (1997). GLI3 frameshift mutations cause autosomal dominant Pallister-Hall syndrome. Nature genetics *15*, 266-268.

Kannu, P., Savarirayan, R., Ozoemena, L., White, S.M., and McGrath, J.A. (2006). Rapp-Hodgkin ectodermal dysplasia syndrome: the clinical and molecular overlap with Hay-Wells syndrome. Am J Med Genet A *140*, 887-891.

Kantaputra, P.N., Hamada, T., Kumchai, T., and McGrath, J.A. (2003). Heterozygous mutation in the SAM domain of p63 underlies Rapp-Hodgkin ectodermal dysplasia. Journal of dental research *82*, 433-437.

Kao, C.C., Lieberman, P.M., Schmidt, M.C., Zhou, Q., Pei, R., and Berk, A.J. (1990). Cloning of a transcriptionally active human TATA binding factor. Science (New York, NY) *248*, 1646-1650.

Karin, M. (1990). Too many transcription factors: positive and negative interactions. The New biologist *2*, 126-131.

Karin, M., Liu, Z., and Zandi, E. (1997). AP-1 function and regulation. Curr Opin Cell Biol *9*, 240-246.

Karni-Schmidt, O., Castillo-Martin, M., Shen, T.H., Gladoun, N., Domingo-Domenech, J., Sanchez-Carbayo, M., Li, Y., Lowe, S., Prives, C., and Cordon-Cardo, C. (2011). Distinct expression profiles of p63 variants during urothelial development and bladder cancer progression. Am J Pathol *178*, 1350-1360.

Kent, W.J. (2002). BLAT--the BLAST-like alignment tool. Genome research *12*, 656-664.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome research *12*, 996-1006.

Kimura, H. (2013). Histone modifications for human epigenome analysis. J Hum Genet *58*, 439-445.

Kljuic, A., Bazzi, H., Sundberg, J.P., Martinez-Mir, A., O'Shaughnessy, R., Mahoney, M.G., Levy, M., Montagutelli, X., Ahmad, W., Aita, V.M., *et al.* (2003). Desmoglein 4 in hair follicle differentiation and epidermal adhesion: evidence from inherited hypotrichosis and acquired pemphigus vulgaris. Cell *113*, 249-260.

Knudsen, K.A., Soler, A.P., Johnson, K.R., and Wheelock, M.J. (1995). Interaction of α-actinin with the cadherin/catenin cell-cell adhesion complex via α-catenin. Journal of Cell Biology *130*, 67-77.

Koch, C.M., Andrews, R.M., Flicek, P., Dillon, S.C., Karaoz, U., Clelland, G.K., Wilcox, S., Beare, D.M., Fowler, J.C., Couttet, P., *et al.* (2007). The landscape of histone modifications across 1% of the human genome in five human cell lines. Genome research *17*, 691-707.

Koch, P.J., Mahoney, M.G., Ishikawa, H., Pulkkinen, L., Uitto, J., Shultz, L., Murphy, G.F., Whitaker-Menezes, D., and Stanley, J.R. (1997). Targeted disruption of the pemphigus vulgaris antigen (desmoglein 3) gene in mice causes loss of keratinocyte cell adhesion with a phenotype similar to pemphigus vulgaris. The Journal of cell biology *137*, 1091-1102.

Konsman, J.P., and Blomqvist, A. (2005). Forebrain patterns of c-Fos and FosB induction during cancer-associated anorexia-cachexia in rat. The European journal of neuroscience *21*, 2752-2766.

Kornberg, R.D., and Lorch, Y. (1991). Irresistible force meets immovable object: transcription and the nucleosome. Cell *67*, 833-836.

Kornblihtt, A.R., Umezawa, K., Vibe-Pedersen, K., and Baralle, F.E. (1985). Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene. The EMBO journal *4*, 1755-1759.

Kosaki, R., Fukuhara, Y., Kosuga, M., Okuyama, T., Kawashima, N., Honna, T., Ueoka, K., and Kosaki, K. (2005). OEIS complex with del(3)(q12.2q13.2). Am J Med Genet A *135*, 224-226.

Koster, M.I., Kim, S., Mills, A.A., DeMayo, F.J., and Roop, D.R. (2004). p63 is the molecular switch for initiation of an epithelial stratification program. Genes Dev *18*, 126-131.

Koster, M.I., and Roop, D.R. (2004). Transgenic mouse models provide new insights into the role of p63 in epidermal development. Cell Cycle *3*, 411-413.

Kouwenhoven, E.N., van Heeringen, S.J., Tena, J.J., Oti, M., Dutilh, B.E., Alonso, M.E., de la Calle-Mustienes, E., Smeenk, L., Rinne, T., Parsaulian, L., *et al.* (2010). Genome-wide profiling of p63 DNA-binding sites identifies an element that regulates gene expression during limb development in the 7q21 SHFM1 locus. PLoS Genet *6*, e1001065.

Kowalczyk, A.P., Stappenbeck, T.S., Parry, D.A., Palka, H.L., Virata, M.L., Bornslaeger, E.A., Nilles, L.A., and Green, K.J. (1994). Structure and function of desmosomal transmembrane core and plaque molecules. Biophys Chem *50*, 97-112.

Krause, M., Fire, A., Harrison, S.W., Priess, J., and Weintraub, H. (1990). CeMyoD accumulation defines the body wall muscle cell fate during C. elegans embryogenesis. Cell *63*, 907-919.

Kubista, M., Andrade, J.M., Bengtsson, M., Forootan, A., Jonak, J., Lind, K., Sindelka, R., Sjoback, R., Sjogreen, B., Strombom, L., *et al.* (2006). The real-time polymerase chain reaction. In Mol Aspects Med (England), pp. 95-125.

Kuhn, R.M., Haussler, D., and Kent, W.J. (2013). The UCSC genome browser and associated tools. Briefings in bioinformatics *14*, 144-161.

Kulozik, A.E., Bellan-Koch, A., Bail, S., Kohne, E., and Kleihauer, E. (1991). Thalassemia intermedia: moderate reduction of beta globin gene transcriptional activity by a novel mutation of the proximal CACCC promoter element. Blood *77*, 2054-2058.

Kunishima, S., Okuno, Y., Yoshida, K., Shiraishi, Y., Sanada, M., Muramatsu, H., Chiba, K., Tanaka, H., Miyazaki, K., Sakai, M.*, et al.* (2013). ACTN1 mutations cause congenital macrothrombocytopenia. Am J Hum Genet *92*, 431-438.

Langman, J., and Sadler, T.W. (1995). Langman's medical embryology, 7 edn (Williams & Wilkins).

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods *9*, 357-359.

Latchman, D.S. (1997). Transcription factors: an overview. The international journal of biochemistry & cell biology *29*, 1305-1312.

Laurikkala, J., Mikkola, M.L., James, M., Tummers, M., Mills, A.A., and Thesleff, I. (2006). p63 regulates multiple signalling pathways required for ectodermal organogenesis and differentiation. Development *133*, 1553-1563.

Lee, M.P., and Feinberg, A.P. (1998). Genomic imprinting of a human apoptosis gene homologue, TSSC3. Cancer Res *58*, 1052-1056.

Lee, T.I., and Young, R.A. (2000). Transcription of eukaryotic protein-coding genes. Annual review of genetics *34*, 77-137.

Leonard, N.J., and Tomkins, D.J. (2002). Diploid/tetraploid/t(1;6) mosaicism in a 17-year-old female with hypomelanosis of Ito, multiple congenital anomalies, and body asymmetry. Am J Med Genet *112*, 86-90.

Leoyklang, P., Siriwan, P., and Shotelersuk, V. (2006). A mutation of the p63 gene in non-syndromic cleft lip. J Med Genet *43*, e28.

Li, Q., Luan, G., Guo, Q., and Liang, J. (2002). A new class of homogeneous nucleic acid probes based on specific displacement hybridization. Nucleic Acids Res *30*, E5.

Lidral, A.C., Romitti, P.A., Basart, A.M., Doetschman, T., Leysens, N.J., Daack-Hirsch, S., Semina, E.V., Johnson, L.R., Machida, J., Burds, A.*, et al.* (1998). Association of MSX1 and TGFB3 with nonsyndromic clefting in humans. Am J Hum Genet *63*, 557-568.

Liebhaber, S.A., Emery, J.G., Urbanek, M., Wang, X.K., and Cooke, N.E. (1990). Characterization of a human cDNA encoding a widely expressed and highly conserved cysteine-rich protein with an unusual zinc-finger motif. Nucleic Acids Res *18*, 3871-3879.

Lin, D., Fiscella, M., O'Connor, P.M., Jackman, J., Chen, M., Luo, L.L., Sala, A., Travali, S., Appella, E., and Mercer, W.E. (1994). Constitutive expression of B-myb can bypass p53-induced Waf1/Cip1-mediated G1 arrest. Proc Natl Acad Sci U S A *91*, 10079-10083.

Lin, H.J., Ndiforchu, F., and Patell, S. (1993). Exstrophy of the cloaca in a 47,XXX child: review of genitourinary malformations in triple-X patients. Am J Med Genet *45*, 761-763.

Ling, J.Q., Li, T., Hu, J.F., Vu, T.H., Chen, H.L., Qiu, X.W., Cherry, A.M., and Hoffman, A.R. (2006). CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science (New York, NY) *312*, 269-272.

Liu, G., Moro, A., Zhang, J.J., Cheng, W., Qiu, W., and Kim, P.C. (2007). The role of Shh transcription activator Gli2 in chick cloacal development. Dev Biol *303*, 448-460.

Liu, K.C., and Cheney, R.E. (2012). Myosins in cell junctions. Bioarchitecture *2*, 158-170.

Liu, K.C., Jacobs, D.T., Dunn, B.D., Fanning, A.S., and Cheney, R.E. (2012). Myosin-X functions in polarized epithelial cells. Molecular biology of the cell *23*, 1675-1687.

Liu, T. (2014). Use model-based Analysis of ChIP-Seq (MACS) to analyze short reads generated by sequencing protein-DNA interactions in embryonic stem cells. Methods in molecular biology (Clifton, NJ) *1150*, 81-95.

Liu, X.L., Xiao, K., Xue, B., Yang, D., Lei, Z., Shan, Y., and Zhang, H.T. (2013). Dual role of TGFBR3 in bladder cancer. Oncology reports *30*, 1301-1308.

Liu, Z., and Garrard, W.T. (2005). Long-range interactions between three transcriptional enhancers, active Vkappa gene promoters, and a 3' boundary sequence spanning 46 kilobases. Molecular and cellular biology *25*, 3220-3231.

Logan, C.Y., and Nusse, R. (2004). The Wnt signaling pathway in development and disease. Annual review of cell and developmental biology *20*, 781-810.

Lohse, M., Bolger, A.M., Nagel, A., Fernie, A.R., Lunn, J.E., Stitt, M., and Usadel, B. (2012). RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Res *40*, W622-627.

Lomvardas, S., Barnea, G., Pisapia, D.J., Mendelsohn, M., Kirkland, J., and Axel, R. (2006). Interchromosomal interactions and olfactory receptor choice. Cell *126*, 403-413.

Lowik, M.M., Groenen, P.J., Levtchenko, E.N., Monnens, L.A., and van den Heuvel, L.P. (2009). Molecular genetic analysis of podocyte genes in focal segmental glomerulosclerosis--a review. European journal of pediatrics *168*, 1291-1304.

Ludwig, M., Ruschendorf, F., Saar, K., Hubner, N., Siekmann, L., Boyadjiev, S.A., and Reutter, H. (2009). Genome-wide linkage scan for bladder exstrophy-epispadias complex. Birth Defects Res A Clin Mol Teratol *85*, 174-178.

Ludwig, M., Utsch, B., and Reutter, H. (2005). [Genetic and molecular biological aspects of the bladder exstrophy-epispadias complex (BEEC)]. Der Urologe Ausg A *44*, 1037-1038, 1040-1034.

Maestrini, E., Patrosso, C., Mancini, M., Rivella, S., Rocchi, M., Repetto, M., Villa, A., Frattini, A., Zoppe, M., Vezzoni, P.*, et al.* (1993). Mapping of two genes encoding isoforms of the actin binding protein ABP-280, a dystrophin like protein, to Xq28 and to chromosome 7. Human molecular genetics *2*, 761-766.

Mahfuz, I., Cheng, W., and White, S.J. (2013a). Identification of Streptococcus parasanguinis DNA contamination in human buccal DNA samples. BMC Res Notes *6*, 481.

Mahfuz, I., Darling, T., Wilkins, S., White, S., and Cheng, W. (2013b). New insights into the pathogenesis of bladder exstrophy-epispadias complex. J Pediatr Urol *9*, 996-1005.

Manner, J., and Kluth, D. (2003). A chicken model to study the embryology of cloacal exstrophy. J Pediatr Surg *38*, 678-681.

Mardis, E.R. (2007). ChIP-seq: welcome to the new frontier. Nat Meth *4*, 613-614.

Mardis, E.R. (2008). Next-generation DNA sequencing methods. Annual review of genomics and human genetics *9*, 387-402.

Marguerat, S., Wilhelm, B.T., and Bahler, J. (2008). Next-generation sequencing: applications beyond genomes. Biochemical Society transactions *36*, 1091-1096.

Marques, M.R., Horner, J.S., Ihrie, R.A., Bronson, R.T., and Attardi, L.D. (2005). Mice lacking the p53/p63 target gene Perp are resistant to papilloma development. Cancer Res *65*, 6551-6556.

Marques, M.R., Ihrie, R.A., Horner, J.S., and Attardi, L.D. (2006). The requirement for perp in postnatal viability and epithelial integrity reflects an intrinsic role in stratified epithelia. J Invest Dermatol *126*, 69-73.

Marshall, V., and Muecke, E. (1968). Congenital Abnormalities of the Bladder. In Malformations (Springer Berlin Heidelberg), pp. 165-223.

Martin, M., Rodriguez, K., Sanchez-Sauco, M., Zambudio-Carmona, G., and Ortega-Garcia, J.A. (2009). Household exposure to pesticides and bladder exstrophy in a newborn baby boy: a case report and review of the literature. Journal of medical case reports *3*, 6626.

Matera, A.G., Wu, W., Imai, H., O'Keefe, C.L., and Chan, E.K. (1997). Molecular cloning of the RNA polymerase I transcription factor hUBF/NOR-90 (UBTF) gene and localization to 17q21.3 by fluorescence in situ hybridization and radiation hybrid mapping. Genomics *41*, 135-138.

Mather, J.P., and Roberts, P.E. (1998). Standard Cell Culture Techniques. In Introduction to Cell and Tissue Culture (Springer US), pp. 63-87.

Matsui, M., Tokuhara, M., Konuma, Y., Nomura, N., and Ishizaki, R. (1990). Isolation of human fos-related genes and their expression during monocyte-macrophage differentiation. Oncogene *5*, 249-255.

Mauney, J.R., Ramachandran, A., Yu, R.N., Daley, G.Q., Adam, R.M., and Estrada, C.R. (2010). All-trans retinoic acid directs urothelial specification of murine embryonic stem cells via GATA4/6 signaling mechanisms. PloS one *5*, e11513.

McGrath, J.A., Duijf, P.H., Doetsch, V., Irvine, A.D., de Waal, R., Vanmolkot, K.R., Wessagowit, V., Kelly, A., Atherton, D.J., Griffiths, W.A.*, et al.* (2001). Hay-Wells syndrome is caused by heterozygous missense mutations in the SAM domain of p63. Human molecular genetics *10*, 221-229.

McKeon, F. (2004). p63 and the epithelial stem cell: more than status quo? Genes Dev *18*, 465-469.

MedlinePlus (2014). Ureterocele (National Institute of Health).

Mendelsohn, C. (2004). Functional obstruction: the renal pelvis rules. J Clin Invest *113*, 957-959.

Metzker, M.L. (2010). Sequencing technologies - the next generation. Nature reviews Genetics *11*, 31-46.

Miceli-Richard, C., Gestermann, N., Ittah, M., Comets, E., Loiseau, P., Puechal, X., Hachulla, E., Gottenberg, J.E., Lebon, P., Becquemont, L.*, et al.* (2009). The CGGGG insertion/deletion polymorphism of the IRF5 promoter is a strong risk factor for primary Sjogren's syndrome. Arthritis and rheumatism *60*, 1991-1997.

Mikles, D.C., Schuchardt, B.J., Bhat, V., McDonald, C.B., and Farooq, A. (2014). Role of promoter DNA sequence variations on the binding of EGR1 transcription factor. Archives of biochemistry and biophysics *549*, 1-11.

Milatovich, A., Travis, A., Grosschedl, R., and Francke, U. (1991). Gene for lymphoid enhancer-binding factor 1 (LEF1) mapped to human chromosome 4 (q23-q25) and mouse chromosome 3 near Egf. Genomics *11*, 1040-1048.

Mildenberger, H., Kluth, D., and Dziuba, M. (1988). Embryology of bladder exstrophy. J Pediatr Surg *23*, 166-170.

Mills, A.A., Zheng, B., Wang, X.J., Vogel, H., Roop, D.R., and Bradley, A. (1999). p63 is a p53 homologue required for limb and epidermal morphogenesis. Nature *398*, 708-713.

Mincheva, A., Lichter, P., Schutz, G., and Kaestner, K.H. (1997). Assignment of the human genes for hepatocyte nuclear factor 3-alpha, -beta, and -gamma (HNF3A, HNF3B, HNF3G) to 14q12-q13, 20p11, and 19q13.2-q13.4. Genomics *39*, 417-419.

Mishra, A. (2007). Renal agenesis: report of an interesting case. The British journal of radiology *80*, e167-169.

Mita, K., Tsuji, H., Morimyo, M., Takahashi, E., Nenoi, M., Ichimura, S., Yamauchi, M., Hongo, E., and Hayashi, A. (1995). The human gene encoding the largest subunit of RNA polymerase II. Gene *159*, 285-286.

Mitchell, L.E., Adzick, N.S., Melchionne, J., Pasquariello, P.S., Sutton, L.N., and Whitehead, A.S. (2004). Spina bifida. Lancet *364*, 1885-1895.

Mitchell, P.J., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science (New York, NY) *245*, 371-378.

Monje, P., Hernandez-Losa, J., Lyons, R.J., Castellone, M.D., and Gutkind, J.S. (2005). Regulation of the transcriptional activity of c-Fos by ERK. A novel role for the prolyl isomerase PIN1. The Journal of biological chemistry *280*, 35081-35084.

Moore, K.L., and Persaud, T.V.N. (2003). The developing human: clinically oriented embryology, 7th edn (University of Michigan: Saunders).

Moretti, G., Mazzaglia, E., D'Anieri, A., Merlino, V., Magaudda, L., Mondello, M.R., Santoro, G., Vaccaro, M., and Albanese, A. (1995). Epidermolysis bullosa junctionalis associated with urinary bladder exstrophy: a case report. Pediatric dermatology *12*, 239-241.

Morozova, O., and Marra, M.A. (2008). Applications of next-generation sequencing technologies in functional genomics. Genomics *92*, 255-264.

Morrisey, E.E., Ip, H.S., Lu, M.M., and Parmacek, M.S. (1996). GATA-6: a zinc finger transcription factor that is expressed in multiple cell lineages derived from lateral mesoderm. Dev Biol *177*, 309-322.

Morrisey, E.E., Tang, Z., Sigrist, K., Lu, M.M., Jiang, F., Ip, H.S., and Parmacek, M.S. (1998). GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. Genes Dev *12*, 3579-3590.

Muecke, E.C. (1964). THE ROLE OF THE CLOACAL MEMBRANE IN EXSTROPHY: THE FIRST SUCCESSFUL. The Journal of urology *92*, 659-667.

Mueller, H., and Franke, W.W. (1983). Biochemical and immunological characterization of desmoplakins I and II, the major polypeptides of the desmosomal plaque. J Mol Biol *163*, 647-671.

Muller, F., Demeny, M.A., and Tora, L. (2007). New problems in RNA polymerase II transcription initiation: matching the diversity of core promoters with a variety of promoter recognition factors. The Journal of biological chemistry *282*, 14685-14689.

Muller, F.J., Laurent, L.C., Kostka, D., Ulitsky, I., Williams, R., Lu, C., Park, I.H., Rao, M.S., Shamir, R., Schwartz, P.H.*, et al.* (2008). Regulatory networks define phenotypic classes of human stem cell lines. Nature *455*, 401-405.

Muller, S., van den Boom, D., Zirkel, D., Koster, H., Berthold, F., Schwab, M., Westphal, M., and Zumkeller, W. (2000). Retention of imprinting of the human apoptosis-related gene TSSC3 in human brain tumors. Human molecular genetics *9*, 757-763.

Muller, T., Bain, G., Wang, X., and Papkoff, J. (2002). Regulation of epithelial cell migration and tumor formation by beta-catenin signaling. Experimental cell research *280*, 119-133.

Mundel, P., Heid, H.W., Mundel, T.M., Kruger, M., Reiser, J., and Kriz, W. (1997). Synaptopodin: an actin-associated protein in telencephalic dendrites and renal podocytes. The Journal of cell biology *139*, 193-204.

Muroya, K., Hasegawa, T., Ito, Y., Nagai, T., Isotani, H., Iwata, Y., Yamamoto, K., Fujimoto, S., Seishu, S., Fukushima, Y.*, et al.* (2001). GATA3 abnormalities and the phenotypic spectrum of HDR syndrome. J Med Genet *38*, 374-380.

Murre, C., Bain, G., van Dijk, M.A., Engel, I., Furnari, B.A., Massari, M.E., Matthews, J.R., Quong, M.W., Rivera, R.R., and Stuiver, M.H. (1994). Structure and function of helix-loop-helix proteins. BBA - Gene Structure and Expression *1218*, 129-135.

Murre, C., McCaw, P.S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell *56*, 777-783.

Nakamura, K., Jeong, S.Y., Uchihara, T., Anno, M., Nagashima, K., Nagashima, T., Ikeda, S., Tsuji, S., and Kanazawa, I. (2001). SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. Human molecular genetics *10*, 1441-1448.

Nakamura, Y., Tsiairis, C.D., Ozbek, S., and Holstein, T.W. (2011). Autoregulatory and repressive inputs localize Hydra Wnt3 to the head organizer. Proc Natl Acad Sci U S A *108*, 9137-9142.

Nazar, R.N. (2004). Ribosomal RNA processing and ribosome biogenesis in eukaryotes. IUBMB life *56*, 457-465.

Nelson, C.P., Dunn, R.L., and Wei, J.T. (2005). Contemporary epidemiology of bladder exstrophy in the United States. J Urol *173*, 1728-1731.

Nenci, A., Becker, C., Wullaert, A., Gareus, R., van Loo, G., Danese, S., Huth, M., Nikolaev, A., Neufert, C., Madison, B*., et al.* (2007). Epithelial NEMO links innate immunity to chronic intestinal inflammation. Nature *446*, 557-561.

Ni, C., Ye, Y., Wang, M., Qian, H., Song, Z., Jia, X., and Zhou, J. (2009). A six-nucleotide insertion-deletion polymorphism in the CASP8 promoter is associated with risk of coal workers' pneumoconiosis. Journal of toxicology and environmental health Part A *72*, 712-716.

Ni, Z., Olsen, J.B., Guo, X., Zhong, G., Ruan, E.D., Marcon, E., Young, P., Guo, H., Li, J., Moffat, J*., et al.* (2011). Control of the RNA polymerase II phosphorylation state in promoter regions by CTD interaction domain-containing proteins RPRD1A and RPRD1B. Transcription *2*, 237-242.

Nicholls, G., and Duffy, P.G. (1998). Anatomical correction of the exstrophy-epispadias complex: analysis of 34 patients. British journal of urology *82*, 865-869.

Niemann, S., Zhao, C., Pascu, F., Stahl, U., Aulepp, U., Niswander, L., Weber, J.L., and Muller, U. (2004). Homozygous WNT3 mutation causes tetra-amelia in a large consanguineous family. Am J Hum Genet *74*, 558-563.

Nikolov, D.B., and Burley, S.K. (1997). RNA polymerase II transcription initiation: a structural view. Proc Natl Acad Sci U S A *94*, 15-22.

Noben-Trauth, K., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Sonoda, G., Testa, J.R., and Klempnauer, K.H. (1996). Mybl2 (Bmyb) maps to mouse chromosome 2 and human chromosome 20q 13.1. Genomics *35*, 610-612.

Nolan, G.P., Ghosh, S., Liou, H.C., Tempst, P., and Baltimore, D. (1991). DNA binding and I kappa B inhibition of the cloned p65 subunit of NF-kappa B, a rel-related polypeptide. Cell *64*, 961-969.

Nolis, I.K., McKay, D.J., Mantouvalou, E., Lomvardas, S., Merika, M., and Thanos, D. (2009). Transcription factors mediate long-range enhancer-promoter interactions. Proc Natl Acad Sci U S A *106*, 20222-20227.

Nomura, N., Ide, M., Sasamoto, S., Matsui, M., Date, T., and Ishizaki, R. (1990). Isolation of human cDNA clones of jun-related genes, jun-B and jun-D. Nucleic Acids Res *18*, 3047-3048.

Nowrousian, M. (2010). Next-Generation Sequencing Techniques for Eukaryotic Microorganisms: Sequencing-Based Solutions to Biological Problems▽. Eukaryot Cell *9*, 1300-1310.

Oldberg, A., and Ruoslahti, E. (1986). Evolution of the fibronectin gene. Exon structure of cell attachment domain. The Journal of biological chemistry *261*, 2113-2116.

Ong, C.-T., and Corces, V.G. (2011). Enhancer function: new insights into the regulation of tissue-specific gene expression. Nature reviews Genetics *12*, 283-293.

Oottamasathien, S., Wang, Y., Williams, K., Franco, O.E., Wills, M.L., Thomas, J.C., Saba, K., Sharif-Afshar, A.R., Makari, J.H., Bhowmick, N.A.*, et al.* (2007). Directed differentiation of embryonic stem cells into bladder tissue. Dev Biol *304*, 556-566.

Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. Genes Dev *10*, 2657-2683.

Ortt, K., Raveh, E., Gat, U., and Sinha, S. (2008). A chromatin immunoprecipitation screen in mouse keratinocytes reveals Runx1 as a direct transcriptional target of DeltaNp63. Journal of cellular biochemistry *104*, 1204-1219.

Panov, K.I., Friedrich, J.K., Russell, J., and Zomerdijk, J.C. (2006). UBF activates RNA polymerase I transcription by stimulating promoter escape. The EMBO journal *25*, 3310-3322.

Papetti, M., and Augenlicht, L.H. (2011). MYBL2, a link between proliferation and differentiation in maturing colon epithelial cells. J Cell Physiol *226*, 785-791.

Parviz, F., Li, J., Kaestner, K.H., and Duncan, S.A. (2002). Generation of a conditionally null allele of hnf4alpha. Genesis (New York, NY : 2000) *32*, 130-133.

Parviz, F., Matullo, C., Garrison, W.D., Savatski, L., Adamson, J.W., Ning, G., Kaestner, K.H., Rossi, J.M., Zaret, K.S., and Duncan, S.A. (2003). Hepatocyte nuclear factor 4alpha controls the development of a hepatic epithelium and liver morphogenesis. Nature genetics *34*, 292-296.

Patten, B.M., and Barry, A. (1952). The genesis of exstrophy of the bladder and epispadias. The American journal of anatomy *90*, 35-57.

Payne, A.S., Yan, A.C., Ilyas, E., Li, W., Seykora, J.T., Young, T.L., Pawel, B.R., Honig, P.J., Camacho, J., Imaizumi, S., *et al.* (2005). Two novel TP63 mutations associated with the ankyloblepharon, ectodermal defects, and cleft lip and palate syndrome: a skin fragility phenotype. Archives of dermatology *141*, 1567-1573.

Peifer, M., and Polakis, P. (2000). Wnt signaling in oncogenesis and embryogenesis - A look outside the nucleus. Science (New York, NY) *287*, 1606-1609.

Persson, H., and Leder, P. (1984). Nuclear localization and DNA binding properties of a protein expressed by human c-myc oncogene. Science (New York, NY) *225*, 718-721.

Peukert, K., Staller, P., Schneider, A., Carmichael, G., Hanel, F., and Eilers, M. (1997). An alternative pathway for gene regulation by Myc. The EMBO journal *16*, 5672-5686.

Pingault, V., Bondurand, N., Kuhlbrodt, K., Goerich, D.E., Prehu, M.O., Puliti, A., Herbarth, B., Hermans-Borgmeyer, I., Legius, E., Matthijs, G., *et al.* (1998). SOX10 mutations in patients with Waardenburg-Hirschsprung disease. Nature genetics *18*, 171-173.

Pittman, A.M., Broderick, P., Sullivan, K., Fielding, S., Webb, E., Penegar, S., Tomlinson, I., and Houlston, R.S. (2008). CASP8 variants D302H and -652 6N ins/del do not influence the risk of colorectal cancer in the United Kingdom population. British journal of cancer *98*, 1434-1436.

Polakis, P. (2000). Wnt signaling and cancer. Genes and Development *14*, 1837-1851.

Polakova, K.M., Lopotova, T., Klamova, H., and Moravcova, J. (2008). High-resolution melt curve analysis: initial screening for mutations in BCR-ABL kinase domain. Leukemia research *32*, 1236-1243.

Pollard, K.S., Hubisz, M.J., Rosenbloom, K.R., and Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. Genome research *20*, 110-121.

Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. Nature biotechnology *28*, 1057-1068.

Ptashne, M., and Gann, A. (1997). Transcriptional activation by recruitment. Nature *386*, 569-577.

Qi, L., Chen, K., Hur, D.J., Yagnik, G., Lakshmanan, Y., Kotch, L.E., Ashrafi, G.H., Martinez-Murillo, F., Kowalski, J., Naydenov, C., *et al.* (2011). Genome-wide expression profiling of urinary bladder implicates desmosomal and cytoskeletal dysregulation in the bladder exstrophy-epispadias complex. Int J Mol Med *27*, 755-765.

Qian, D., Jones, C., Rzadzinska, A., Mark, S., Zhang, X., Steel, K.P., Dai, X., and Chen, P. (2007). Wnt5a functions in planar cell polarity regulation in mice. Dev Biol *306*, 121-133.

Qian, N., Frank, D., O'Keefe, D., Dao, D., Zhao, L., Yuan, L., Wang, Q., Keating, M., Walsh, C., and Tycko, B. (1997). The IPL gene on chromosome 11p15.5 is imprinted in humans and mice and is similar to TDAG51, implicated in Fas expression and apoptosis. Human molecular genetics *6*, 2021-2029.

Qiao, Y., Zhang, Z., Huang, W., Pang, S., Xing, Q., and Yan, B. (2014). Two functional sequence variants of the GATA6 gene promoter in patients with indirect inguinal hernia. Gene *547*, 86-90.

Quiroz-Guerrero, J., Badillo, M., Munoz, N., Anaya, J., Rico, G., and Maldonado-Valadez, R. (2009). Bladder augmentation in a young adult female exstrophy patient with associated omphalocele: an extremely unusual case. J Pediatr Urol *5*, 330-332.

Raboch, J. (1975). Incidence of hypospadia and epispadia in chromatin-positive men. Andrologia *7*, 237-239.

Radiopaedia.org (2014a). Ectopic ureter H. Knipe, and A.A. Rabou, eds. (Radiopaedia).

Radiopaedia.org (2014b). Renal agenesis, A. Goel, and R. Esedov, eds.

Randhawa, G.S., Bell, D.W., Testa, J.R., and Feinberg, A.P. (1998). Identification and mapping of human histone acetylation modifier gene homologues. Genomics *51*, 262-269.

Rando, O.J. (2012). Combinatorial complexity in chromatin structure and function: revisiting the histone code. Current opinion in genetics & development *22*, 148-155.

Raschella, G., Negroni, A., Sala, A., Pucci, S., Romeo, A., and Calabretta, B. (1995). Requirement of b-myb function for survival and differentiative potential of human neuroblastoma cells. The Journal of biological chemistry *270*, 8540-8545.

Rauch, G.J., Hammerschmidt, M., Blader, P., Schauerte, H.E., Strahle, U., Ingham, P.W., McMahon, A.P., and Haffter, P. (1997). Wnt5 is required for tail formation in the zebrafish embryo. Cold Spring Harbor symposia on quantitative biology *62*, 227-234.

Reidy, K., and Tufro, A. (2011). Semaphorins in kidney development and disease: modulators of ureteric bud branching, vascular morphogenesis, and podocyte-endothelial crosstalk. Pediatric nephrology (Berlin, Germany) *26*, 1407-1412.

Reiner, W.G., and Gearhart, J.P. (2004). Discordant sexual identity in some genetic males with cloacal exstrophy assigned to female sex at birth. N Engl J Med *350*, 333-341.

Reutter, H., Boyadjiev, S.A., Gambhir, L., Ebert, A.K., Rösch, W.H., Stein, R., Schröder, A., Boemers, T.M., Bartels, E., Vogt, H., *et al.* (2011). Phenotype Severity in the Bladder Exstrophy-Epispadias Complex: Analysis of Genetic and Nongenetic Contributing Factors in 441 Families from North America and Europe. J Pediatr.

Reutter, H., Draaken, M., Pennimpede, T., Wittler, L., Brockschmidt, F.F., Ebert, A.K., Bartels, E., Rosch, W., Boemers, T.M., Hirsch, K.*, et al.* (2014). Genome-wide association study and mouse expression data identify a highly conserved 32 kb intergenic region between WNT3 and WNT9b as possible susceptibility locus for isolated classic exstrophy of the bladder. Human molecular genetics, 1-9.

Reutter, H., Hoischen, A., Ludwig, M., Stein, R., Radlwimmer, B., Engels, H., Wolffenbuttel, K.P., and Weber, R.G. (2007a). Genome-wide analysis for micro-aberrations in familial exstrophy of the bladder using array-based comparative genomic hybridization. BJU Int *100*, 646-650.

Reutter, H., Qi, L., Gearhart, J.P., Boemers, T., Ebert, A.K., Rösch, W., Ludwig, M., and Boyadjiev, S.A. (2007b). Concordance analyses of twins with bladder exstrophy-epispadias complex suggest genetic etiology. Am J Med Genet A *143A*, 2751-2756.

Reutter, H., Ruschendorf, F., Mattheisen, M., Draaken, M., Bartels, E., Hubner, N., Hoffmann, P., Payabvash, S., Saar, K., Nothen, M.M., *et al.* (2010). Evidence for linkage of the bladder exstrophy-epispadias complex on chromosome 4q31.21-22 and 19q13.31-41 from a consanguineous Iranian family. Birth Defects Res A Clin Mol Teratol *88*, 757-761.

Ricketts, R.R., Woodard, J.R., Zwiren, G.T., Andrews, H.G., and Broecker, B.H. (1991). Modern treatment of cloacal exstrophy. J Pediatr Surg *26*, 444-448; discussion 448-450.

Rimm, D.L., Koslov, E.R., Kebriaei, P., Cianci, C.D., and Morrow, J.S. (1995). α1(E)-catenin is an actin-binding and -bundling protein mediating the attachment of F-actin to the membrane adhesion complex. Proceedings of the National Academy of Sciences of the United States of America *92*, 8813-8817.

Rinne, T., Brunner, H.G., and van Bokhoven, H. (2007). p63-associated disorders. Cell Cycle *6*, 262-268.

Rinne, T., Hamel, B., van Bokhoven, H., and Brunner, H.G. (2006). Pattern of p63 mutations and their phenotypes--update. Am J Med Genet A *140*, 1396-1406.

Rivera, C.M., and Ren, B. (2013). Mapping human epigenomes. Cell *155*, 39-55.

Robertson, K.D. (2005). DNA methylation and human disease. Nature reviews Genetics *6*, 597-610.

Roeder, R.G. (1996). The role of general initiation factors in transcription by RNA polymerase II. Trends in biochemical sciences *21*, 327-335.

Romano, R.A., Birkaya, B., and Sinha, S. (2006). Defining the regulatory elements in the proximal promoter of DeltaNp63 in keratinocytes: Potential roles for Sp1/Sp3, NF-Y, and p63. J Invest Dermatol *126*, 1469-1479.

Rubio, E.D., Reiss, D.J., Welcsh, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. Proc Natl Acad Sci U S A *105*, 8309-8314.

Rudney, J.D., Chen, R., and Zhang, G. (2005). Streptococci dominate the diverse flora within buccal cells. Journal of dental research *84*, 1165-1171.

Sadvakassova, G., Dobocan, M.C., Difalco, M.R., and Congote, L.F. (2009). Regulator of differentiation 1 (ROD1) binds to the amphipathic C-terminal peptide of thrombospondin-4 and is involved in its mitogenic activity. J Cell Physiol *220*, 672-679.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science (New York, NY) *230*, 1350-1354.

Sala, A., and Watson, R. (1999). B-Myb protein in cellular proliferation, transcription control, and cancer: latest developments. J Cell Physiol *179*, 245-250.

Sandelin, A., Carninci, P., Lenhard, B., Ponjavic, J., Hayashizaki, Y., and Hume, D.A. (2007). Mammalian RNA polymerase II core promoters: insights from genome-wide studies. Nature reviews Genetics *8*, 424-436.

Sanger, F., and Coulson, A.R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol *94*, 441-448.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A *74*, 5463-5467.

Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. Nature *419*, 407-411.

Sardet, C., Vidal, M., Cobrinik, D., Geng, Y., Onufryk, C., Chen, A., and Weinberg, R.A. (1995). E2F-4 and E2F-5, two members of the E2F family, are expressed in the early phases of the cell cycle. Proc Natl Acad Sci U S A *92*, 2403-2407.

Savinkova, L., Drachkova, I., Arshinova, T., Ponomarenko, P., Ponomarenko, M., and Kolchanov, N. (2013). An experimental verification of the predicted effects of promoter TATA-box polymorphisms associated with human diseases on interactions between the TATA boxes and TATA-binding protein. PloS one *8*, e54626.

Sayan, B.S., Yang, A.L., Conforti, F., Tucci, P., Piro, M.C., Browne, G.J., Agostini, M., Bernardini, S., Knight, R.A., Mak, T.W., *et al.* (2010). Differential control of TAp73 and DeltaNp73 protein stability by the ring finger ubiquitin ligase PIR2. Proc Natl Acad Sci U S A *107*, 12877-12882.

Schlondorff, D. (1987). The glomerular mesangial cell: an expanding role for a specialized pericyte. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *1*, 272-281.

Schneider, R., Bannister, A.J., Myers, F.A., Thorne, A.W., Crane-Robinson, C., and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol *6*, 73-77.

Schober, J.M., Carmichael, P.A., Hines, M., and Ransley, P.G. (2002). The ultimate challenge of cloacal exstrophy. J Urol *167*, 300-304.

Senoo, M., Pinto, F., Crum, C.P., and McKeon, F. (2007). p63 Is essential for the proliferative potential of stem cells in stratified epithelia. Cell *129*, 523-536.

Sethi, I., Sinha, S., and Buck, M.J. (2014). Role of chromatin and transcriptional co-regulators in mediating p63-genome interactions in keratinocytes. BMC genomics *15*, 1042.

Shapiro, E., Lepor, H., and Jeffs, R.D. (1984). The inheritance of the exstrophy-epispadias complex. J Urol *132*, 308-310.

Sharma, P.K., Pandey, P.K., Vijay, M.K., Bera, M.K., Singh, J.P., and Saha, K. (2013). Squamous cell carcinoma in exstrophy of the bladder. Korean journal of urology *54*, 555-557.

Shaulian, E., and Karin, M. (2002). AP-1 as a regulator of cell life and death. Nat Cell Biol *4*, E131-136.

Shendure, J., and Ji, H. (2008). Next-generation DNA sequencing. Nature biotechnology *26*, 1135-1145.

Shiroyanagi, Y., Liu, B., Cao, M., Agras, K., Li, J., Hsieh, M.H., Willingham, E.J., and Baskin, L.S. (2007). Urothelial sonic hedgehog signaling plays an important role in bladder smooth muscle formation. Differentiation *75*, 968-977.

Shotelersuk, V., Janklat, S., Siriwan, P., and Tongkobpetch, S. (2005). De novo missense mutation, S541Y, in the p63 gene underlying Rapp-Hodgkin ectodermal dysplasia syndrome. Clinical and experimental dermatology *30*, 282-285.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., *et al.* (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome research *15*, 1034-1050.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive

chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nature genetics *38*, 1348-1354.

Singh, J.K., Mahajan, J.K., Bawa, M., and Rao, K.L.N. (2011). Covered exstrophy with anorectal malformation and vaginal duplication. J Indian Assoc Pediatr Surg *16*, 26-28.

Sinnatamby, C.S. (2011). Last's anatomy: regional and applied (Elsevier Health Sciences).

Sitzmann, J., Noben-Trauth, K., Kamano, H., and Klempnauer, K.H. (1996). Expression of B-Myb during mouse embryogenesis. Oncogene *12*, 1889-1894.

Skandalakis, J.E., and Gray, S.W. (1994). Embryology for surgeons : the embryological basis for the treatment of congenital anomalies (Baltimore: Williams & Wilkins).

Slaughenhoupt, B.L., Chen, C.J., and Gearhart, J.P. (1996). Creation of a model of bladder exstrophy in the fetal lamb. J Urol *156*, 816-818.

Slavotinek, A.M., Tanaka, J., Winder, A., Vargervik, K., Haggstrom, A., and Bamshad, M. (2005). Acro-dermato-ungual-lacrimal-tooth (ADULT) syndrome: report of a child with phenotypic overlap with ulnar-mammary syndrome and a new mutation in TP63. Am J Med Genet A *138a*, 146-149.

Smale, S.T., and Kadonaga, J.T. (2003). The RNA polymerase II core promoter. Annual review of biochemistry *72*, 449-479.

Smith, L.M., Fung, S., Hunkapiller, M.W., Hunkapiller, T.J., and Hood, L.E. (1985). The synthesis of oligonucleotides containing an aliphatic amino group at the 5' terminus: synthesis of fluorescent DNA primers for use in DNA sequence analysis. Nucleic Acids Res *13*, 2399-2412.

Smith, L.M., Sanders, J.Z., Kaiser, R.J., Hughes, P., Dodd, C., Connell, C.R., Heiner, C., Kent, S.B., and Hood, L.E. (1986). Fluorescence detection in automated DNA sequence analysis. Nature *321*, 674-679.

Socransky, S.S., Manganiello, A.D., Propas, D., Oram, V., and van Houte, J. (1977). Bacteriological studies of developing supragingival dental plaque. Journal of periodontal research *12*, 90-106.

Sorasio, L., Ferrero, G.B., Garelli, E., Brunello, G., Martano, C., Carando, A., Belligni, E., Dianzani, I., and Cirillo Silengo, M. (2006). AEC syndrome: further evidence of a common genetic etiology with Rapp-Hodgkin syndrome. Eur J Med Genet *49*, 520-522.

Southgate, J., Hutton, K.A., Thomas, D.F., and Trejdosiewicz, L.K. (1994). Normal human urothelial cells in vitro: proliferation and induction of stratification. Laboratory investigation; a journal of technical methods and pathology *71*, 583-594.

Sozen, M.A., Suzuki, K., Tolarova, M.M., Bustos, T., Fernandez Iglesias, J.E., and Spritz, R.A. (2001). Mutation of PVRL1 is associated with sporadic, non-syndromic cleft lip/palate in northern Venezuela. Nature genetics *29*, 141-142.

Spicuglia, S., and Vanhille, L. (2012). Chromatin signatures of active enhancers. Nucleus (Austin, Tex) *3*, 126-131.

Spilianakis, C.G., and Flavell, R.A. (2006). Molecular biology. Managing associations between different chromosomes. Science (New York, NY) *312*, 207-208.

Srivastava, D., Thomas, T., Kirby, M.L., Brown, D., and Olson, E.N. (1997). Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND. Nature genetics *16*, 154-160.

Strachan, T., and Read, A.P. (1999). Human gene expression. In Human molecular genetics (Wiley-Liss), p. 576.

218

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Su, Z., Ning, B., Fang, H., Hong, H., Perkins, R., Tong, W., and Shi, L. (2011). Next-generation sequencing and its applications in molecular diagnostics. Expert review of molecular diagnostics *11*, 333-343.

Sukhatme, V.P., Cao, X.M., Chang, L.C., Tsai-Morris, C.H., Stamenkovich, D., Ferreira, P.C., Cohen, D.R., Edwards, S.A., Shows, T.B., Curran, T.*, et al.* (1988). A zinc finger-encoding gene coregulated with c-fos during growth and differentiation, and after cellular depolarization. Cell *53*, 37-43.

Sun, L., Li, P., Carr, A.L., Gorsuch, R., Yarka, C., Li, J., Bartlett, M., Pfister, D., Hyde, D.R., and Li, L. (2014). Transcription of the SCL/TAL1 interrupting locus (Stil) is required for cell proliferation in adult zebrafish retinas. Journal of Biological Chemistry *289*, 6934-6940.

Sun, T., Gao, Y., Tan, W., Ma, S., Shi, Y., Yao, J., Guo, Y., Yang, M., Zhang, X., Zhang, Q.*, et al.* (2007). A six-nucleotide insertion-deletion polymorphism in the CASP8 promoter is associated with susceptibility to multiple cancers. Nature genetics *39*, 605-613.

Suzuki, E., Evans, T., Lowry, J., Truong, L., Bell, D.W., Testa, J.R., and Walsh, K. (1996). The human GATA-6 gene: structure, chromosomal location, and regulation of expression by tissue-specific and mitogen-responsive signals. Genomics *38*, 283-290.

Svanvik, N., Westman, G., Wang, D., and Kubista, M. (2000). Light-up probes: thiazole orange-conjugated peptide nucleic acid for detection of target nucleic acid in homogeneous solution. Analytical biochemistry *281*, 26-35.

Tang, Y., Ma, C.X., Cui, W., Chang, V., Ariet, M., Morse, S.B., Resnick, M.B., and Roth, J. (2006). The risk of birth defects in multiple births: a population-based study. Matern Child Health J *10*, 75-81.

Thauvin-Robinet, C., Faivre, L., Cusin, V., Khau Van Kien, P., Callier, P., Parker, K.L., Fellous, M., Borgnon, J., Gounot, E., Huet, F.*, et al.* (2004). Cloacal exstrophy in an infant with 9q34.1-qter deletion resulting from a de novo unbalanced translocation between chromosome 9q and Yq. Am J Med Genet A *126A*, 303-307.

The Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207-214.

Thiel, G., and Cibelli, G. (2002). Regulation of life and death by the zinc finger transcription factor Egr-1. Journal of Cellular Physiology *193*, 287-292.

Thomas, G.J., and Barton, J.C. (1936). ECtopic pelvic kidney. Journal of the American Medical Association *106*, 197-201.

Thomas, J.C., Oottamasathien, S., Makari, J.H., Honea, L., Sharif-Afshar, A.R., Wang, Y., Adams, C., Wills, M.L., Bhowmick, N.A., Adams, M.C.*, et al.* (2008). Temporal-spatial protein expression in bladder tissue derived from embryonic stem cells. J Urol *180*, 1784-1789.

Thomson, E., Ferreira-Cerca, S., and Hurt, E. (2013). Eukaryotic ribosome biogenesis at a glance. J Cell Sci *126*, 4815-4821.

Tinanoff, N., Gross, A., and Brady, J.M. (1976). Development of plaque on enamel. Parallel investigations. Journal of periodontal research *11*, 197-209.

Tindall, E.A., Petersen, D.C., Woodbridge, P., Schipany, K., and Hayes, V.M. (2009). Assessing high-resolution melt curve analysis for accurate detection of gene variants in complex DNA fragments. Human mutation *30*, 876-883.

Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell *10*, 1453-1465.

Tsutsui, K., Asai, Y., Fujimoto, A., Yamamoto, M., Kubo, M., and Hatta, N. (2003). A novel p63 sterile alpha motif (SAM) domain mutation in a Japanese patient with ankyloblepharon, ectodermal defects and cleft lip and palate (AEC) syndrome without ankyloblepharon. The British journal of dermatology *149*, 395-399.

Tyagi, S., Bratu, D.P., and Kramer, F.R. (1998). Multicolor molecular beacons for allele discrimination. Nature biotechnology *16*, 49-53.

Tyagi, S., and Kramer, F.R. (1996). Molecular beacons: probes that fluoresce upon hybridization. Nature biotechnology *14*, 303-308.

Ueno, K., Hirata, H., Majid, S., Yamamura, S., Shahryari, V., Tabatabai, Z.L., Hinoda, Y., and Dahiya, R. (2012). Tumor suppressor microRNA-493 decreases cell motility and migration ability in human bladder cancer cells by downregulating RhoC and FZD4. Molecular cancer therapeutics *11*, 244-253.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3--new capabilities and interfaces. Nucleic Acids Res *40*, e115.

Urologyhealth.org (2014a). Ectopic ureter (Urology Care Foundation Inc.).

Urologyhealth.org (2014b). Horseshoe Kidney (Renal Fusion) (Urology Care Foundation Inc.).

Urologyhealth.org (2014c). Megaureter (Urology Care Foundation Inc.).

van Bokhoven, H., and Brunner, H.G. (2002). Splitting p63. Am J Hum Genet *71*, 1-13.

van Bokhoven, H., Hamel, B.C., Bamshad, M., Sangiorgi, E., Gurrieri, F., Duijf, P.H., Vanmolkot, K.R., van Beusekom, E., van Beersum, S.E., Celli, J.*, et al.* (2001). p63 Gene mutations in eec syndrome, limb-mammary syndrome, and isolated split hand-

split foot malformation suggest a genotype-phenotype correlation. Am J Hum Genet *69*, 481-492.

van Bokhoven, H., Jung, M., Smits, A.P., van Beersum, S., Ruschendorf, F., van Steensel, M., Veenstra, M., Tuerlings, J.H., Mariman, E.C., Brunner, H.G., *et al.* (1999). Limb mammary syndrome: a new genetic disorder with mammary hypoplasia, ectrodactyly, and other Hand/Foot anomalies maps to human chromosome 3q27. Am J Hum Genet *64*, 538-546.

van Nuland, R., Schram, A.W., van Schaik, F.M., Jansen, P.W., Vermeulen, M., and Marc Timmers, H.T. (2013). Multivalent engagement of TFIID to nucleosomes. PloS one *8*, e73495.

Varley, C.L., Bacon, E.J., Holder, J.C., and Southgate, J. (2009). FOXA1 and IRF-1 intermediary transcriptional regulators of PPARgamma-induced urothelial cytodifferentiation. Cell Death Differ *16*, 103-114.

Vasioukhin, V., Bowers, E., Bauer, C., Degenstein, L., and Fuchs, E. (2001). Desmoplakin is essential in epidermal sheet formation. Nat Cell Biol *3*, 1076-1085.

Vilar, J.M., and Saiz, L. (2005). DNA looping in gene regulation: from the assembly of macromolecular complexes to the control of transcriptional noise. Current opinion in genetics *&* development *15*, 136-144.

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F*., et al.* (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature *457*, 854-858.

Walker, A.H., Najarian, D., White, D.L., Jaffe, J.F., Kanetsky, P.A., and Rebbeck, T.R. (1999). Collection of genomic DNA by buccal swabs for polymerase chain reaction-based biomarker assays. Environmental health perspectives *107*, 517-520.

Walsh, D.J., Corey, A.C., Cotton, R.W., Forman, L., Herrin, G.L., Jr., Word, C.J., and Garner, D.D. (1992). Isolation of deoxyribonucleic acid (DNA) from saliva and forensic science samples containing saliva. Journal of forensic sciences *37*, 387-395.

Wang, H., Meyer, C.A., Fei, T., Wang, G., Zhang, F., and Liu, X.S. (2013a). A systematic approach identifies FOXA1 as a key factor in the loss of epithelial traits during the epithelial-to-mesenchymal transition in lung cancer. BMC genomics *14*, 680.

Wang, J., Zhuang, J., Iyer, S., Lin, X., Whitfield, T.W., Greven, M.C., Pierce, B.G., Dong, X., Kundaje, A., Cheng, Y.*, et al.* (2012). Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. Genome research *22*, 1798-1812.

Wang, J., Zhuang, J., Iyer, S., Lin, X.Y., Greven, M.C., Kim, B.H., Moore, J., Pierce, B.G., Dong, X., Virgil, D.*, et al.* (2013b). Factorbook.org: a Wiki-based database for transcription factor-binding data generated by the ENCODE consortium. Nucleic Acids Res *41*, D171-176.

Weinhold, B. (2006). Epigenetics: the science of change. Environ Health Perspect *114*, A160-167.

Weiskirchen, R., and Gunther, K. (2003). The CRP/MLP/TLP family of LIM domain proteins: acting by connecting. BioEssays : news and reviews in molecular, cellular and developmental biology *25*, 152-162.

Whittock, N.V., and Bower, C. (2003). Genetic evidence for a novel human desmosomal cadherin, desmoglein 4. J Invest Dermatol *120*, 523-530.

Wilkins, S., Zhang, K.W., Mahfuz, I., Quantin, R., D'Cruz, N., Hutson, J., Ee, M., Bagli, D., Aitken, K., Fong, F.N.*, et al.* (2012). Insertion/deletion polymorphisms in the DeltaNp63 promoter are a risk factor for bladder exstrophy epispadias complex. PLoS Genet *8*, e1003070.

Williams, A., Spilianakis, C.G., and Flavell, R.A. (2010). Interchromosomal association and gene regulation in trans. Trends Genet *26*, 188-197.

Williams, S.C. (2013). Epigenetics. Proc Natl Acad Sci U S A *110*, 3209.

Wittwer, C.T. (2009). High-resolution DNA melting analysis: advancements and limitations. Human mutation *30*, 857-859.

Wittwer, C.T., Reed, G.H., Gundry, C.N., Vandersteen, J.G., and Pryor, R.J. (2003). High-resolution genotyping by amplicon melting analysis using LCGreen. Clinical chemistry *49*, 853-860.

Wold, B., and Myers, R.M. (2008). Sequence census methods for functional genomics. Nature methods *5*, 19-21.

Women's Health Matters (2015). Medical Description. In Gynecological health (Health Centers: Women's College Hospital).

Wong, M.L., and Medrano, J.F. (2005). Real-time PCR for mRNA quantitation. In BioTechniques (United States), pp. 75-85.

Woodage, T., Basrai, M.A., Baxevanis, A.D., Hieter, P., and Collins, F.S. (1997). Characterization of the CHD family of proteins. Proc Natl Acad Sci U S A *94*, 11472-11477.

Woodhouse, C.R., and Hinsch, R. (1997). The anatomy and reconstruction of the adult female genitalia in classical exstrophy. British journal of urology *79*, 618-622.

Woychik, N.A., and Hampsey, M. (2002). The RNA polymerase II machinery: structure illuminates function. Cell *108*, 453-463.

Wu, K.J., Polack, A., and Dalla-Favera, R. (1999). Coordinated regulation of iron-controlling genes, H-ferritin and IRP2, by c-MYC. Science (New York, NY) *283*, 676-679.

Wu, X.R., Lin, J.H., Walz, T., Haner, M., Yu, J., Aebi, U., and Sun, T.T. (1994). Mammalian uroplakins. A group of highly conserved urothelial differentiation-related membrane proteins. The Journal of biological chemistry *269*, 13716-13724.

Yamashita, M., Ukai-Tadenuma, M., Miyamoto, T., Sugaya, K., Hosokawa, H., Hasegawa, A., Kimura, M., Taniguchi, M., DeGregori, J., and Nakayama, T. (2004). Essential role of GATA3 for the maintenance of type 2 helper T (Th2) cytokine production and chromatin remodeling at the Th2 cytokine gene loci. The Journal of biological chemistry *279*, 26983-26990.

Yang, A., Kaghad, M., Wang, Y., Gillett, E., Fleming, M.D., Dötsch, V., Andrews, N.C., Caput, D., and McKeon, F. (1998). p63, a p53 homolog at 3q27-29, encodes multiple products with transactivating, death-inducing, and dominant-negative activities. Mol Cell *2*, 305-316.

Yang, A., and McKeon, F. (2000). p63 and p73: p53 mimics, menaces and more. Nat Rev Mol Cell Biol *1*, 199-207.

Yang, A., Schweitzer, R., Sun, D., Kaghad, M., Walker, N., Bronson, R.T., Tabin, C., Sharpe, A., Caput, D., Crum, C.*, et al.* (1999). p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. Nature *398*, 714-718.

Yang, A., Zhu, Z., Kapranov, P., McKeon, F., Church, G.M., Gingeras, T.R., and Struhl, K. (2006). Relationships between p63 binding, DNA sequence, transcription activity, and biological function in human cells. Mol Cell *24*, 593-602.

Yang, C., and Glass, W.F., 2nd (2008). Expression of alpha-actinin-1 in human glomerular mesangial cells in vivo and in vitro. Experimental biology and medicine (Maywood, NJ) *233*, 689-693.

Yang, L., Chen, Y., Cui, T., Knosel, T., Zhang, Q., Albring, K.F., Huber, O., and Petersen, I. (2012). Desmoplakin acts as a tumor suppressor by inhibition of the Wnt/beta-catenin signaling pathway in human lung cancer. Carcinogenesis *33*, 1863-1870.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., and Madden, T.L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC bioinformatics *13*, 134.

Yeaman, C., Grindstaff, K.K., and Nelson, W.J. (2004). Mechanism of recruiting Sec6/8 (exocyst) complex to the apical junctional complex during polarization of epithelial cells. J Cell Sci *117*, 559-570.

Youssoufian, H., McAfee, M., and Kwiatkowski, D.J. (1990). Cloning and chromosomal localization of the human cytoskeletal alpha-actinin gene reveals linkage to the beta-spectrin gene. Am J Hum Genet *47*, 62-72.

Yu, B., and Zhang, C. (2011). In silico PCR analysis. Methods in molecular biology (Clifton, NJ) *760*, 91-107.

Yu, J., Manabe, M., Wu, X.R., Xu, C., Surya, B., and Sun, T.T. (1990). Uroplakin I: a 27-kD protein associated with the asymmetric unit membrane of mammalian urothelium. The Journal of cell biology *111*, 1207-1216.

Yu, Y.T., Breitbart, R.E., Smoot, L.B., Lee, Y., Mahdavi, V., and Nadal-Ginard, B. (1992). Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted MADS box transcription factors. Genes Dev *6*, 1783-1798.

Zaki, M.S., Gillessen-Kaesbach, G., Vater, I., Caliebe, A., Siebert, R., Kamel, A.K., Mohamed, A.M., and Mazen, I. (2012). Bladder exstrophy and extreme genital anomaly in a patient with pure terminal 1q deletion: expansion of phenotypic spectrum. Eur J Med Genet *55*, 43-48.

Zayats, T., Young, T.L., Mackey, D.A., Malecaze, F., Calvas, P., and Guggenheim, J.A. (2009). Quality of DNA extracted from mouthwashes. PloS one *4*, e6165.

Zenteno, J.C., Berdon-Zapata, V., Kofman-Alfaro, S., and Mutchinick, O.M. (2005). Isolated ectrodactyly caused by a heterozygous missense mutation in the transactivation domain of TP63. Am J Med Genet A *134a*, 74-76.

Zhang, W., Cui, H., and Wong, L.J. (2014). Application of next generation sequencing to molecular diagnosis of inherited diseases. Topics in current chemistry *336*, 19-45.

Zipper, H., Brunner, H., Bernhagen, J., and Vitzthum, F. (2004). Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications. Nucleic Acids Res *32*, e103.

Zucchero, T.M., Cooper, M.E., Maher, B.S., Daack-Hirsch, S., Nepomuceno, B., Ribeiro, L., Caprau, D., Christensen, K., Suzuki, Y., Machida, J., *et al.* (2004). Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. N Engl J Med *351*, 769-780.

Appendices

Appendices

Appendix Table I: Primers used for PCR and sequencing of *ΔNP63* promoter regions.

Primers	Forward (5′-3′)	Reverse (5′-3′)
ΔNP63.PP1	CCACATATAAGCATTGTTGTTGAGT	AGGCTCAGGGTCCCAAGTAT
ΔNP63.PP2	CTGATAGCTGAGGTGGCACA	GATTTGCCTCCCCTTTACTTCT
ΔNP63.PP3	AAGTAGAAGTAAAGGGGAGGCA	GCCAGTGGTTTCATAACAAGC
ΔNP63.PP4	ACATTACATTGGGTGACCGAGA	TGGATGTTTGTTTGTTTTTGTAAGT

Appendix Table II: Primers used for HRM and sequencing of *TAP63* promoter regions

Primers	Forward (5′-3′)	Reverse (5′-3′)	Targeted Regions
TAP63.PP1	CCCTATTGCTTTTAGCCTCCCGGC	AGGGCAGTACTGTAGGGTGGCA	R1
TAP63.PP2	GTCAATTGATGAATCTCATTGGCTA	GCAATAGGGTCAAATGCTATCAGA	R2
TAP63.PP3	AGTCCAGGCTGCTGAAATTAAAC	TTCTTACCAGCTGGGCATCCT	R3
TAP63.PP4	AACAGGCGGTTGGCTGAAAG	CCCTGCTCCAGCTACATAACAC	R4
TAP63.PP5	GCTGAGAGGAAGAGTGAGTTC	CAGTTTCCCTTTCAGCCAACC	R5
TAP63.PP6	GGGTGAAGGGAAGATGTGAACTC	TGACCTACTTCGAAAGGAAACCA	R6
TAP63.PP7	TAGCAGGGAAGAAAGCCAAAC	CCCAGCCAAATCAAACAGCTT	R7
TAP63.PP8	TTGTGGAGCAAGCTGTTTGA	TTTGTTGCTTCAGCCAGGAT	R8

Appendix Table III: Primers used in the screening of *PERP* promoter and exons.

Exons	Primers		Annealing Temp.
	Forward (5′-3′)	Reverse (5′-3′)	
PERP **Exons**			
Exon 1	GCTCGTCCTAAACAGGCATT	ATTCACTGAGCTCCCCTTGA	58°C
Exon 2	ACCATGTTGATACTGTGATAAGTAATT	GGAATTGACCATTACTAAGTCGAT	66°C
Exon 3	TGCATTCATACCCTGGGATTT	TCTCAGCAGCAGCGATTT	65°C
PERP **Promoter**			
PERP.PPx1	GCTCGTCCTAAACAGGCATT	ATTCACTGAGCTCCCCTTGA	58°C
PERP.Rare	CCCACCTAGATTCCGGTGAC	GCGCTGTGCAGATAAACAAG	62°C

Appendix Table IV: Primers used in the screening of *DSP* promoter and exons.

Exons	Primers		Annealing Temp.
	Forward (5′-3′)	Reverse (5′-3′)	
DSP Exons			
Exon 1	TGGGAAGAAACCGGCCAG	GAGTCTCCCAGTCGTCCA	63°C
Exon 2	GATGTCTGGTTTCTCTGTGT	CTTTGGGAGAAGCGATGG	65°C
Exon 3	ATGTGTTTTCCTTCATGTGAGT	CCCGATCCATGCTCTGT	60°C
Exon 4	ACGGGTTTTCATAGGCTG	CATCGTGCTGGAACGTC	60°C
Exon 5	GGGGCAACAACAAAGGG	CCTTCGGTAGATATTTTGAAGACTC	65°C
Exon 6	CACAAGGGGATTTATATCTACCTG	AGAAAGTCGATGCACTTGAG	65°C
Exon 7	AATGCTGCCTTTGAACCT	GACAAGCATCTACAGCCG	60°C
Exon 8	GGGGTGATGGAAGTTTAATGATG	AGAAAATAAAGTTGGCACGGCT	63°C
Exon 9	TGTACAACTGAGCTAGGCTAAG	CAATGAAGAGGATTTATGCAACA	65°C
Exon 10	GCTGTTCATTCACTGATCACT	CTGCTGAGGTTCTGGGTC	61°C
Exon 11	CTCTAAAACTCACAGGGTATCTAT	TGTGTGTAAAGACAGGGAAAAC	65°C
Exon 12	ACTTATGAATAAAGCCAAACCCT	CATGCGTGGCTTTGGGAA	61°C
Exon 13	AAGCCTGGCTCTAGCAT	TCCCTCCAGGCAGCCTA	61°C
Exon 14	GGGCAGTCATAAATCCCAAAT	ATGGTATGGATGGTTGATACAAAG	61°C
Exon 15	AGGCCTAGCACCTTGATAC	ACAGGGTAACCAGGCAAA	61°C
Exon 16	TCAGCTAAATCAAAAGAGCTTTCC	ATCTAAAGCAAATGAAGACCCAG	65°C
Exon 17	ACAGACTAATTATGTCACTGGC	AATTAGTTCCTGTAAGTTGGGC	65°C
Exon 18	AGCATACAATGGGAGAAGGGA	GGGAGAAGATGGGGAAGGAG	63°C
Exon 19	GTGAATTTCTGGGTGATTCTATGT	CACCCAATTAATCTCAAAACAGC	63°C
Exon 20	CAGTGGTAAAAGAGTGAAATCACA	CAAATCAGCAAGTCATGCAGTTA	65°C
Exon 21	CTGTAGCTGTTAGTGATGCTTT	GGAAGAAGCAGAAGGAGGG	63°C
Exon 22	AAATCGGTCAAATTACATAGGACT	GACGCTACAAGAAAAGACAGTAT	65°C
Exon 23.1	GGTCTGGGAGGGGAAAAGTA	GCATTATCTCAGAGCCACAGG	63°C
Exon 23.2	TCAGGCTCAATGACAGCATC	ATCCTCAGGCGTGTCAGTTC	63°C
Exon 23.3	ACGCCTGAGGATCGACTATG	AGCAGAAAGAGCACAGTTCAGA	63°C

| Exons | Primers | | Annealing Temp. |
	Forward (5′-3′)	Reverse (5′-3′)	
Exon 24.1	TCACAGTGTATCCAGGGACAA	TCTTTCTCTTGGCCTCTACCA	61°C
Exon 24.2	TGAAATCCAGCCATTCCTTC	CCAGCAACTCCAGAGCAGTA	61°C
Exon 24.3	TGGCTTAGTCCGACCTGGTA	CCGACACCATTTTTCAAGGA	61°C
Exon 24.4	CAGCCTCACTCAATTTGCTG	TAGCTGCTGGTGTCTTGCAG	61°C
Exon 24.5	CAGCAGAGGCAGTGAAAGAA	AGCAACCACTCCCAACTGAC	61°C
DSP Promoter			
DSP.PP1	GCAGTGTTTTCCTCTTCATGC	ACACCTGGCCGGTTTCTT	63°C
DSP.PP2	CCGCCCGGTGTTACTCAG	CGCGCCGAGAATAGTACATC	61°C

Appendix Table V: Primers used in QPCR experiments on targeted gene loci to check ChIP efficiency.

Primers	Forward	Reverse
ACTB+	TCCAAAGGAGACTCAGGTCAG	CGCCCTTTCTCACTGGTTC
ΔNP63+	CAGGGCCTGTGCTATCGTAT	AGCAAGCCTGCTTGAATGTT
PERP+	GGCTTCCCACTGTGAAAGAA	GTATGCCATTTTCCCAGCAT
DSP+	ATCTCTCTGGGAAGGGGAAA	GATCCCCGGAAAGAATGAC
FOS-	GGCCAAGAAGTGATTATGTGG	CCCATCCTTAGAACTACAAAGACC
MYOG-	AAGTTTGACAAGTTCAAGCACCTG	TGGCACCATGCTTCTTTAAGTC

Appendix Table VI: Description and type of BEEC candidate gene found containing H3K4me3 and H3K27ac enrichment.

No.	Gene Name	Gene Description	Gene Type
1.	*ABRACL*	ABRA C-terminal like	Protein-coding
2.	*ACTN1*	Actinin, alpha 1	Protein-coding
3.	*AFF3*	AF4/FMR2 family, member 3	Protein-coding
4.	*AIM1*	Absent in melanoma 1	Protein-coding
5.	*ALDH1B1*	Aldehyde dehydrogenase 1 family, member B1	Protein-coding
6.	*AMMECR1*	Alport syndrome, mental retardation, midface hypoplasia and elliptocytosis chromosomal region gene 1	Protein-coding
7.	*APRT*	Adenine phosphoribosyltransferase	Protein-coding
8.	*AQP3*	Aquaporin 3 (Gill blood group)	Protein-coding
9.	*ASS1*	Argininosuccinate synthase 1	Protein-coding
10.	*BAG2*	BCL2-associated athanogene 2	Protein-coding
11.	*CA12*	Carbonic anhydrase XII	Protein-coding
12.	*CALD1*	Caldesmon 1	Protein-coding
13.	*CAMK2G*	Calcium/calmodulin-dependent protein kinase II gamma	Protein-coding
14.	*CD24*	CD24 molecule	Protein-coding
15.	*CDC42EP3*	CDC42 effector protein (Rho GTPase binding) 3	Protein-coding
16.	*CDCA7*	Cell division cycle associated 7	Protein-coding
17.	*CDK1*	Cyclin-dependent kinase 1	Protein-coding
18.	*CFL2*	Cofilin 2 (muscle)	Protein-coding
19.	*CLIC4*	Chloride intracellular channel 4	Protein-coding
20.	*CRABP2*	Cellular retinoic acid binding protein 2	Protein-coding
21.	*CSRP1*	Cysteine and glycine-rich protein 1	Protein-coding
22.	*CSTB*	Cystatin B (stefin B)	Protein-coding
23.	*CTSC*	Cathepsin C	Protein-coding
24.	*DDX39A*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 39A	Protein-coding
25.	*DES*	Desmin	Protein-coding

No.	Gene Name	Gene Description	Gene Type
26.	DHCR24	24-dehydrocholesterol reductase	Protein-coding
27.	DMPK	Dystrophia myotonica-protein kinase	Protein-coding
28.	DSP	Desmoplakin	Protein-coding
29.	DTL	Denticleless E3 ubiquitin protein ligase homolog (Drosophila)	Protein-coding
30.	DTNA	Dystrobrevin, alpha	Protein-coding
31.	EFNB2	Ephrin-B2	Protein-coding
32.	ERMP1	Endoplasmic reticulum metallopeptidase 1	Protein-coding
33.	FABP5	Fatty acid binding protein 5 (psoriasis-associated)	Protein-coding
34.	FAM129A	Family with sequence similarity 129, member A	Protein-coding
35.	FERMT2	Fermitin family member 2	Protein-coding
36.	FHL1	Four and a half LIM domains 1	Protein-coding
37.	FLNA	Filamin A, alpha	Protein-coding
38.	FN1	Fibronectin 1	Protein-coding
39.	GCOM1	GRINL1A complex locus 1	Protein-coding
40.	GEM	GTP binding protein overexpressed in skeletal muscle	Protein-coding
41.	GGH	Gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)	Protein-coding
42.	GJC1	Gap junction protein, gamma 1, 45kDa	Protein-coding
43.	GPR126	G protein-coupled receptor 126	Protein-coding
44.	GREM1	Gremlin 1, DAN family BMP antagonist	Protein-coding
45.	HGF	Hepatocyte growth factor (hepapoietin A; scatter factor)	Protein-coding
46.	HN1	Hematological and neurological expressed 1	Protein-coding
47.	IGFBP2	Insulin-like growth factor binding protein 2, 36kDa	Protein-coding
48.	INPP5A	Inositol polyphosphate-5-phosphatase, 40kDa	Protein-coding
49.	ITGA2	Integrin, alpha 2 (CD49B, alpha 2 subunit of VLA-2 receptor)	Protein-coding
50.	ITGA5	Integrin, alpha 5 (fibronectin receptor, alpha polypeptide)	Protein-coding

No.	Gene Name	Gene Description	Gene Type
51.	ITGA6	Integrin, alpha 6	Protein-coding
52.	JAM3	Junctional adhesion molecule 3	Protein-coding
53.	KCTD1	Potassium channel tetramerisation domain containing 1	Protein-coding
54.	KIAA1217	KIAA1217	Protein-coding
55.	KLF5	Kruppel-like factor 5 (intestinal)	Protein-coding
56.	KRT18	Keratin 18	Protein-coding
57.	LPP	LIM domain containing preferred translocation partner in lipoma	Protein-coding
58.	MCM4	Minichromosome maintenance complex component 4	Protein-coding
59.	MCM5	Minichromosome maintenance complex component 5	Protein-coding
60.	MLF1IP	MLF1 interacting protein	Protein-coding
61.	MREG	Melanoregulin	Protein-coding
62.	MSMO1	Methylsterol monooxygenase 1	Protein-coding
63.	MSRB3	Methionine sulfoxide reductase B3	Protein-coding
64.	MYL9	Myosin, light chain 9, regulatory	Protein-coding
65.	MYO10	Myosin X	Protein-coding
66.	NANS	N-acetylneuraminic acid synthase	Protein-coding
67.	NDFIP2-AS1	NDFIP2 antisense RNA 1	Miscrna
68.	NEXN	Nexilin (F actin binding protein)	Protein-coding
69.	NUDT5	Nudix (nucleoside diphosphate linked moiety X)-type motif 5	Protein-coding
70.	NUP50	Nucleoporin 50kDa	Protein-coding
71.	NUSAP1	Nucleolar and spindle associated protein 1	Protein-coding
72.	PALLD	Palladin, cytoskeletal associated protein	Protein-coding
73.	PCNA	Proliferating cell nuclear antigen	Protein-coding
74.	PDGFA	Platelet-derived growth factor alpha polypeptide	Protein-coding
75.	PDLIM1	PDZ and LIM domain 1	Protein-coding
76.	PDLIM7	PDZ and LIM domain 7	Protein-coding

No.	Gene Name	Gene Description	Gene Type
77.	*PERP*	PERP, TP53 apoptosis effector	Protein-coding
78.	*PEX14*	Peroxisomal biogenesis factor 14	Protein-coding
79.	*PHLDA2*	Pleckstrin homology-like domain, family A, member 2	Protein-coding
80.	*PLEKHO1*	Pleckstrin homology domain containing, family O member 1	Protein-coding
81.	*POLR2M*	Polymerase (RNA) II (DNA directed) polypeptide M	Protein-coding
82.	*PPIB*	Peptidylprolyl isomerase B (cyclophilin B)	Protein-coding
83.	*PPP1R12A*	Protein phosphatase 1, regulatory subunit 12A	Protein-coding
84.	*PRDX4*	Peroxiredoxin 4	Protein-coding
85.	*PRKACB*	Protein kinase, cAMP-dependent, catalytic, beta	Protein-coding
86.	*PRUNE2*	Prune homolog 2 (Drosophila)	Protein-coding
87.	*PTBP2*	Polypyrimidine tract binding protein 2	Protein-coding
88.	*PTBP3*	Polypyrimidine tract binding protein 3	Protein-coding
89.	*PTPLA*	Protein tyrosine phosphatase-like (proline instead of catalytic arginine), member A	Protein-coding
90.	*RAB23*	RAB23, member RAS oncogene family	Protein-coding
91.	*RAB27B*	RAB27B, member RAS oncogene family	Protein-coding
92.	*REXO2*	REX2, RNA exonuclease 2 homolog (S. cerevisiae)	Protein-coding
93.	*RHOB*	Ras homolog family member B	Protein-coding
94.	*RNF144B*	Ring finger protein 144B	Protein-coding
95.	*RRM2*	Ribonucleotide reductase M2	Protein-coding
96.	*S100A11*	S100 calcium binding protein A11	Protein-coding
97.	*S100A16*	S100 calcium binding protein A16	Protein-coding
98.	*SACS*	Spastic ataxia of Charlevoix-Saguenay (sacsin)	Protein-coding
99.	*SCHIP1*	Schwannomin interacting protein 1	Protein-coding
100.	*SCUBE3*	Signal peptide, CUB domain, EGF-like 3	Protein-coding
101.	*SEMA3C*	Sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3C	Protein-coding

No.	Gene Name	Gene Description	Gene Type
102.	*SFN*	Stratifin	Protein-coding
103.	*SGCB*	Sarcoglycan, beta (43kDa dystrophin-associated glycoprotein)	Protein-coding
104.	*SH3YL1*	SH3 domain containing, Ysc84-like 1 (S. cerevisiae)	Protein-coding
105.	*SHMT2*	Serine hydroxymethyltransferase 2 (mitochondrial)	Protein-coding
106.	*SLC16A1*	Solute carrier family 16, member 1 (monocarboxylic acid transporter 1)	Protein-coding
107.	*SLC25A4*	Solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 4	Protein-coding
108.	*SLC8A1*	Solute carrier family 8 (sodium/calcium exchanger), member 1	Protein-coding
109.	*SLMAP*	Sarcolemma associated protein	Protein-coding
110.	*SMTN*	Smoothelin	Protein-coding
111.	*SOX9*	SRY (sex determining region Y)-box 9	Protein-coding
112.	*SVIL*	Supervillin	Protein-coding
113.	*SYNM*	Synemin, intermediate filament protein	Protein-coding
114.	*SYNPO2*	Synaptopodin 2	Protein-coding
115.	*TECR*	Trans-2,3-enoyl-CoA reductase	Protein-coding
116.	*TFRC*	Transferrin receptor (p90, CD71)	Protein-coding
117.	*TGM2*	Transglutaminase 2 (C polypeptide, protein-glutamine-gamma-glutamyltransferase)	Protein-coding
118.	*TIMM13*	Translocase of inner mitochondrial membrane 13 homolog (yeast)	Protein-coding
119.	*TMED3*	Transmembrane emp24 protein transport domain containing 3	Protein-coding
120.	*TNS1*	Tensin 1	Protein-coding
121.	*TOP2A*	Topoisomerase (DNA) II alpha 170kDa	Protein-coding
122.	*TPM1*	Tropomyosin 1 (alpha)	Protein-coding
123.	*TPM2*	Tropomyosin 2 (beta)	Protein-coding
124.	*TSPO*	Translocator protein (18kDa)	Protein-coding
125.	*TUBA4A*	Tubulin, alpha 4A	Protein-coding

No.	Gene Name	Gene Description	Gene Type
126.	*TXNDC17*	Thioredoxin domain containing 17	Protein-coding
127.	*TYMS*	Thymidylate synthetase	Protein-coding
128.	*UHRF1*	Ubiquitin-like with PHD and ring finger domains 1	Protein-coding
129.	*WNT5A*	Wingless-type MMTV integration site family, member 5A	Protein-coding
130.	*ZAK*	Sterile alpha motif and leucine zipper containing kinase AZK	Protein-coding
131.	*ZEB1*	Zinc finger E-box binding homeobox 1	Protein-coding
132.	*ZNF185*	Zinc finger protein 185 (LIM domain)	Protein-coding
133.	*ZWINT*	ZW10 interactor, kinetochore protein	Protein-coding

Appendix Table VII: Enrichment of H3K27ac near *PERP* and *DSP* genes.

Gene	Distance to TSS	Genomic Regions	Location of the Peaks	Peak Score
	-68453	Intergenic	Chr6:7472951-7473884	14.3
	-59913	Intergenic	Chr6:7480228-7483687	40.8
	-56877	Intergenic	Chr6:7484356-7485631	18.9
	-52760	Intergenic	Chr6:7488860-7489360	6.1
	-48260	Intergenic	Chr6:7493360-7493860	6.6
	-46143	Intergenic	Chr6:7495290-7496165	10.7
	-43958	Intergenic	Chr6:7497662-7498162	7.1
	-42275	Intergenic	Chr6:7498997-7500194	17.3
	-36962	Intergenic	Chr6:7500830-7508987	**134.1**
	-31703	Intergenic	Chr6:7509729-7510605	11.7
DSP	-28000	Intergenic	Chr6:7513068-7514672	20.9
	-21730	Intergenic	Chr6:7519827-7520454	8.7
	-17632	Intergenic	Chr6:7523243-7525233	25
	-11231	Intergenic	Chr6:7525981-7535297	127
	-5465	Intergenic	Chr6:7535961-7536849	11.7
	-3566	Intergenic	Chr6:7537685-7538924	18.9
	1429	Intron 1	Chr6:7539447-7547152	**120.3**
	7924	Intron 1	Chr6:7548614-7550974	28.6
	10267	Intron 1	Chr6:7551747-7552527	11.7
	12032	Intron 1	Chr6:7553585-7554219	9.2
	-502	Promoter-TSS	Chr6:138428912-138429412	6.6
	1445	Intron 1	Chr6:138426873-138427558	10.2
PERP	18517	3' UTR	Chr6:138409893-138410393	7.1
	73389	Intergenic	Chr6:138355021-138355521	7.6

Appendix Table VIII: H3K4me3 peak locations near the most down-regulated BEEC candidate gene.

Gene	Distance to TSS	Genomic Regions	Location of the Peaks	Peak Score
ACTN1	1365	Intron 1	Chr14:69440795-69448642	**205**
	7166	Intron 1	Chr14:69438667-69439167	8.7
	11922	Intron 1	Chr14:69433911-69434411	9.5
	27811	Intron 1	Chr14:69418022-69418522	8.7
CALD1	-13594	Intron 3	Chr7:134562307-134562807	8
	-1738	Intron 3	Chr7:134574163-134574663	8
	1116	Intron 1	Chr7:134463364-134467197	**106.2**
	1328	Intron 1	Chr7:134575808-134579151	**99.6**
	8327	Intron 1	Chr7:134584068-134584888	11.6
CSRP1	122	Promoter-TSS	Chr1:201474060-201478470	**86.5**
	2841	Intron 1	Chr1:201472876-201473376	8
FLNA	502	Intron 1	ChrX:153602050-153602958	13.8
	2044	Intron 1	ChrX:153600590-153601334	12.4
	4187	Intron 2	ChrX:153598234-153599404	21.8
FLNC	524	Exon 1	Chr7:128467896-128474118	**207.9**
FN1	-183927	Intron 8	Chr2:216484391-216485045	10.9
	-25749	Intergenic	Chr2:216326290-216326790	7.3
	-12460	Intergenic	Chr2:216313001-216313501	8
	396	Exon 1	Chr2:216296060-216304731	**196.3**
LPP	41242	Intron 2	Chr3: 187984106-187984765	11.6
	70053	Intron 2	Chr3:188012832-188013661	**13.1**
TPM1	714	Intron 1	Chr15:63340718-63341983	17.4
	1238	Intron 2	Chr15:63332747-63339406	**169.4**

Appendix Table IX: H3K27ac enrichment near the most down-regulated BEEC candidate gene.

Gene	Distance to TSS	Genomic Regions	Location of the Peaks	Peak Score
	4325	Intron 1	Chr14:69441123-69442393	18.9
	5715	Intron 1	Chr14:69440118-69440618	6.6
	12059	Intron 1	Chr14:69433625-69434424	11.2
	15976	Intron 1	Chr14:69429734-69430480	10.2
	19336	Intron 1	Chr14:69426497-69426997	7.1
	27139	Intron 1	Chr14:69418694-69419194	7.6
ACTN1	30651	Intron 1	Chr14:69415105-69415760	8.2
	31856	Intron 1	Chr14:69413977-69414477	8.2
	33871	Intron 1	Chr14:69411889-69412536	8.7
	35130	Intron 1	Chr14:69410703-69411203	7.1
	38177	Intron 1	Chr14:69407269-69408543	17.3
	40232	Intron 1	Chr14:69405601-69406101	6.1
	41663	Intron 1	Chr14:69404170-69404670	7.6
	43068	Intron 1	Chr14:69402765-69403265	6.6
	-44794	Intron 2	Chr7:134531107-134531607	6.6
	1431	Intron 1	Chr7:134464452-134466738	28
CALD1	1600	Intron 1	Chr7:134576758-134578745	25
	9357	Intron 1	Chr7:134585258-134585758	6.6
	11760	Intron 1	Chr7:134587338-134588484	15.8
	53345	Intron 1	Chr7:134517259-134517759	7.1
	-4616	Intergenic	Chr1:201480615-201481392	9.7
	-3819	Intron 1	Chr1:201469186-201469854	9.2
	-1971	Intron 1	Chr1:201467422-201467922	7.1
	-366	Promoter-TSS	Chr1:201476503-201477003	6.6
CSRP1	1457	Intron 1	Chr1:201474045-201474975	15.8
	2138	Intron 2	Chr1:201463313-201463813	6.1
	3981	Intron 2	Chr1:201461470-201461970	7.6
	4312	Intron 1	Chr1:201471405-201471905	7.1
	7332	Intron 3	Chr1:201456750-201459988	45.4

Gene	Distance to TSS	Genomic Regions	Location of the Peaks	Peak Score
	12494	Exon 6	Chr1:201452957-201453457	7.1
DES	-8962	Intergenic	Chr2:220273887-220274387	9.2
	-8674	Intron 9	Chr7:128461559-128462059	6.6
	-3976	Intergenic	Chr7:128466257-128466757	7.6
FLNC	586	Intron 1	Chr7:128470692-128471447	11.2
	1797	Intron 1	Chr7:128472030-128472530	7.1
	4490	Intron 1	Chr7:128474556-128475391	11.2
	-30124	Intergenic	Chr2:216330561-216331269	10.7
	-26560	Intergenic	Chr2:216326948-216327754	11.7
FN1	-1358	Intergenic	Chr2:216301899-216302399	6.6
	1013	Intron 1	Chr2:216299462-216300095	10.2
	-91147	Intergenic	Chr3:187780266-187780766	8.2
	-1156	Exon 1	Chr3:187869855-187871159	15.3
	60454	Intron 2	Chr3-188003397-188003897	7.6
LPP	69757	Intron 2	Chr3-188012700-188013200	7.1
	448957	Intron 6	Chr3-188391900-188392400	7.1
	452878	Intron 6	Chr3-188395821-188396321	7.6
	455077	Intron 6	Chr3-188398020-188398520	6.6
	-11190	Intergenic	Chr2-218819279-218820694	19.9
	-2474	Intergenic	Chr2-218811020-218811520	6.6
TNS1	19316	Intron 1	Chr2-218789067-218789893	11.2
	40249	Intron 4	Chr2-218768297-218768797	6.1
	42432	Intron 4	Chr2-218766114-218766614	7.1
	-109118	Intergenic	Chr15-63225470-63225970	7.6
	-29240	Intergenic	Chr15-63305348-63305848	6.6
TPM1	-22456	Intergenic	Chr15-63312132-63312632	7.1
	-14005	Intergenic	Chr15-63320448-63321218	9.7
	-1603	Intron 2	Chr15-63338717-63339350	9.2

Published Articles

Istiak Mahfuz

Published Articles

I. **Mahfuz, I.**, Darling, T., Wilkins, S., White, S., and Cheng, W. (2013). New insightsinto the pathogenesis of bladder exstrophy-epispadias complex. J Pediatr Urol *9*,996-1005.

II. Wilkins, S., Zhang, K. W., **Mahfuz, I.**, Quantin, R., D'Cruz, N., Hutson, J. . . . Cheng, W. 2012. Insertion/Deletion Polymorphisms in the *ΔNP63* Promoter Are a Risk Factorfor Bladder Exstrophy Epispadias Complex. PLOS Genetics *8(12)*. doi: 10.1371/journal.pgen.1003070.

III. Darling, T., **Mahfuz, I.**, White, S.J., and Cheng, W. (2013). No *TAP63* promoter mutation is detected in bladder exstrophy-epispadias complex patients. J Pediatr Surg *48*, 2393-2400.

IV. **Mahfuz, I.**, Cheng, W., and White, S.J. (2013). Identification of *Streptococcus parasanguinis* DNA contamination in human buccal DNA samples. BMC Res Notes *6*,481.

V. Kabir, A. H., Ahmed, I., Ahmed, M. B., **Mahfuz, I.** 2014. Cloning and site directed mutagenesis in *NALP3* gene. Minerva Biotecnologica. *16(1)*, 17-21.

Submitted for Publication

VI. **Mahfuz I.**, White, S., Cheng, W. (2014). Mutation screening of TP63 related target genes; *PERP* and *DSP* in Bladder Exstrophy Epispadias Complex. BMC Res Notes (Submitted for publication).

REVIEW ARTICLE

New insights into the pathogenesis of bladder exstrophy–epispadias complex

Istiak Mahfuz [a], Tom Darling [a], Simon Wilkins [a,b], Stefan White [a],
Wei Cheng [a,c,d,*]

[a] Monash Institute of Medical Research, Faculty of Medicine, Nursing and Health Sciences, Monash
University, Australia
[b] Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive
Medicine, Monash University, Australia
[c] Department of Paediatrics, Southern Medical School, Faculty of Medicine, Nursing and Health
Sciences, Monash University, Australia
[d] Department of Surgery, Southern Medical School, Faculty of Medicine, Nursing and Health Sciences,
Monash University, Australia

Abstract Bladder exstrophy–epispadias complex (BEEC) is a complex and debilitating congenital disease. Familial and twin studies suggest a possible genetic component in BEEC pathogenesis. Bladder mesenchyme (detrusor) development requires induction by a signal from bladder urothelium, and we and others have shown the Shh–Gli–Bmp4 signalling pathway is likely to be involved. P63 is a master regulator in epithelial stratification and is expressed in urothelium. We have shown that p63 knock-out mice undergo excessive urothelial apoptosis. Failure of mesenchymal induction by epithelium leads to BEEC. We further demonstrated that insertion/deletion (in/del) polymorphisms (1 base pair (bp) ins and 4 bp ins., and 12 bp del) in the ΔNP63 promoter reduce transcriptional efficiency, and are associated with a statistically significant increase in the risk of BEEC in humans. Furthermore, a Genome-Wide Expression Profiling (GWEP) study suggests possible involvement of PERP in human BEEC. Intriguingly, PERP is a direct target of p63 during development, and is also involved in epithelial stratification. PERP co-localizes with desmosome, and both PERP and desmosome are essential for maintaining tissue integrity by cellular adhesion and epithelial stratification. A recent study showed that PERP and desmosome expression levels are abnormal in human BEEC patients. This review describes the role of the

* Corresponding author. Monash Institute of Medical Research, Faculty of Medicine, Nursing and Health Sciences, Monash University, Australia.

P63 > *PERP* > desmosome pathway in the development of human bladder during embryogenesis. We hypothesize that disruption of this pathway may increase the risk of BEEC.

Introduction

Bladder exstrophy—epispadias complex (BEEC) is a serious congenital urological abnormality, which greatly impairs both bladder function and the quality of life of those affected. BEEC is manifested as a cluster of ventral midline defects [1] (Fig. 1A and B), which include: i) epispadias, ii) separation of pubic bones and the rectus abdominis muscles, iii) bladder exstrophy, iv) exomphalos, v) ventrally displaced or imperforate anus, and/or vi) cloacal exstrophy. The spectrum of severity ranges from epispadias (E) to classic bladder exstrophy (CBE) to cloacal exstrophy (CE), the most severe form. Without treatment, the affected babies suffer from a progressive decline in renal function, a constant stench, and severe psychosocial strain for both patients and parents. Untreated BEEC patients have a 700 fold increased chance of developing bladder cancer [2]. Even after mostly successful surgical correction, complications such as urinary and faecal incontinence [3], reduced bladder volume, genital disfigurement, sexual dysfunction, and psychological stress [4] are common and debilitating.

Pathogenesis

Previous theories regarding BEEC pathogenesis

The cause of BEEC was conventionally thought to be due to the failure of the cloacal membrane being reinforced by ingrowth of mesoderm [5]. Marshall and Muecke [6] in 1968 explained that: "the defect is an abnormal over-development of the cloacal membrane, which prevents medial migration of the mesenchymal tissue and proper lower abdominal wall development". Several other theories exist concerning the cause of BEEC, focussing on the involvement of cloacal membrane and mesenchymal tissues during their defective evolution in embryogenesis [7—10]. These theories do not explain the molecular aspects of the pathogenesis.

Due to the rarity of BEEC in humans, a number of animal models were developed to explain its pathogenesis. One model is in the sheep, where foetal lambs underwent *in utero* surgical creation of classic bladder exstrophy [11]. Another model is the chick embryo, where the fertilized eggs were

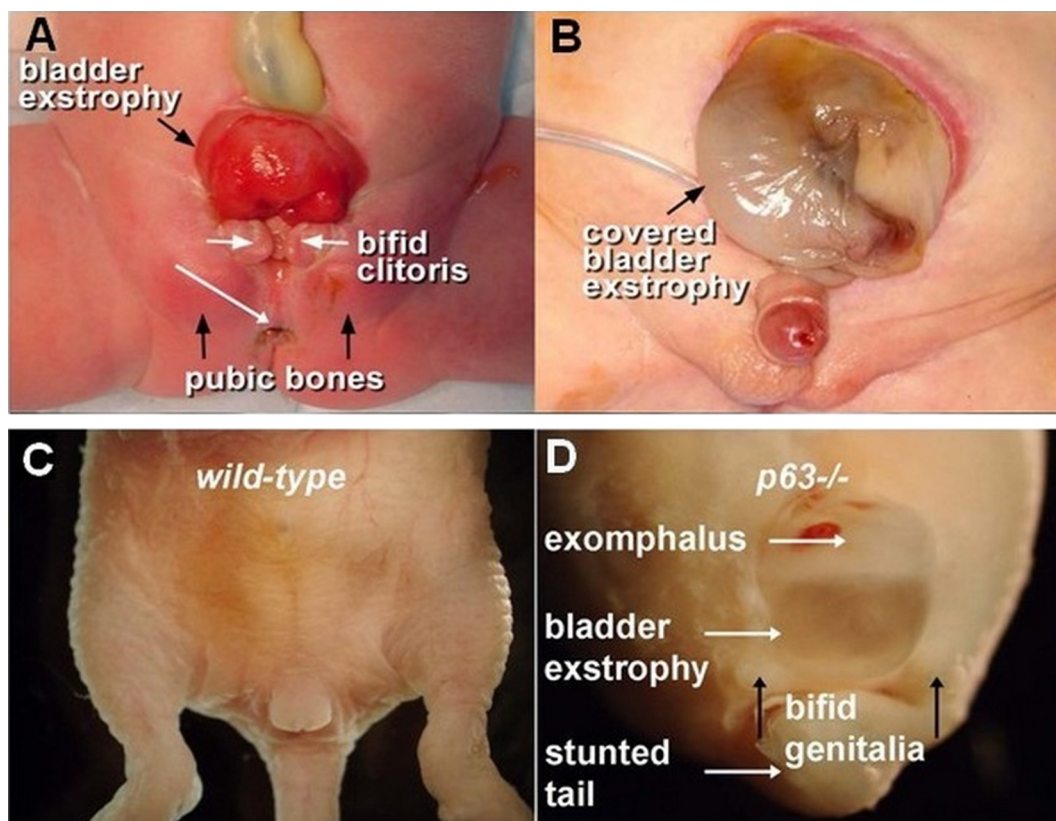**Figure 1**   A: BEEC in a girl with exposed bladder and separated pubic symphysis. B: BEEC variant in a boy with bifid scrotum and covered bladder exstrophy. C: Wild-type E18 mouse embryo. D: $p63^{-/-}$ mouse embryo with bladder exstrophy, exomphalos (umbilical hernia), bifid genitalia (epispadias), stunted tail and absence of hind limbs. Adapted with permission from Development, 2006 [14].

Table 1   BEEC-associated anomalies seen in BEEC patients and the $p63^{-/-}$ mice.

	BEEC patients	$p63^{-/-}$ knock-out mice
Related anomaly	Exomphalos [15]	Exomphalos [14]
	Ano-rectal malformations [16,17]	Ano-rectal malformations [18]
	Spina bifida, sacral hypoplasia, myelo-meningocele [19]	Stunted tail [20]
	Cleft palate, median cleft face syndrome [21,22]	Cleft palate [23]
	Epidermolysis bullosa [24]	Non-stratified skin epidermis [23]
	Absence of feet, tibial deformity [21]	Absent or stunted limb buds [20]

Adapted with permission from PLoS Genetics, 2012 [75].

windowed externally to inject suramin and/or trypan blue drug to induce cloacal exstrophy [12]. These models do not satisfactorily explain the complete BEEC spectrum reported in children, especially the pelvic bone anomaly. However, a better explanation of the pelvic abnormality was offered using a rabbit embryo model [13]. The above BEEC models generally explain anatomical and/or embryological aspects of the condition, but not the molecular basis. We have recently established the only genetic model (a $p63^{-/-}$ knock-out mouse), which shows the complete spectrum of BEEC, i.e. bladder exstrophy, epispadias, separation of pubic bones, imperforate anus and exomphalos (Fig. 1C and D) [14]. More interestingly, these mice also develop the associated anomalies of BEEC (Table 1).

Environmental factors

Data on environmental risk factors in BEEC pathogenesis are limited and inconclusive [25–27]. Due to the rarity of the condition, prospective population-based studies are not possible for identifying possible environmental factors. Recent studies have suggested that maternal smoking and medical radiation during the first trimester may be associated with severe exstrophy phenotype [28]. Birth weight, gestational age and maternal reproductive history are not significantly different from those in the general population [29] in disease pathogenesis.

Genetic explanation

The overall incidence of complete BEEC spectrum in children of European descent is 1 in 10,000 live births, with a greater proportion of males affected [28–31]. Birth prevalence for specific BEEC subtypes (includes terminated pregnancies) are as follows: epispadias was estimated to occur in 1 in 117,000 males and 1 in 484,000 females [32]; for CBE and CE, it is 1 in 37,000 [33] and 1 in 200,000 to 1 in 400,000 [34], respectively. Among siblings and offspring of BEEC patients, the risks of the disease increase dramatically from 1 in 10,000–50,000 to 1 in 100 and 1 in 70 respectively, representing a 500-fold increase in incidence [31,35]. Further evidence of genetic involvement comes from twin studies, with monozygotic twins significantly more prone (62%) to be concordant for BEEC than dizygotic twins (11%) [31]. Males were more common in both the

Table 2   Reported chromosomal anomalies in BEEC patients.

Chromosomal location	Type of mutation/genotype	Phenotype	Reference
—	47,XXY	E	Raboch [37]
9p	Duplication [dup(9p)] of the short arm (47,XY)	E	Chipail et al. [38]
	Trisomy (47,XXX)	CE	Lin et al. [39]
4p	Deletion in the short arm of chromosome 4 (46,XY)	E	Nicholls and Duffy [40]
21	Duplication of chromosome 21 [dup(21)] (47,XX)	CE	Husmann and
—	45,X0/46,XX mosaicism	CE	Vandersteen [41]
—	Trisomy (47,XXX)	CE	
—	Diploid/Tetraploid/t(1; 6) mosaicism [in fibroblasts: 16% (3 cells) 92,XXXX; 11% (2 cells) 46,XX,t(1; 6)(p32; q13); 73% (14 cells) 46,XX]	CE with hypomelanosis of Ito	Leonard and Tomkins [42]
9q34.1-qter Deletion	*De novo* unbalanced translocation between chromosome 9q and Yq. 46,Xder(Y)t(Y; 9)(q11.23; q34.1)−del(Y)(q11.2),der(9)t(Y; 9)	CE	Thauvin-Robinet et al. [43]
3q12.2−13.2	*De novo* deletion (46,XY)	CE	Kosaki et al. [44]
t(8;9)(p11.2;q13)	Translocation between 8p11.2 and 9q13 (46,XY)	CBE	Boyadjiev et al. [45]
t(2;9)(q13;q32)	Translocation between 2q13 and 9q32 (46,XY)	CBE	Ludwig et al. [46]
22q11.21	*De novo* microduplication	CBE	Draaken et al. [47]
1q43q44	*De novo* 10.4 Mb deletion	CBE with absence phallus	Zaki et al. [48]
19p13.12	*De novo* 0.9 Mb microduplication	CBE	Draaken et al. [49]

Table 3 Genetic loci associated with BEEC pathogenesis.

Chromosomal location	Description	Reference
4q31.21—22 19q13.31—41	Harbor genes for an autosomal recessive form of BEEC	Reutter et al. [50]
2p22.1—p21 2p25.2—p25.1 4q23—q32.3 7q21.3—q33 7q34—q36.1 14q31.1—q32.2 19q13.33—q13.43	Evidence for possible risk/modifying loci Classic Bladder exstrophy on chromosomes	Ludwig et al. [51]

epispadias group (M/F, 2.2, 29 patients) and the classic bladder exstrophy group (M/F 1.8, 164 patients), but the sex ratio was close to unity in the cloacal exstrophy group (1.1, 15 patients). BEEC is significantly more prevalent in children born to older mothers [29] and within Caucasian populations [29,36]. Some chromosomal anomalies have been reported in a small number of patients, and are listed in Table 2. There have also been several genetic loci associated with BEEC pathogenesis (Table 3).

New insights

Employing developmental biology and molecular genetic technologies enables BEEC to be explained from a different perspective.

Sonic hedgehog (*SHH*) signalling pathway mediates epithelial—mesenchymal interaction and bladder organogenesis

Baskin and co-workers showed that epithelial—mesenchymal interaction is essential in bladder mesenchymal (smooth muscle) development [52,53], *i.e.* a diffusible signal from the urothelium is responsible for inducing mesenchyme development (Fig. 2A and B). The mesenchyme further develops into the detrusor muscle. Haraguchi et al. [54] identified *Shh* as a diffusible signal involved in the external genitalia development. In the bladder, we and others have demonstrated that *Shh* is involved in development (Fig. 3) through the *Shh—Ptc1—Gli2—Bmp4* pathway [55,56]. *Shh* activates the signalling pathway by binding to its membrane-bound receptor, *Ptc1* (Patched), and initiates a cascade of intracellular signal transduction [57]. A downstream gene, Glioma 2 (*Gli2*) is responsible for mesenchymal proliferation. *Gli2* also induces Bone morphogenetic protein 4 (*Bmp4*). When the *Gli2* and *Bmp4* gradients decrease at the periphery of the bladder, the mesenchyme differentiates into smooth muscle (detrusor) [55,56]. This has been corroborated by Baskin and colleagues [52].

TP63 (*P63*) gene is expressed in urothelium

TP63, tumour suppressor protein 63 (*P63*) is a member of the *P53* tumour suppressor family [58] and a master regulator of stratified epithelial development [59]. It is expressed in all stratified epithelia, including the bladder urothelium and the skin overlying the external genitalia during development [14]. *TP63* is located at 3q28, and has 15 exons and 2 promoters, *TAP63* and *ΔNP63*, located upstream of exon 1 and exon 3 respectively. The 2 promoters of *P63* generate two protein isoform groups (TA and ΔN) with opposing properties. The *TAP63* isoforms (full length) are pro-apoptotic and the truncated *ΔNP63* isoforms are anti-apoptotic. These isoforms are antagonistic, and compete for the same target genes [58,60].

The urothelium is lined with a urine-proof stratified epithelium. The basal layer of urothelium expresses *p63* [23,61], and *p63* expression is also found in the intermediate cells of urothelium [62]. Several studies have shown that *p63* plays a pivotal role in the commitment to epithelial stratification, terminal differentiation, cell proliferation, cell—cell adhesion and epithelial—mesenchymal signalling [14,58—60,63—71]. Carroll et al. [72] observed that
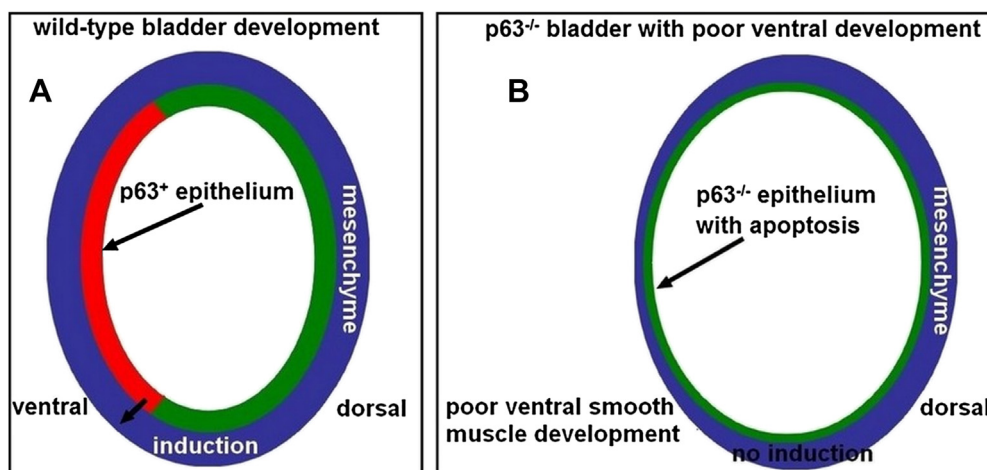
Figure 2 Epithelial—mesenchymal interaction in bladder development. A: *ΔNp63* is normally expressed in ventral bladder urothelium and is required for differentiation and production of signals required for inducing the normal developmental programme in adjacent mesenchyme. B: $p63^{-/-}$ bladder where ventral urothelium undergoes apoptosis. Absence of the epithelial signal results in the lack of mesenchymal induction and a lack of smooth muscle development.
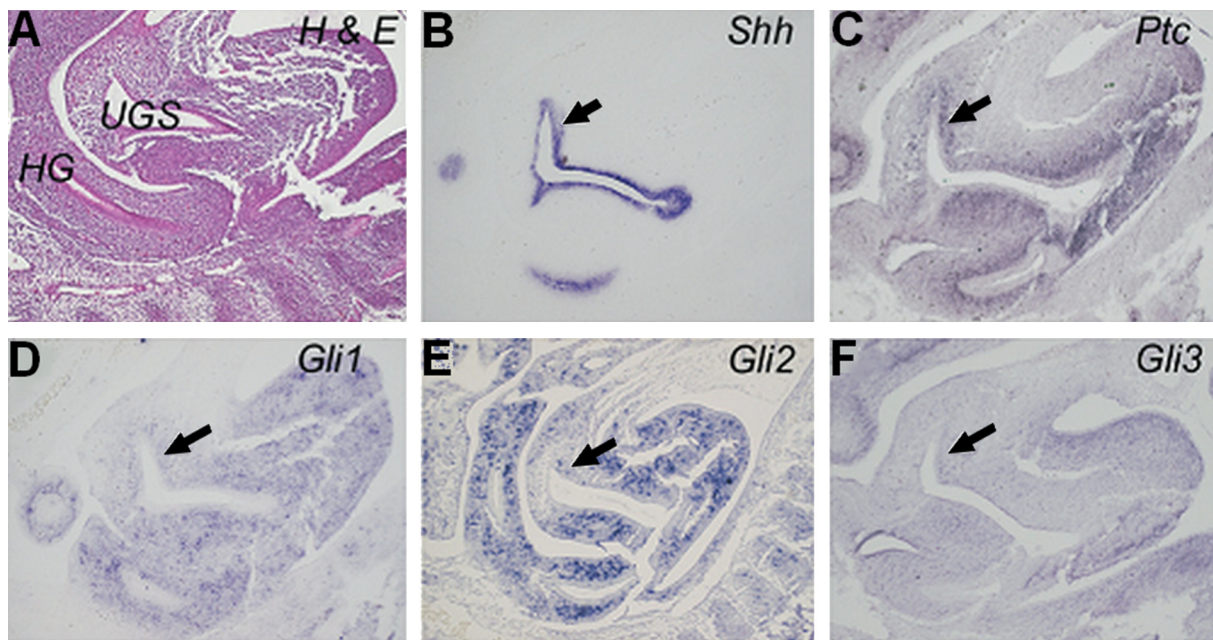
Figure 3 *Shh* signalling pathway. (A) Haematoxylin and eosin [H & E] staining of sagittal section of E12.5 WT embryo. The expressions of *Shh* (B), its receptor *Ptc* (C) and transcriptional factors *Gli1*, *Gli2* and *Gli3* (D, E and F) in E12.5 mouse urogenital sinus (UGS, primitive bladder) and hind gut (HG). Adapted with permission from Urology, 2008 [56].

knockdown of *p63* down-regulates the cell–cell adhesion genes, and conversely, that the over expression of *p63* (*TAp63* or *ΔNp63*) up-regulates the adhesion complexes. These findings implicate *p63* as a key regulator of cellular adhesion in stratified epithelial tissues.

ΔNP63 is responsible for bladder urothelial survival

P63 is expressed in urothelium during bladder organogenesis. Since urothelium is responsible for bladder mesenchymal induction, and *P63* is responsible for urothelium survival, we postulated that *P63* may play a role in bladder development. Our genetic model showed that the anti-apoptotic isoform of *p63*, *ΔNp63*, is predominantly expressed in foetal murine bladder epithelium. The *ΔNp63* isoform is expressed throughout bladder organogenesis, especially along the genital tubercle and the ventral

urothelium [14]. Without *ΔNp63*, the ventral bladder epithelium is neither stratified nor differentiated. In the knock-out mice, the bladder epithelium lacks the proper stratification (Fig. 4A and B) and undergoes apoptosis. As a result, $p63^{-/-}$ murine mutants lack stratified skin, and have a short tail, truncated limbs, and a cleft palate. They die shortly after birth, presumably due to fluid loss as a consequence of poorly stratified skin epithelium [20,23]. Loss of *ΔNp63* expression in bladder results in an increase in apoptosis, decrease in cell proliferation and a loss of smooth muscle development in ventral bladder. We also showed that the apoptotic ventral bladder epithelium in $p63^{-/-}$ mice may lead to reduced mesenchymal induction (evidenced by reduced expression of the mesenchymal markers *Msx-1* and *Fgf8*) [14]. Moreover, the apoptotic activity of ventral urothelium of $p63^{-/-}$ bladder is markedly increased (Fig. 5A and B). The reduction in mesenchymal cell proliferation and failure of smooth muscle
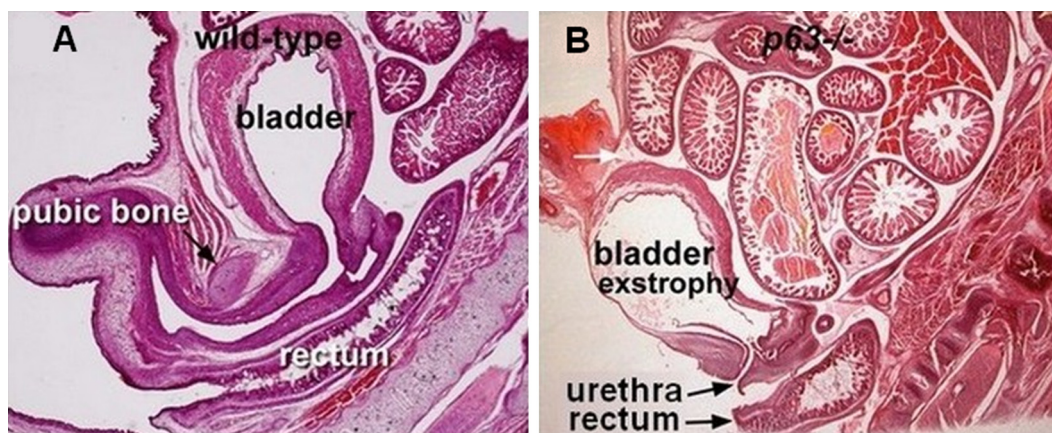

Figure 4 A: Histology of E18 wild-type foetus. B: $p63^{-/-}$ foetus with thin ventral bladder wall. Adapted with permission from Development, 2006 [14].
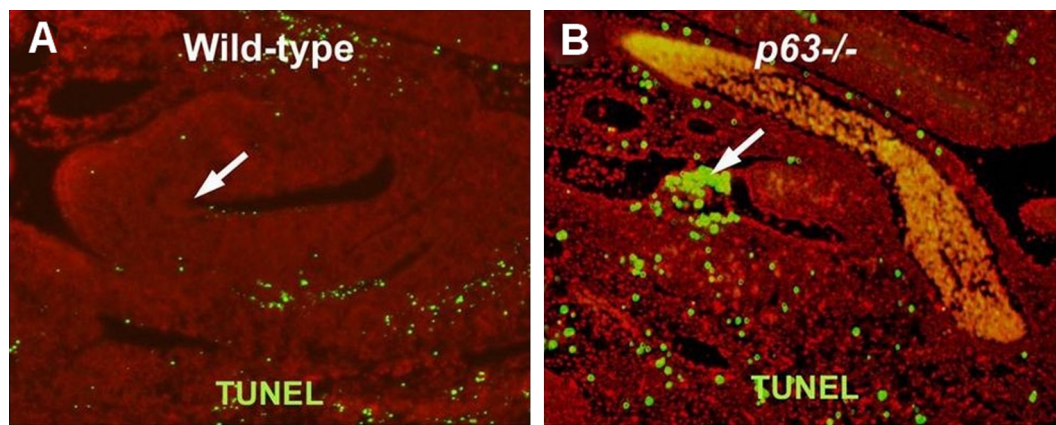
Figure 5 Fluorescent TUNEL staining showing apoptosis (arrows) of wild-type (A) and p63$^{-/-}$ urogenital sinus (200×) (B). The DAPI staining of nuclei is shown in red to increase colour contrast. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) Adapted with permission from Development, 2006 [14].

formation may in turn result in ventral midline defects, i.e. BEEC. In zebrafish, *ΔNp63* is required for the specification of the ventral tissue. Bakkers et al., found that loss of *ΔNp63* results in reduction of the ventral ectoderm [73]. The evidence above strongly suggests that *P63* is a candidate gene for human BEEC pathogenesis [14].

ΔNP63 promoter in/del polymorphism increases the risk of human BEEC

We postulated whether the findings of *p63* knock-out BEEC model could be translated into a better understanding of BEEC in humans. We first confirmed that, like our murine model, *ΔNP63* is the dominant isoform expressed in human bladder urothelium and foreskin epithelium. However, sequencing of *P63* exons in BEEC patients did not reveal any mutations, which was confirmed by the Reutter group [74,75]. Both groups also observed that *TAP63* is over-expressed in BEEC, whereas the *ΔNP63* isoform was either under-expressed or not expressed at all [74,75]. Therefore, we first focused our search on the *ΔNP63* promoter and compared the *ΔNP63* promoter sequence between 163 BEEC patients (98% are CBE) and 285 ethnicity-matched controls. Our research identified 7 single nucleotide polymorphisms (SNPs) and 4 insertion/deletion (in/del) polymorphisms. No SNPs were significantly associated with BEEC incidence. However, 3 out of 4 in/del polymorphisms were associated with a statistically significant increased risk of BEEC (Table 4) [75]. To confirm the effect of in/del

polymorphisms on transcription, we performed luciferase assays and demonstrated a consistent and statistically significant reduction in transcriptional efficiency of *ΔNP63* promoters containing the in/del polymorphisms. As increased *TAP63* isoforms may also explain increased apoptosis, we have also analyzed the *TAP63* promoter region for possible mutations. High resolution melting studies and sequencing of the TA promoters were carried out. No variants that might contribute to the risk of BEEC were identified [76].

PERP is a downstream gene of *p63* signalling pathway

Genome-Wide Expression Profiling (GWEP) study of normal and exstrophic human bladder tissues, as well as human and mouse embryologic bladder precursor tissues, identified 162 differentially expressed genes. This study also reported *PERP* to be a possible candidate gene for BEEC [77]. *PERP* was originally identified as an apoptosis-associated target of *P53*. The *PERP* promoter is bound not only by *P53*, but also by its family member, *P63*, suggesting that *PERP* is responsive to *P63* signalling [78,79]. It is a critical component of the desmosome, and is also required for maintaining the integrity of the stratified epithelia during embryogenesis [63,70,79−82]. Study of *Perp*$^{-/-}$ knock-out mice shows that *Perp* plays an essential role in cell−cell adhesion, by enabling desmosome function. Intriguingly, *Perp* specifically localizes to the desmosomes, which extends through the plasma membrane to link with

Table 4 Comparison of genotype frequencies between BEEC patients and normal controls.

Indel polymorphism	Heterozygous			Homozygous		
	p	OR	95% CI	*p*	OR	95% CI
12-bp Del. (rs6148242)	0.0291	1.763	1.057−2.942	0.0068	10.80	1.300−89.73
1-bp Ins. (rs5855273)	0.9465	—	—	<0.0001	8.794	2.872−26.93
4-bp Ins. (novel)	0.8018	—	—	0.0220	2.722	1.137−6.516

p: *p*-Value. OR: odds ratio. CI: 95% confidence intervals. Del: deletion. Ins: insertion. bp: Base pair. Data were analyzed with contingency tables, χ-square, odds ratio and 95% confidence intervals.

Reproduced with permission from PLoS Genetics, 2012 [75].

neighbouring cells [63]. In the absence of *Perp*, desmosomal complexes are defective and do not assemble properly [63]. In a study of three postnatal BEEC bladder samples (exstrophic bladder, EB) and three normal bladder (NB) samples, Qi and co-authors found that *PERP* is over-expressed in EB samples compared to NB, and is present in all embryonic mesenchyme surrounding the urogenital sinus [77]. Over expression of *PERP* may be a compensatory response to the absence, or reduced expression, of *ΔNP63* in BEEC patients (possibly due to a negative feedback loop).

Recently, Ihrie and co-workers have shown that *Perp* null mice die post-natally (within 10 days of birth) of dehydration, and develop blisters in stratified epithelia, including the skin and oral mucosa [63]. This phenotype complex is similar to that of *p63* null mice, which similarly die of dehydration shortly after birth, presumably due to absence of stratified epithelia. Wound healing in *Perp* deficient mice takes longer than the controls, suggesting that *Perp* is necessary for cell—cell adhesion during wound closure [83]. As with *P63*, *PERP* is highly expressed in stratified epithelia suggesting a possible involvement of *PERP* in BEEC pathogenesis.

Desmosome assembly and cell adherence are facilitated by *PERP*

Desmosomes provide multicellular structures and strength, and are essential for tissues undergoing mechanical stress like myocardium, bladder, gastrointestinal mucosa, and skin [84—86]. The essential element of functional desmosomes is Desmoplakin (*Dsp*). It is also a major constituent of the desmosomal plaques present in epithelial and myocardial cells of diverse species [87]. *DSP* is ubiquitously expressed [87,88], and has an important role in early embryogenesis [89]. The GWEP study of Qi and colleagues suggests potential desmosomal activity in BEEC patients, as 30% of identified genes are directly associated with desmosome [77]. They found *DSP* to be the sixth most over-expressed gene in exstrophic bladder samples. They also found that Desmin (*DES*) and Desmulin (*DMN*) are the two most under-expressed genes in BEEC [77]. These two genes encode muscle-specific intermediate filament (IF) proteins, which directly interact with the C-terminal domain of *DSP*, where *DSP* works as a connection between IFs and desmosomal plaques [88,90]. This interaction finally forms an inner dense plaque, which tethers the keratin/cytoskeletal network to the adhesion complex [86,91—93]. It has been reported that the association between *DSP* and Keratin IFs is essential for developing both epidermis and adult stratified epidermal tissue [94]. Jonkman and co-workers showed that the loss of *DSP* tail (C-terminal) lead to epidermolysis bullosa [95], which has been reported in BEEC patients [24].

Ihrie and colleagues found that *Perp* exclusively co-localizes with the desmosomal protein desmoplakin [63]. It has been reported that the N-terminus (head) of *Dsp* is required for localization [70]. From this finding, it is evident that *Perp* and desmosomes work intimately in maintaining the epithelial integrity, and that *Perp* has an important function in desmosome assembly [96]. Moreover, *Perp* null [97] and desmosomal component deficient mice [98] showed a similar phenotype of blistering. Furthermore, desmosomes in *Dsp*-null embryos do not attach to IFs, so a collapse of keratin network was found

[89]. The electron microscope study on morphometric differences of *Perp*$^{-/-}$ mice skin showed that desmosome structures were perturbed in these specimens [63], suggesting a key role of both *Perp* and desmosome in stratified epithelial development by cellular adhesion. It is obvious that the role of *PERP* and desmosome within the *P63* developmental programme is to maintain epithelial integrity [70]. This is further confirmed by down-regulation of *PERP* expression in some patients with ankyloblepharon, ectodermal dysplasia and cleft lip/palate (AEC) syndrome [99]. Ectodermal dysplasia and cleft lip/palate are part of human EEC syndrome (ectodactyly, ectodermal dysplasia and cleft lip/palate), caused by heterozygous *P63* missense mutations [21].

From the data described above, it is evident that *PERP* is the direct molecular target of *P63*, and is required for stratified epithelial function *in vivo* [63,70,81,82]. *PERP* also interacts with *DSP*, which links IFs to the cytoskeletal network [93,98]. All these studies suggest that *P63*, *PERP* and *DSP* have an important role in BEEC pathogenesis. However, the limitation of our proposed *P63* > *PERP* > desmosome pathway abnormality as an explanation of BEEC is that, it fails to examine the genome—environment interplay.

Additional factors and pathways

A GWEP study showed that *SYNPO2* and genes of *WNT* pathway may also play a role in BEEC pathogenesis [77]. *SYNPO2* is an actin-associated protein involved in actin-based motility in renal podocytes, which have a high capacity for protein synthesis and post-translational modifications [100,101]. It is possible that disruption in *SYNPO2* gene may increase the risk of BEEC by interrupting protein synthesis. WNT protein family genes have many functions, including the regulation of cell-to-cell interactions during embryogenesis. Abnormal *WNT* signalling and/or mutations in these genes can cause a range of birth defects and other diseases, including cancer [102—104]. Recent developments in *WNT* signal transduction research showed that it is essential for many critical developmental processes [105—108]. In BEEC patients, *WNT5A* gene was found to be highly up-regulated (14-fold over expression). It has been reported that in cloacal mesoderm over expression of mouse *Gli2* induced the expression of *Wnt5A*. More interestingly, *Gli2* also up-regulated *p63* expression in cloacal endoderm, suggesting a common upstream pathway is involved in the regulation of both *p63* and *Wnt5A* [109]. The literature also suggests potential involvement of the *SHH* signalling pathway in BEEC during bladder development [54,56]. However, more research is needed to confirm the involvement of *SYNPO2* and *WNT5A*.

Conclusion and working hypothesis

Based on the evidence of i) bladder epithelial—mesenchymal induction, ii) *p63* knock-out murine model and human in/del

Figure 6   *Tp63* > *Perp* > desmosome pathway during development of bladder.

polymorphism association, and iii) laboratory and human data of *P63* downstream genes, *PERP* and desmosomes, we hypothesize that the *P63* > *PERP* > desmosome signalling pathway is involved in human BEEC pathogenesis (Fig. 6).

Conflict of interest

The authors have no conflict of interest.

Funding sources

Acknowledgement

References

[1] Gearhart JP, Jeffs RD. State-of-the-art reconstructive surgery for bladder exstrophy at the Johns Hopkins Hospital. Am J Dis Child 1989;143:1475−8.

[2] Smeulders N, Woodhouse CR. Neoplasia in adult exstrophy patients. BJU Int 2001;87:623−8.

[3] Ricketts RR, Woodard JR, Zwiren GT, Andrews HG, Broecker BH. Modern treatment of cloacal exstrophy. J Pediatr Surg 1991;26:444−8. discussion 48−50.

[4] Reiner WG, Gearhart JP. Discordant sexual identity in some genetic males with cloacal exstrophy assigned to female sex at birth. N Engl J Med 2004;350:333−41.

[5] Muecke EC. The role of the cloacal membrane in exstrophy: the first successful experimental study. J Urol 1964;92:659−67.

[6] Marshall V, Muecke EC. Congenital abnormalities of the bladder. In Handbuch de Urologic. New York: Springer; 1968. p. 165.

[7] Ambrose SS, O'Brien 3rd DP. Surgical embryology of the exstrophy−epispadias complex. Surg Clin North Am 1974;54:1379−90.

[8] Bruch SW, Adzick NS, Goldstein RB, Harrison MR. Challenging the embryogenesis of cloacal exstrophy. J Pediatr Surg 1996;31:768−70.

[9] Mildenberger H, Kluth D, Dziuba M. Embryology of bladder exstrophy. J Pediatr Surg 1988;23:166−70.

[10] Patten BM, Barry A. The genesis of exstrophy of the bladder and epispadias. Am J Anat 1952;90:35−57.

[11] Slaughenhoupt BL, Chen CJ, Gearhart JP. Creation of a model of bladder exstrophy in the fetal lamb. J Urol 1996;156:816−8.

[12] Manner J, Kluth D. A chicken model to study the embryology of cloacal exstrophy. J Pediatr Surg 2003;38:678−81.

[13] Beaudoin S, Barbet P, Bargy F. Pelvic development in the rabbit embryo: implications in the organogenesis of bladder exstrophy. Anat Embryol (Berl) 2004;208:425−30.

[14] Cheng W, Jacobs WB, Zhang JJ, Moro A, Park JH, Kushida M, et al. DeltaNp63 plays an anti-apoptotic role in ventral bladder development. Development 2006;133:4783−92.

[15] Quiroz-Guerrero J, Badillo M, Munoz N, Anaya J, Rico G, Maldonado-Valadez R. Bladder augmentation in a young adult female exstrophy patient with associated omphalocele: an extremely unusual case. J Pediatr Urol 2009;5:330−2.

[16] Singh JK, Mahajan JK, Bawa M, Rao KLN. Covered exstrophy with anorectal malformation and vaginal duplication. J Indian Assoc Pediatr Surg 2011;16:26−8.

[17] Boemers TML, de Jong TPVM, Rövekamp MH, Bax NMA, van Gool JD. Covered exstrophy associated with an anorectal malformation: a rare variant of classical bladder exstrophy. Pediatr Surg Int 1994;9:438−40.

[18] Ince TA, Cviko AP, Quade BJ, Yang A, McKeon FD, Mutter GL, et al. p63 Coordinates anogenital modeling and epithelial cell differentiation in the developing female urogenital tract. Am J Pathol 2002;161:1111−7.

[19] Cadeddu JA, Benson JE, Silver RI, Lakshmanan Y, Jeffs RD, Gearhart JP. Spinal abnormalities in classic bladder exstrophy. Br J Urol 1997;79:975−8.

[20] Mills AA, Zheng B, Wang XJ, Vogel H, Roop DR, Bradley A. p63 is a p53 homologue required for limb and epidermal morphogenesis. Nature 1999;398:708−13.

[21] Celli J, Duijf P, Hamel BC, Bamshad M, Kramer B, Smits AP, et al. Heterozygous germline mutations in the p53 homolog p63 are the cause of EEC syndrome. Cell 1999;99:143−53.

[22] Rinne T, Hamel B, van Bokhoven H, Brunner HG. Pattern of p63 mutations and their phenotypes − update. Am J Med Genet A 2006;140:1396−406.

[23] Yang A, Schweitzer R, Sun D, Kaghad M, Walker N, Bronson RT, et al. p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. Nature 1999;398:714−8.

[24] Moretti G, Mazzaglia E, D'Anieri A, Merlino V, Magaudda L, Mondello MR, et al. Epidermolysis bullosa junctionalis associated with urinary bladder exstrophy: a case report. Pediatr Dermatol 1995;12:239−41.

[25] Caton AR, Bloom A, Druschel CM, Kirby RS. Epidemiology of bladder and cloacal exstrophies in New York State, 1983−1999. Birth Defects Res A Clin Mol Teratol 2007;79:781−7.

[26] Tang Y, Ma CX, Cui W, Chang V, Ariet M, Morse SB, et al. The risk of birth defects in multiple births: a population-based study. Matern Child Health J 2006;10:75−81.

[27] Anonymous. Epidemiology of bladder exstrophy and epispadias: a communication from the International Clearinghouse for Birth Defects Monitoring Systems. Teratology 1987;36:221−7.

[28] Reutter H, Boyadjiev SA, Gambhir L, Ebert AK, Rösch WH, Stein R, et al. Phenotype severity in the bladder exstrophy−epispadias complex: analysis of genetic and nongenetic contributing factors in 441 families from North America and Europe. J Pediatr 2011.

[29] Boyadjiev SA, Dodson JL, Radford CL, Ashrafi GH, Beaty TH, Mathews RI, et al. Clinical and molecular characterization of the bladder exstrophy−epispadias complex: analysis of 232 families. BJU Int 2004;94:1337−43.

[30] Ebert AK, Reutter H, Ludwig M, Rösch WH. The exstrophy−epispadias complex. Orphanet J Rare Dis 2009;4:23.

[31] Reutter H, Qi L, Gearhart JP, Boemers T, Ebert AK, Rösch W, et al. Concordance analyses of twins with bladder exstrophy−epispadias complex suggest genetic etiology. Am J Med Genet A 2007;143A:2751−6.

[32] Gearhart JP, Jeffs RD. Exstrophy−epispadias complex and bladder anomalies. In: Walsh PC, Retik AB, Vaughan ED, Wein AJ, editors. Campbell's urology. 7th ed. Philadelphia: WB Saunders Co; 1998. p. 1939−90.

[33] Wiesel A, Queisser-Luft A, Clementi M, Bianca S, Stoll C. Prenatal detection of congenital renal malformations by fetal ultrasonographic examination: an analysis of 709,030 births in 12 European countries. Eur J Med Genet 2005;48: 131—44.

[34] Hurwitz RS, Manzoni GA, Ransley PG, Stephens FD. Cloacal exstrophy: a report of 34 cases. J Urol 1987;138:1060—4.

[35] Shapiro E, Lepor H, Jeffs RD. The inheritance of the exstrophy—epispadias complex. J Urol 1984;132:308—10.

[36] Nelson CP, Dunn RL, Wei JT. Contemporary epidemiology of bladder exstrophy in the United States. J Urol 2005;173: 1728—31.

[37] Raboch J. Incidence of hypospadia and epispadia in chromatin-positive men. Andrologia 1975;7:237—9.

[38] Chipail A, Constantinescu V, Covic M, Angheloni T. Phenotypic and cytogenetic analysis of an unusual malformative syndrome (trisomy 9 p+). Rev Pediatr Obstet Ginecol Pediatr 1976;25:201—10.

[39] Lin HJ, Ndiforchu F, Patell S. Exstrophy of the cloaca in a 47, XXX child: review of genitourinary malformations in triple-X patients. Am J Med Genet 1993;45:761—3.

[40] Nicholls G, Duffy PG. Anatomical correction of the exstrophy—epispadias complex: analysis of 34 patients. Br J Urol 1998;82:865—9.

[41] Husmann DA, Vandersteen DR. Anatomy of the cloacal exstrophy. In: Gearhard JP, Matthews R, editors. The epispadias—exstrophy complex. New York: Kluwer Academic/Plenum; 1999. p. 199—206.

[42] Leonard NJ, Tomkins DJ. Diploid/tetraploid/t(1;6) mosaicism in a 17-year-old female with hypomelanosis of Ito, multiple congenital anomalies, and body asymmetry. Am J Med Genet 2002;112:86—90.

[43] Thauvin-Robinet C, Faivre L, Cusin V, Khau Van Kien P, Callier P, Parker KL, et al. Cloacal exstrophy in an infant with 9q34.1-qter deletion resulting from a de novo unbalanced translocation between chromosome 9q and Yq. Am J Med Genet A 2004;126A:303—7.

[44] Kosaki R, Fukuhara Y, Kosuga M, Okuyama T, Kawashima N, Honna T, et al. OEIS complex with del(3)(q12.2q13.2). Am J Med Genet A 2005;135:224—6.

[45] Boyadjiev SA, South ST, Radford CL, Patel A, Zhang G, Hur DJ, et al. A reciprocal translocation 46, XY, t(8;9)(p11.2; q13) in a bladder exstrophy patient disrupts CNTNAP3 and presents evidence of a pericentromeric duplication on chromosome 9. Genomics 2005;85:622—9.

[46] Ludwig M, Utsch B, Reutter H. Genetische und molekularbiologische Aspekte des Blasenekstrophie-Epispadie-Komplexes (BEEK). Der Urologe 2005;44:1037—44.

[47] Draaken M, Reutter H, Schramm C, Bartels E, Boemers TM, Ebert AK, et al. Microduplications at 22q11.21 are associated with non-syndromic classic bladder exstrophy. Eur J Med Genet 2010;53:55—60.

[48] Zaki MS, Gillessen-Kaesbach G, Vater I, Caliebe A, Siebert R, Kamel AK, et al. Bladder exstrophy and extreme genital anomaly in a patient with pure terminal 1q deletion: expansion of phenotypic spectrum. Eur J Med Genet 2012;55: 43—8.

[49] Draaken M, Mughal SS, Pennimpede T, Wolter S, Wittler L, Ebert AK, et al. Isolated bladder exstrophy associated with a de novo 0.9 Mb microduplication on chromosome 19p13.12. Birth Defects Res A Clin Mol Teratol 2013.

[50] Reutter H, Ruschendorf F, Mattheisen M, Draaken M, Bartels E, Hubner N, et al. Evidence for linkage of the bladder exstrophy—epispadias complex on chromosome 4q31.21—22 and 19q13.31—41 from a consanguineous Iranian family. Birth Defects Res A Clin Mol Teratol 2010;88:757—61.

[51] Ludwig M, Rüschendorf F, Saar K, Hübner N, Siekmann L, Boyadjiev SA, et al. Genome-wide linkage scan for bladder exstrophy—epispadias complex. Birth Defects Res A Clin Mol Teratol 2009;85:174—8.

[52] Baskin LS, Hayward SW, Young P, Cunha GR. Role of mesenchymal—epithelial interactions in normal bladder development. J Urol 1996;156:1820—7.

[53] Baskin LS, Hayward SW, Sutherland RA, DiSandro MJ, Thomson AA, Goodman J, et al. Mesenchymal—epithelial interactions in the bladder. World J Urol 1996;14:301—9.

[54] Haraguchi R, Motoyama J, Sasaki H, Satoh Y, Miyagawa S, Nakagata N, et al. Molecular analysis of coordinated bladder and urogenital organ formation by Hedgehog signaling. Development (England) 2007:525—33.

[55] Shiroyanagi Y, Liu B, Cao M, Agras K, Li J, Hsieh MH, et al. Urothelial sonic hedgehog signaling plays an important role in bladder smooth muscle formation. Differentiation 2007;75: 968—77.

[56] Cheng W, Yeung CK, Ng YK, Zhang JR, Hui CC, Kim PC. Sonic Hedgehog mediator Gli2 regulates bladder mesenchymal patterning. J Urol 2008;180:1543—50.

[57] Tasian G, Cunha G, Baskin L. Smooth muscle differentiation and patterning in the urinary bladder. Differentiation 2010: 106—17. England: 2010 International Society of Differentiation. Published by Elsevier B.V.

[58] Yang A, Kaghad M, Wang Y, Gillett E, Fleming MD, Dötsch V, et al. p63, a p53 homolog at 3q27—29, encodes multiple products with transactivating, death-inducing, and dominant-negative activities. Mol Cell 1998;2:305—16.

[59] Koster MI, Kim S, Mills AA, DeMayo FJ, Roop DR. p63 is the molecular switch for initiation of an epithelial stratification program. Genes Dev 2004;18:126—31.

[60] Koster MI, Roop DR. Transgenic mouse models provide new insights into the role of p63 in epidermal development. Cell Cycle 2004;3:411—3.

[61] Karni-Schmidt O, Castillo-Martin M, HuaiShen T, Gladoun N, Domingo-Domenech J, Sanchez-Carbayo M, et al. Distinct expression profiles of p63 variants during urothelial development and bladder cancer progression. Am J Pathol 2011; 178:1350—60.

[62] Castillo-Martin M, Domingo-Domenech J, Karni-Schmidt O, Matos T, Cordon-Cardo C. Molecular pathways of urothelial development and bladder tumorigenesis. Urologic Oncol Semin Original Invest 2010;28:401—8.

[63] Ihrie RA, Marques MR, Nguyen BT, Horner JS, Papazoglu C, Bronson RT, et al. Perp is a p63-regulated gene essential for epithelial integrity. Cell 2005;120:843—56.

[64] Laurikkala J, Mikkola ML, James M, Tummers M, Mills AA, Thesleff I. p63 regulates multiple signalling pathways required for ectodermal organogenesis and differentiation. Development 2006;133:1553—63.

[65] Senoo M, Pinto F, Crum CP, McKeon F. p63 is essential for the proliferative potential of stem cells in stratified epithelia. Cell 2007;129:523—36.

[66] Barbieri CE, Pietenpol JA. p63 and epithelial biology. Exp Cell Res United States 2006:695—706.

[67] Blanpain C, Fuchs E. p63: revving up epithelial stem-cell potential. Nat Cell Biol 2007;9:731—3.

[68] Di Como CJ, Urist MJ, Babayan I, Drobnjak M, Hedvat CV, Teruya-Feldstein J, et al. p63 expression profiles in human normal and tumor tissues. Clin Cancer Res 2002;8: 494—501.

[69] Guerrini L, Costanzo A, Merlo GR. A symphony of regulations centered on p63 to control development of ectoderm-derived structures. J Biomed Biotechnol 2011;2011:864904.

[70] Ihrie RA, Attardi LD. A new Perp in the lineup: linking p63 and desmosomal adhesion. Cell Cycle 2005;4:873—6.

[71] Koster MI, Roop DR. The role of p63 in development and differentiation of the epidermis. J Dermatol Sci 2004;34: 3—9.

[72] Carroll DK, Carroll JS, Leong CO, Cheng F, Brown M, Mills AA, et al. p63 regulates an adhesion programme and cell survival in epithelial cells. Nat Cell Biol 2006;8:551—61.

[73] Bakkers J, Hild M, Kramer C, Furutani-Seiki M, Hammerschmidt M. Zebrafish DeltaNp63 is a direct target of Bmp signaling and encodes a transcriptional repressor blocking neural specification in the ventral ectoderm. Dev Cell 2002;2:617—27.

[74] Ching BJ, Wittler L, Proske J, Yagnik G, Qi L, Draaken M, et al. p63 (TP73L) a key player in embryonic urogenital development with significant dysregulation in human bladder exstrophy tissue. Int J Mol Med 2010;26:861—7.

[75] Wilkins S, Zhang KW, Mahfuz I, Quantin R, D'Cruz N, Hutson J, et al. Insertion/deletion polymorphisms in the ΔNp63 promoter are a risk factor for bladder exstrophy epispadias complex. PLOS Genet 2012;8:e1003070.

[76] Darling T, Wilkins S, Mahfuz I, White S, Cheng W. No TAp63 promoter mutation is detected in bladder exstrophy-—epispadias complex patients. In: 46th Conference of the Pacific Association of Pediatric Surgery. Hunter Valley, Australia 7—11 April 2013.

[77] Qi L, Chen K, Hur DJ, Yagnik G, Lakshmanan Y, Kotch LE, et al. Genome-wide expression profiling of urinary bladder implicates desmosomal and cytoskeletal dysregulation in the bladder exstrophy—epispadias complex. Int J Mol Med 2011; 27:755—65.

[78] Flores ER, Tsai KY, Crowley D, Sengupta S, Yang A, McKeon F, et al. p63 and p73 are required for p53-dependent apoptosis in response to DNA damage. Nature 2002;416:560—4.

[79] Ihrie RA, Reczek E, Horner JS, Khachatrian L, Sage J, Jacks T, et al. Perp is a mediator of p53-dependent apoptosis in diverse cell types. Curr Biol 2003;13:1985—90.

[80] Marques MR, Horner JS, Ihrie RA, Bronson RT, Attardi LD. Mice lacking the p53/p63 target gene Perp are resistant to papilloma development. Cancer Res 2005;65:6551—6.

[81] Ihrie RA, Bronson RT, Attardi LD. Adult mice lacking the p53/p63 target gene Perp are not predisposed to spontaneous tumorigenesis but display features of ectodermal dysplasia syndromes. Cell Death Differ 2006;13:1614—8.

[82] Marques MR, Ihrie RA, Horner JS, Attardi LD. The requirement for perp in postnatal viability and epithelial integrity reflects an intrinsic role in stratified epithelia. J Invest Dermatol 2006;126:69—73.

[83] Beaudry VG, Ihrie RA, Jacobs SB, Nguyen B, Pathak N, Park E, et al. Loss of the desmosomal component perp impairs wound healing in vivo. Dermatol Res Pract 2010;2010:759731.

[84] Holthofer B, Windoffer R, Troyanovsky S, Leube RE. Structure and function of desmosomes. Int Rev Cytol 2007;264:65—163.

[85] Getsios S, Huen AC, Green KJ. Working out the strength and flexibility of desmosomes. Nat Rev Mol Cell Biol 2004;5: 271—81.

[86] Delva E, Tucker DK, Kowalczyk AP. The desmosome. Cold Spring Harb Perspect Biol 2009;1:a002543.

[87] Mueller H, Franke WW. Biochemical and immunological characterization of desmoplakins I and II, the major polypeptides of the desmosomal plaque. J Mol Biol 1983;163: 647—71.

[88] Borrmann CM, Mertens C, Schmidt A, Langbein L, Kuhn C, Franke WW. Molecular diversity of plaques of epithelial-adhering junctions. Ann N Y Acad Sci 2000;915:144—50.

[89] Gallicano GI, Kouklis P, Bauer C, Yin M, Vasioukhin V, Degenstein L, et al. Desmoplakin is required early in development for assembly of desmosomes and cytoskeletal linkage. J Cell Biol 1998;143:2009—22.

[90] Bornslaeger EA, Corcoran CM, Stappenbeck TS, Green KJ. Breaking the connection: displacement of the desmosomal plaque protein desmoplakin from cell—cell interfaces disrupts anchorage of intermediate filament bundles and alters intercellular junction assembly. J Cell Biol 1996;134: 985—1001.

[91] Garrod D, Chidgey M. Desmosome structure, composition and function. Biochim Biophys Acta Neth 2008:572—87.

[92] Kowalczyk AP, Stappenbeck TS, Parry DA, Palka HL, Virata ML, Bornslaeger EA, et al. Structure and function of desmosomal transmembrane core and plaque molecules. Biophys Chem 1994;50:97—112.

[93] Huen AC, Park JK, Godsel LM, Chen X, Bannon LJ, Amargo EV, et al. Intermediate filament-membrane attachments function synergistically with actin-dependent contacts to regulate intercellular adhesive strength. J Cell Biol 2002;159: 1005—17.

[94] Vasioukhin V, Bowers E, Bauer C, Degenstein L, Fuchs E. Desmoplakin is essential in epidermal sheet formation. Nat Cell Biol 2001;3:1076—85.

[95] Jonkman MF, Pasmooij AM, Pasmans SG, van den Berg MP, Ter Horst HJ, Timmer A, et al. Loss of desmoplakin tail causes lethal acantholytic epidermolysis bullosa. Am J Hum Genet 2005;77:653—60.

[96] Bektas M, Rubenstein DS. Perp and pemphigus: a disease of desmosome destabilization. J Invest Dermatol 2009;129: 1606—8.

[97] Koch PJ, Mahoney MG, Ishikawa H, Pulkkinen L, Uitto J, Shultz L, et al. Targeted disruption of the pemphigus vulgaris antigen (desmoglein 3) gene in mice causes loss of keratinocyte cell adhesion with a phenotype similar to pemphigus vulgaris. J Cell Biol 1997;137:1091—102.

[98] Green KJ, Gaudry CA. Are desmosomes more than tethers for intermediate filaments? Nat Rev Mol Cell Biol 2000;1:208—16.

[99] Beaudry VG, Pathak N, Koster MI, Attardi LD. Differential PERP regulation by TP63 mutants provides insight into AEC pathogenesis. Am J Med Genet A 2009;149A:1952—7.

[100] Lowik MM, Groenen PJ, Levtchenko EN, Monnens LA, van den Heuvel LP. Molecular genetic analysis of podocyte genes in focal segmental glomerulosclerosis — a review. Eur J Pediatr 2009;168:1291—304.

[101] Mundel P, Heid HW, Mundel TM, Kruger M, Reiser J, Kriz W. Synaptopodin: an actin-associated protein in telencephalic dendrites and renal podocytes. J Cell Biol 1997;139: 193—204.

[102] Angers S, Moon RT. Proximal events in Wnt signal transduction. Nat Rev Mol Cell Biol 2009;10:468—77.

[103] Logan CY, Nusse R. The Wnt signaling pathway in development and disease. Annu Rev Cell Dev Biol 2004;20: 781—810.

[104] Andre P, Wang Q, Wang N, Gao B, Schilit A, Halford MM, et al. The Wnt coreceptor Ryk regulates Wnt/planar cell polarity by modulating the degradation of the core planar cell polarity component Vangl2. J Biol Chem 2012;287.

[105] Gao B, Song H, Bishop K, Elliot G, Garrett L, English MA, et al. Wnt signaling gradients establish planar cell polarity by inducing Vangl2 phosphorylation through Ror2. Dev Cell 2011; 20:163—76.

[106] Heisenberg CP, Tada M, Rauch GJ, Saude L, Concha ML, Geisler R, et al. Silberblick/Wnt11 mediates convergent extension movements during zebrafish gastrulation. Nature 2000;405:76—81.

[107] Qian D, Jones C, Rzadzinska A, Mark S, Zhang X, Steel KP, et al. Wnt5a functions in planar cell polarity regulation in mice. Dev Biol 2007;306:121—33.

[108] Rauch GJ, Hammerschmidt M, Blader P, Schauerte HE, Strahle U, Ingham PW, et al. Wnt5 is required for tail formation in the zebrafish embryo. Cold Spring Harb Symp Quant Biol 1997;62:227—34.

[109] Liu G, Moro A, Zhang JJ, Cheng W, Qiu W, Kim PC. The role of Shh transcription activator Gli2 in chick cloacal development. Dev Biol 2007;303:448—60.

PLOS | GENETICS

Insertion/Deletion Polymorphisms in the *ΔNp63* Promoter Are a Risk Factor for Bladder Exstrophy Epispadias Complex

Simon Wilkins[1][¤]*, Ke Wei Zhang[1], Istiak Mahfuz[1], Renaud Quantin[1], Nancy D'Cruz[1], John Hutson[2], Michael Ee[3], Darius Bagli[4], Karen Aitken[4], Fion Nga-Yin Fong[5], Patrick Kwok-Shing Ng[5], Stephen Kwok-Wing Tsui[5], Wendy Yin-Wan Fung[5], Tahmina Banu[6], Atul Thakre[7], Kaid Johar[7], Enrique Jaureguizar[8], Long Li[9], Wei Cheng[1,10,11]*

1 Monash Institute of Medical Research, Faculty of Medicine, Nursing, and Health Sciences, Monash University, Melbourne, Australia, 2 Department of Paediatric Urology, University of Melbourne, Melbourne, Australia, 3 Women's and Children's Clinical Services, Royal Hobart Hospital, Hobart, Australia, 4 Division of Urology, Hospital for Sick Children and University of Toronto, Toronto, Canada, 5 School of Biomedical Sciences, The Chinese University of Hong Kong, Hong Kong, China, 6 Department of Pediatric Surgery, Chittagong Medical College and Hospital, Chittagong, Bangladesh, 7 Iladevi Cataract and Intraocular Lens Research Centre, Civil Hospital, Ahmedabad, India, 8 Department of Urology, Hospital Universitario La Paz, Madrid, Spain, 9 Department of Surgery, Capital Institute of Pediatrics, Beijing, China, 10 Department of Paediatrics, Department of Surgery, Southern Medical School, Faculty of Medicine, Nursing, and Health Sciences, Monash University, Melbourne, Australia, 11 Department of Paediatric Surgery, Monash Children's, Southern Health, Melbourne, Australia

Abstract

Bladder exstrophy epispadias complex (BEEC) is a severe congenital anomaly; however, the genetic and molecular mechanisms underlying the formation of BEEC remain unclear. *TP63*, a member of *TP53* tumor suppressor gene family, is expressed in bladder urothelium and skin over the external genitalia during mammalian development. It plays a role in bladder development. We have previously shown that $p63^{-/-}$ mouse embryos developed a bladder exstrophy phenotype identical to human BEEC. We hypothesised that *TP63* is involved in human BEEC pathogenesis. RNA was extracted from BEEC foreskin specimens and, as in mice, *ΔNp63* was the predominant *p63* isoform. *ΔNp63* expression in the foreskin and bladder epithelium of BEEC patients was reduced. DNA was sequenced from 163 BEEC patients and 285 ethnicity-matched controls. No exon mutations were detected. Sequencing of the *ΔNp63* promoter showed 7 single nucleotide polymorphisms and 4 insertion/deletion (indel) polymorphisms. Indel polymorphisms were associated with an increased risk of BEEC. Significantly the sites of indel polymorphisms differed between Caucasian and non-Caucasian populations. A 12-base-pair deletion was associated with an increased risk with only Caucasian patients (p = 0.0052 Odds Ratio (OR) = 18.33), whereas a 4-base-pair insertion was only associated with non-Caucasian patients (p = 0.0259 OR = 4.583). We found a consistent and statistically significant reduction in transcriptional efficiencies of the promoter sequences containing indel polymorphisms in luciferase assays. These findings suggest that indel polymorphisms of the *ΔNp63* promoter lead to a reduction in p63 expression, which could lead to BEEC.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Bladder exstrophy epispadias complex (BEEC; MIM600057) is a serious congenital anomaly present in 1 in 36,000 live births [1]. BEEC is manifested as a cluster of ventral midline defects including: 1) ventral bladder and abdominal wall defects, 2) epispadias (split external genitalia), 3) separation of the pubic bones and the rectus abdominis muscles, 4) exomphalos, and 5), ventrally displaced anus [2]. Treatment of BEEC requires a series of major reconstructive surgeries with high morbidity rate. Without treatment, the affected babies continuously leak urine through the bladder defects, resulting in skin excoriation, a progressive decline in renal function, a constant stench, and severe psychosocial strain for both patients and parents. Untreated BEEC patients develop chronic bladder mucosal irritation and have a 700-fold increased chance of developing bladder cancer. The genetic and molecular mechanisms underlying the formation of BEEC remain unclear. To date, studies have shown an increased incidence of BEEC in children of older mothers [3], in Caucasian populations [4], and a familial genetic component of pathogenesis [5]. Chromosomal abnormalities have been suggested as possible cause of BEEC and a genome wide linkage study suggested more than one gene was involved [6]. Among siblings and offspring of BEEC patients, the risk of BEEC increase dramatically from 1 in

Author Summary

Bladder exstrophy epispadias complex is a severe congenital abnormality. The affected babies' bladders are born open, leaking urine constantly. Treatment involves multiple major reconstructive surgeries and the need for lifelong care for the complications of the disease. Although a number of studies have suggested a genetic cause of the disease, the genetic and molecular mechanism underlying the formation of BEEC remains unknown. One gene, *TP63*, plays a crucial role in the early bladder development. Two different genetic promoters of *TP63* produce different forms of the protein with opposing properties. We have shown mice lacking *p63* displayed a deformity complex identical to human BEEC. There are no genetic mutations in the p63 protein in BEEC, so genetic variants in the promoter could alter protein expression. Our hypothesis was that loss of p63 expression due to sequence polymorphisms in a promoter is a risk factor for BEEC. We found promoter sequence variants that were statistically associated with the disease and the sequence variant location varied between Caucasian and non-Caucasian patients. This is particularly important as Caucasian populations have a higher risk of BEEC. These findings provide an explanation of BECC and a base for further study of *TP63* related genes in this disease.

10,000–50,000 to 1 in 100 and 1 in 70 respectively, representing a 500-fold increase in incidence [7]. The concordance rate among the monozygotic twins is much higher than that of dizygotic twins (62% vs. 11%), representing a 4500-fold increase in incidence compared to that of the general population [8]. Brought together these data clearly suggest a genetic component in BEEC pathogenesis.

A member of the p53 tumor suppressor family, the protein p63, is expressed in all stratified epithelia, including the bladder urothelium and the skin overlying the external genitalia during development [9]. The protein p63 plays a key role in initiating epithelial stratification during development [10]. Expression of *TP63* is regulated by two promoters, *TAp63* and *ΔNp63*, located upstream to exon 1 and exon 3 respectively [11]. *TAp63* protein isotypes are pro-apoptotic whereas *ΔNp63* isotypes are anti-apoptotic and both isotypes may compete for the same set of target genes [11]. Mouse embryos lacking *p63* have thin, non-stratified skin, a short tail, truncated limbs, and cleft palate and are perinatally fatal [11,12]. *p63* is required for stratification of epithelia, including the bladder urothelium [10]; furthermore the epithelial-mesenchymal interaction is instrumental in bladder mesenchymal (smooth muscle) development [13]. We have demonstrated that $p63^{-/-}$ mouse embryos exhibit ventral midline defects identical to those of human BEEC, including ventral bladder and abdominal wall defects, separation of external genitalia, separation of pubic bones and rectus abdominis muscles, exomphalos, and ventrally translocated anus [9]. *p63* is expressed in the bladder urothelium during development, *TAp63* at an earlier stage (E9–11) and *ΔNp63* later (E11–19). *ΔNp63* is preferentially expressed along the ventral midline in the epithelium overlying the genital tubercle and ventral bladder. Moreover, the apoptotic activity of ventral urothelium of $p63^{-/-}$ bladder is markedly increased whereas cell proliferation is much reduced. We believe that bladder epithelial apoptosis and failure of induction of the adjacent mesenchyme lead to the development of a BEEC-like phenotype in mice [9]. This study investigates whether the loss of p63 expression due to genetic variants in the *ΔNp63* is a risk factor for BEEC. In this study we show the

expression of the two different *TP63* promoters in BEEC patient tissue, and explore sequence of the *ΔNp63* promoter in BEEC patients and controls. Our studies suggest that *ΔNp63* is the dominant promoter in human tissue and its expression is significantly reduced in BEEC tissue in early bladder formation. We also find many sequence variants within the *ΔNp63* promoter and three insertion/deletion polymorphisms were significantly associated with increased risk of BEEC. Furthermore the sites of these polymorphisms varied between Caucasian and non-Caucasian patients.

Results/Discussion

To establish if *TP63* plays a role in human bladder exstrophy, we first confirmed by real time-PCR that normal human foreskin (from circumcision) predominantly expresses *ΔNp63* mRNA, whereas the *TAp63* isoform was expressed at lower levels (Figure 1A). Compared with normal controls, the *ΔNp63* expression in the dorsal foreskin (adjacent to the epispadias) was decreased and also decreased compared to the patient's own ventral foreskin (opposite side) (Figure 1B). Down-regulation of *ΔNp63* appears to be mainly in the foreskin mucosa (non-keratinized epithelium) in an 8-year old BEEC patient (Figure 1C). Strikingly, in a tissue sample taken from the ventral bladder of a 2 day old Caucasian BEEC patient, *ΔNp63* expression was decreased whereas *TAp63* was increased compared with normal controls (Figure 1D and 1E). Tissue taken from the bladder of a<1 year old Caucasian BEEC patient also had deceased *ΔNp63* and increased *TAp63* compared with normal controls (Figure 1F). Expression of *ΔNp63* and *TAp63* in older Caucasian BEEC patients showed more varied expression (Figure 1G). The older patients haven undergone Mitroffanof procedures (using appendix as a conduit for bladder catheterization) and bladder augmentation (using small bowel patch to increase bladder volume). The irritation from small bowel mucus secretion and various degree of cystitis (from bacteria introduced by catheters) may affect *TP63* expression in the urothelia sampled. The genotypes of patient samples from Figure 1D–1G are shown in Table 1. We found that *ΔNp63* expression in some of the BEEC patients' urothelia is reduced. Although the post-natal expression does not necessarily represent that of the developing embryos and the sample size of post-natal bladder urothelia will not be large enough to draw statistically convincing conclusion, our data does demonstrate that *TP63* is expressed in neonatal bladder urothelium and neonatal foreskin of normal individuals and BEEC patients. The possible reduced *ΔNp63* expression during development could be one of the possible mechanisms of BEEC pathogenesis.

Our data corroborate a recently published study where three out of the five BEEC patients had no *ΔNp63* expression detected in their bladders [14]. Our results showing reduced levels of *ΔNp63* expression led us to examine if any mutations were present within the coding sequence of the *TP63* gene in BEEC patients. We therefore sequenced all 15 exons of *TP63* gene [11] in 15 BEEC patients but found no mutations (data not shown). This finding confirms that of a recent study where no exon mutations were found in a study of 22 BEEC patients [14].

To explain reduced *ΔNp63* expression in the absence of any exon mutation, the *ΔNp63* promoter (2700 nucleotides upstream of exon 3) was sequenced in BEEC patients and normal controls. DNA was extracted from buccal swab samples from 163 BEEC patients and 285 ethnicity-matched controls from India, Bangladesh, China, Australia, Spain, Canada and USA. We found 7 single nucleotide polymorphisms (SNPs; 2 of which were novel ss#541026548 and ss#541027120) and 4 insertion/deletion
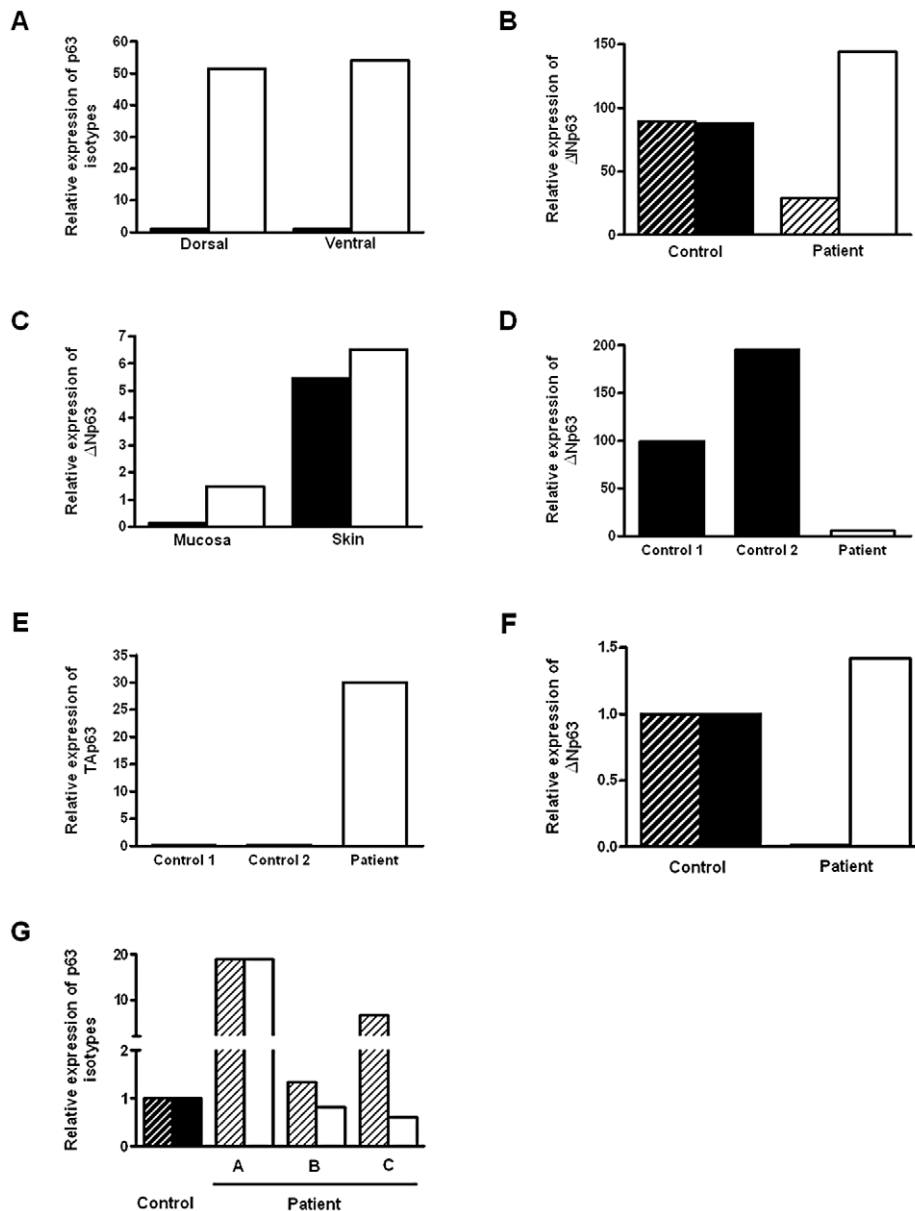
Figure 1. Expression of *p63* isoforms in BEEC patient tissue and controls. (A) Real-time qPCR of *ΔNp63* (white column) and *TAp63* (black column) expression dorsal and ventral foreskin from normal control (2 yr. male). (B) Real time qPCR of *ΔNp63* expression dorsal (black hatched column) and ventral foreskin (black column) from 2 normal controls and dorsal (hatched column) and ventral foreskin (white column) from a BEEC patient. (C) Real-time qPCR of *ΔNp63* expression in from BEEC patient (8 yr. male) dorsal (black column) and ventral (white column) foreskin mucosa and skin. (D) Real-time qPCR showing reduced *ΔNp63* mRNA expression in BEEC bladder epithelium compared with 2 normal controls. (E) Increased *TAp63* mRNA expression in BEEC bladder epithelium compared with 2 normal controls. (F) Real-time qPCR showing reduced *ΔNp63* mRNA and increased *TAp63* mRNA expression in bladder mucosa from a Caucasian BEEC patient (<1 yr. male) compared with normal controls (normalized to 1). (G) Real-time qPCR showing *ΔNp63* and *TAp63* mRNA expression in bladder mucosa from a three Caucasian BEEC patients (A, 13 yr. male; B, 14 yr. female; C, 6 yr. male) compared with normal controls (normalized to 1).
doi:10.1371/journal.pgen.1003070.g001

(indel; 1 novel ss#541028600) polymorphisms in the *ΔNp63* promoter in both BEEC and control sequences (Table 2). There were no significant deviations from Hardy-Weinberg equilibrium in the controls' genotype distributions when tested by a goodness of fit χ-square test [15]. In a number of cases we were unable to obtain complete sequences from patient or control samples.

Significantly, three indel polymorphisms (rs6148242, rs5855273, ss#541028600) were associated with increased risk of BEEC (Table 3). The indel polymorphisms showed no linkage disequilibrium (12 base pair (bp) deletion (del) vs. 4 bp insertion (ins),

$r^2 = 0.0754$; 1 bp vs 4 bp, $r^2 = 0.145$; 12 bp vs 1 bp, $r^2 = 0.0273$) [16]. We stratified the cohort into Caucasian and non-Caucasian groups based on the increased incidence in Caucasian populations [4]. The non-Caucasian group contained mainly Indian and Chinese patients and there were no significant differences in the three indel polymorphisms between the Indian and Chinese patients. A significant ethnicity bias of *ΔNp63* promoter base pair (bp) indel polymorphisms was observed in patients compared to their ethnically matched controls. The 12 bp deletion (rs6148242) was only associated with increased risk of BEEC among Caucasians

Table 1. Genotypes of BEEC patient samples used in real-time expression experiments.

Promoter Position	Genotype of patient samples used in real-time expression				
	Figure 1D/1E	Figure 1F	Figure 1G Patient A	Figure 1G Patient B	Figure 1G Patient C
−2657	A/A	A/A	A/A	A/A	A/A
−2651	C/C	C/C	C/C	C/T	C/T
−2293 to −2282	TCCAGAATCTT/-	TCCAGAATCTT/TCCAGAATCTT	TCCAGAATCTT/TCCAGAATCTT	TCCAGAATCTT/-	TCCAGAATCTT/-
−1944	C/C	C/C	C/C	C/C	C/C
−1287	T/-	T/T	-/-	T/-	T/-
−1209	T/T	T/T	C/C	C/C	C/C
−1059	C/C	C/C	C/C	C/C	C/C
−71	AGAG/-	AGAG/AGAG	AGAG/-	AGAG/AGAG	AGAG/-

doi:10.1371/journal.pgen.1003070.t001

Table 2. Single nucleotide polymorphisms and insertion/deletion polymorphisms in *ΔNp63* promoter found in BEEC patients and normal controls.

Ch. 3 position	NCBI rs or ss# number	Promoter position	Genotype	Numbers		Frequency (%)	
				BEEC	Controls	BEEC	Controls
190987627	ss#541026548*	−2657	A/A	59	147	86.8	68.7
	Novel SNP		A/T	8	61	11.8	28.5
			T/T	1	6	1.5	2.8
190987634	rs2138247	−2651	C/C	53	137	65.4	66.5
			C/T	28	65	34.6	31.6
			T/T	0	4	0.0	1.9
190987853	ss#541027120	−2431	G/G	46	77	64.8	65.3
	Novel SNP		G/A	20	37	28.2	31.4
			A/A	5	6	7.0	3.4
190987991	rs6148242	−2293 to −2282	TCCAGAATCTTT/TCCAGAATCTTT	70	108	56.0	71.5
	12 bp del.		TCCAGAATCTTT/-	48	42	38.4	27.8
			-/-	7	1	5.6	0.7
190988024	rs1554130	−2260	C/C	23	25	74.2	78.1
			C/T	7	7	22.6	21.9
			T/T	1	0	3.2	0.0
190988340	rs55803942	−1944	C/C	33	81	64.7	47.6
			C/T	16	75	31.4	44.1
			T/T	2	14	3.9	8.2
190988997	rs5855273	−1287	-/-	51	78	47.2	57.8
	1 bp ins.		T/-	34	53	31.5	39.3
			T/T	23	4	21.3	3.0
190989073	rs1464118	−1209	T/T	29	77	54.7	44.3
			C/T	18	70	34.0	40.2
			C/C	6	27	11.3	15.5
190989223	rs1464117**	−1059	C/C	29	67	58	41.4
			A/C	13	65	26	40.1
			A/A	8	30	16	18.5
190990212	rs34201045	−71	-/-	83	95	84.7	88.8
	2 bp ins.		-/AG	15	11	15.3	10.3

Table 2. Cont.

Ch. 3 position	NCBI rs or ss# number	Promoter position	Genotype	Numbers		Frequency (%)	
				BEEC	**Controls**	**BEEC**	**Controls**
			AG/AG	0	1	0.0	0.9
190990212	ss#541028600	−71	-/-	36	49	36.7	43.8
	Novel 4 bp ins.		AGAG/-	42	53	42.9	47.3
			AGAG/AGAG	20	10	20.4	8.9

Indel polymorphism/SNP rs numbers from reference sequences of p63 GenBank (Locus: NC_000003). Sumitter SNP (ss) accession numbers from NCBI. Data were sourced from 108 Caucasian patients, 126 Caucasian controls, 55 Non-Caucasian patients and 159 Non-Caucasian controls.
*Heterozygous Novel SNP ss#541026548 (Ch. 3 position 190987627) was associated with a decreased risk of BEEC (p = 0.0071, OR = 2.357, CI 95% = 1.256–4.426).
**Heterozygous SNP rs1464117 (Ch. 3 position 190989223) was associated with a decreased risk of BEEC (p = 0.0379, OR = 2.164, CI 95% = 1.035–4.527). Data were analyzed with contingency tables, χ-square, odds ratio and 95% confidence intervals.
doi:10.1371/journal.pgen.1003070.t002

but not in non-Caucasian patients. Conversely, the novel homozygous 4 bp insertion (ss#541028600) was found to be only associated increased risk of BEEC amongst non-Caucasians. The homozygous 1 bp insertion polymorphism (rs5855273) was significant across both cohorts (Table 4; Table 5). Interestingly, two heterozygous SNPs (ss#541026548, rs1464117) were associated with decreased BEEC risk (p = 0.0043 Odds Ratio (OR) = 3.060 and p = 0.0379 OR = 2.164 respectively) (Table 2; Figure 2).

To assess the effect of different indel polymorphisms on transcriptional efficiency, we performed luciferase assays to test the *ΔNp63* promoters. We sub-cloned four variations of indel polymorphisms (Table 6) into the pGL3 luciferase vector and transfected human embryonic kidney (HEK-293) cells. We found a consistent and statistically significant reduction of the transcriptional efficiencies of the promoters containing indel polymorphisms compared with the control sequence lacking indel polymorphisms (Figure 3; Table 7).

A number of studies suggest a genetic component in the etiology of BEEC. Precise regulation of *ΔNp63* expression is required for development and differentiation of the ventral bladder urothelium during human development. In addition, $p63^{-/-}$ mice also have phenotypes identical to the associated anomalies of BEEC patients (Table 8). The only genetic model of the BEEC provides valuable insight into the possible mechanism of BEEC in humans. The mouse model showed that the anti-apoptotic isotype of p63, *ΔNp63*, is predominantly expressed in fetal murine bladder

Table 3. Comparison of genotype frequencies between BEEC patients and normal controls.

Indel polymorphism	Heterozygous			Homozygous		
	p	OR	95% CI	p	OR	95% CI
12-bp Del. (rs6148242)	0.0291	1.763	1.057–2.942	0.0068	10.80	1.300–89.73
1-bp Ins. (rs5855273)	0.9465	-	-	<0.0001	8.794	2.872–26.93
4-bp Ins. (ss#541028600)	0.8018	-	-	0.0220	2.722	1.137–6.516

Data were sourced from 108 Caucasian patients, 126 Caucasian controls, 55 Non-Caucasian patients and 159 Non-Caucasian controls. Actual numbers for each genotype are shown in Table 2. P: p-value. OR: odds ratio CI: 95% confidence intervals. Del: deletion. Ins: insertion. bp: base pair. Data were analyzed with contingency tables, χ-square, odds ratio and 95% confidence intervals.
doi:10.1371/journal.pgen.1003070.t003

epithelium. Loss of *ΔNp63* expression in $p63^{-/-}$ mice increased apoptosis in the ventral bladder epithelium causing reduced mesenchymal induction (indicated by reduced expression of the mesenchymal markers Msx-1 and Fgf8) [9]. The reduction in mesenchymal cell proliferation and failure of smooth muscle formation in turn resulted in ventral midline defects, i.e. BEEC [9]. Brought together this evidence strongly suggests that *TP63* is a candidate gene for human BEEC and dysregulation of p63 expression could be a contributing factor in BEEC pathogenesis.

Our murine model shows that the urogenital tubercle, the embryological origin of the foreskin, is one of few anatomical sites where *ΔNp63* is expressed [9]. Foreskin in our human study is one of p63 expressing tissues and is logistically more accessible. Our study has shown that *ΔNp63* is the dominant isotype in human foreskin, that *ΔNp63* expression was reduced in the urothelium of BEEC patients, and expression of the pro-apoptotic *TAp63* was increased. No exon mutations have been discovered in *TP63* suggesting that dysregulation may be caused by other regions of the gene such as the promoter regions. Sequence variants in promoters have been shown to be risk factors in numerous diseases including pneumoconiosis [17], auto-immune diseases [18], asthma [19], and β-thalassaemia [20]. A six-nucleotide promoter polymorphism has been described as a risk factor in multiple cancers such as lung, esophagus, stomach, colorectum, breast and cervix in Chinese populations [21]. However, studies on breast, prostate, and colorectal cancer in European and USA populations have not shown the same association with cancer risk [22–24]. Sequencing of the *ΔNp63* promoter region revealed three indel polymorphisms in the 163 BEEC patients associated with a statistically significant increase in BEEC risk. The prevalence and the role of indel polymorphisms differed between Caucasian and non-Caucasian ethnic populations. While further studies are required to explain this, we speculate ethnicity-specific polymorphisms of up-stream or down-stream genes may further modify the final effect of the *ΔNp63* promoter polymorphisms. One possible explanation may be that the polymorphisms may interfere with transcription binding sites which may differ in Caucasian and non-Caucasian populations. An example is the deletion of six-nucleotides in the CASP8 promoter destroys the Sp1 transcription factor binding site [21]. We searched transcription factor binding sites in the *ΔNp63* promoter using MATCH software. Computational analysis predicted the 12 bp indel might affect binding of Hand1, GATA-1,2,3,6, Gfi1, LEF1, TCF1 and SOX10, whereas the 4 bp indel may affect binding of SREBP, EGR and CBF transcription factors. Further studies would be needed to verify these suggested interactions.

Table 4. Comparison of each of the BEEC patient groups against the ethnically-matched controls.

	Caucasians vs. controls	Non-Caucasians vs. controls
Heterozygous 12-bp deletion (rs6148242)	p = 0.0071, OR = 2.357, CI: 1.256–4.426	p = 0.2325
Homozygous 12-bp deletion (rs6148242)	p = 0.0052, OR = 18.33, CI: 0.9869–340.6	p = 0.1754
Homozygous 1-bp insertion (rs5855273)	p<0.0001, OR = 44.10, CI: 2.549–762.8	p = 0.0368, OR = 4.513, CI: 1.036–19.66
Homozygous 4-bp insertion (ss#541028600)	p = 0.0892	p = 0.0259, OR = 4.583, CI: 1.160–18.10

Frequencies for each genotype are shown in Table 5. P: p-value. OR: odds ratio CI: 95% confidence intervals. Del: deletion. Ins: insertion. bp: base pair. Data were analyzed with contingency tables, χ-square, odds ratio and 95% confidence intervals.
doi:10.1371/journal.pgen.1003070.t004

In this study, indel polymorphisms in *ΔNp63* promoter are associated with increased risk of BEEC, most likely due to decreased transcriptional efficiency and therefore decreased expression of anti-apoptotic *ΔNp63* isoforms during bladder development. The consequences of decreased expression of p63 isotypes could be complex and result in stimulation or negative regulation of a number of genes. One such gene, *PERP*, is a p63 regulated gene central to epithelial integrity and homeostasis [25]. In skin deficient in PERP, desmosomal deficits are observed in addition to epithelial blistering [25]. Genome-wide expression profiling has revealed a great number of desmosomal-linked genes in addition to *PERP*, *SYNOP2*, and the Wnt pathway as potentially contributing to BEEC etiology [26]. P63 mutation has been implicated in human disease such as Ectodactyly, Ectodermal dysplasia and Cleft palate/lip syndrome (EEC) [27]. Maas et al., reported that, out of 14 members of a family with EEC syndrome, 10 suffered from micturition problem [28]. After reviewing 24 previous reports of urogenital anomalies in EEC patients, Maas concluded that structural anomalies of urogenital system may be part of EEC syndrome. The report included a histological figure of "atrophic" urothelium, which showed area of thin urothelium, reminiscent of the non-stratified bladder epithelium of $p63^{-/-}$ BEEC knockout model [28]. The association is further supported by a report from Chuangsuiwanich et al., which showed a case of EEC foetus with markedly hypoplastic bladder, lined at it lower part with thin urothelium [29]. Sub-clinical variations of BEEC may be more prevalent in EEC patients and/or other conditions than previously thought.

We conclude that insertion/deletion polymorphisms of *ΔNp63* promoter are associated with increased risk of Bladder Exstrophy Epispadias Complex. We believe our results provide a base for further study of p63 related genes in this debilitating condition

Materials and Methods

DNA Samples

The Research Ethics Committee of Royal Children's Hospital, Melbourne, Australia reviewed and approved the study. Informed consents were obtained from the patients, parents/guardians and the normal controls. Buccal swabs were used to collect DNA extracted with the BuccalAmp DNA Extraction Kit (Epicentre Biotechnologies, Madison, WI, USA). All samples were codified and subject identities kept confidential. Collaborators from overseas centres (Canada, USA, Spain, India, Bangladesh, China, and Malaysia) had the study approved by their institutional ethic committees prior to sample collection.

Table 5. Frequency (%) of indel genotypes in Caucasian and non-Caucasian groups.

Frequency (%) of indel genotypes in Caucasian and non-Caucasian groups

Indel polymorphism	Genotype	Caucasians		Non-Caucasians	
		BEEC	Controls	BEEC	Controls
12-bp Del. (rs6148242)	TCCAGAATCTTT/TCCAGAATCTTT	48.0	71.3	78.0	75.0
	TCCAGAATCTTT/-	45.4	28.7	17.1	25.0
	-/-	6.5	0.0	4.9	0.0
1-bp Ins. (rs5855273)	-/-	52.2	63.9	33.3	40.0
	T/-	27.5	36.0	38.4	52.5
	T/T	20.3	0.0	28.2	7.5
4-bp Ins. (ss#541025600)	-/-	20.7	50.0	15.4	35.5
	AGAG/-	50.9	43.1	50.0	45.1
	AGAG/AGAG	9.4	6.9	38.4	20.0

Data were sourced from 77 Caucasian patients, 94 Caucasian controls, 41 Non-Caucasian patients and 44 Non-Caucasian controls. Del: deletion. Ins: insertion. bp: base pair.
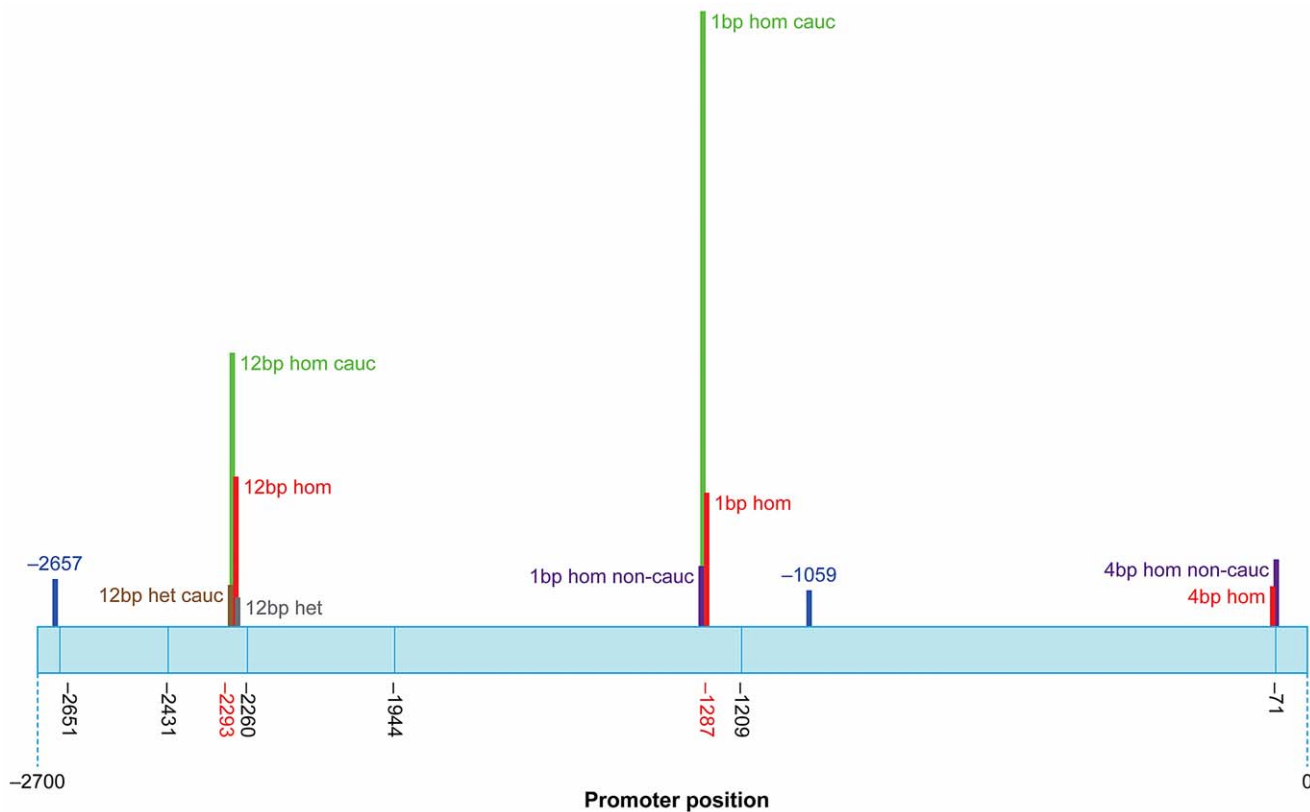doi:10.1371/journal.pgen.1003070.t005

Figure 2. Schematic representation of the human *ΔNp63* promoter showing SNP and indel locations. The height of the bars represents the Odds Ratio (OR). Promoter positions of SNPs and indel that are not associated with any significant change in risk of BEEC are shown in black. OR represented: 12 bp heterozygous (het) all patients (grey; OR = 1.763); 12 bp homozygous (hom) all patients (red; OR = 10.80); 1 bp hom all patients (red; OR = 8.794); 4 bp hom all patients (red; OR = 2.722); 12 bp het Caucasian patients (brown; OR = 2.357); 12 bp hom Caucasian patients (green; OR = 18.33); 1 bp hom Caucasian patients (green; OR = 44.10); 1 bp hom Non-Caucasian patients (purple; OR = 4.513); 4 bp hom Non-Caucasians patients (purple; OR = 4.583). Two SNPs associated with decreased risk of BEEC are shown in blue: SNP −1059 (OR = 2.164); SNP −2657 (OR = 3.060). doi:10.1371/journal.pgen.1003070.g002

PCR and Sequencing

Exon PCR amplification was performed using published protocols [27] and in-house primers designed using Primer 3 and Premier Netprimer (Premier BioSoft International). Four pairs of primers were designed to complete the PCR and sequencing of the 2785 bp promoter in four parts. The extent of the *ΔNp63* promoter by aligning and comparing the 5′ upstream regions of exon 3 of *P63* gene sequences from humans, mouse and pig. The most conserved region 2785 bp (−2696 to +89) was selected as the putative region of *ΔNp63*. Sequencing was carried out in both directions however in a number of cases sequences from patients or controls were unreadable using a number of different

Table 6. Sequences of BEEC patients and controls sub-cloned into the pGL3 luciferase vector.

| Promoter Position | Genotype of each luciferase reporter clone | | | | |
	Control	12 bp (Cauc)	1 bp (Non-Cauc)	4 bp (Non-Cauc)	4 bp+1 bp (Non-Cauc)
−2657	A	A	A	A	T
−2651	C	C	C	C	C
−2293 to −2282	TCCAGAATCTT	-	TCCAGAATCTT	TCCAGAATCTT	TCCAGAATCTT
−1944	C	C	T	C	T
−1287	-	-	T	-	T
−1209	T	T	C	C	C
−1059	C	C	A	A	A
−71	-	-	-	AGAG	AGAG

Cauc: Caucasian patient sequence. Non-Cauc: Non Caucasian patient sequence. Bp: base pair.
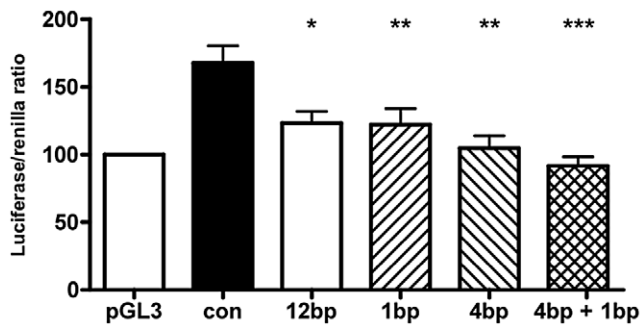doi:10.1371/journal.pgen.1003070.t006

Figure 3. Indel polymorphisms decrease transcriptional efficiency. Luciferase assay of *ΔNp63* promoters with various indel polymorphisms transfected into human embryonic kidney (HEK-293) cells. Normal Control sequence (con), 12 base pair deletion (12 bp; rs6148242), 1 base pair insertion (1 bp, rs5855273), 4 base pair insertion (4 bp, ss#541028600). Data shown is the average of 8 experiments, each examined in triplicate. Data +/− SEM, * : p<0.05, ** p<0.01, *** : p<0.001, compared to normal control, One-way ANOVA, Bonferroni's Multiple Comparison test.
doi:10.1371/journal.pgen.1003070.g003

Table 7. Statistical significance of reduction of transcriptional efficiencies of indel polymorphic *ΔNp63* promoters compared to normal controls.

	Human embryonic kidney cell line: HEK-293		
	P	Mean diff.	95% CI
control vs. 12 bp del.	P<0.05	44.76	6.023 to 83.50
control vs. 1 bp ins.	P<0.01	45.76	7.019 to 84.49
control vs. 4 bp ins.	P<0.001	63.21	24.47 to 101.9
control vs. 4 bp+1 bp ins.	P<0.001	76.18	37.44 to 114.9

P: p-value. CI: 95% confidence intervals. Bp: base pair. Ins: insertion, Del: deletion. Data compared to normal control, One-way ANOVA, Bonferroni's Multiple Comparison test.
doi:10.1371/journal.pgen.1003070.t007

combinations of primers. The PCR products were purified and then used as templates for direct DNA sequencing by the automated ABI Prism 3100 Genetic Analyzer and the Big Dye Terminator kit (Applied Biosystems, Mulgrave, VIC, Australia). Sequences were compared with reference sequences of *p63* GenBank (Locus: NC_000003) using the BioEdit software (Ibis Biosciences). Data were analyzed with contingency tables, χ-square, odds ratio and 95% confidence intervals (GraphPad Prism, GraphPad Software, Inc.).

Quantitative Real-Time PCR

As previously described [9], RNA was extracted from foreskin and bladder urothelial tissue samples from BEEC patients. Circumcision foreskin and normal bladder urothelium from cystectomy (cancer) specimens were used as controls. Tissue samples were snap frozen in liquid nitrogen, ground to powder, mixed with 10 µl of β-mercaptoethanol (Sigma-Aldrich Pty. Ltd, Castle Hill, NSW, Australia) in buffer and centrifuged. Samples were purified with QIAshredder (Qiagen Pty. Ltd., Doncaster, VIC, Australia). RNA was then extracted with RNeasy (Qiagen Pty. Ltd., Doncaster, VIC, Australia) or Trizol (Invitrogen, Mulgrave, VIC, Australia). cDNA synthesis was accomplished using Oligo dT (Invitrogen, Mulgrave, VIC, Australia), dNTPs (Invitrogen, Mulgrave, VIC, Australia), Superscript II First-Strand Synthesis Kit (Invitrogen, Mulgrave, VIC, Australia), and RNAse inhibitor (Invitrogen, Mulgrave, VIC, Australia). Real-time qPCR was carried out using SYBR Green (Applied Biosystems,

Mulgrave, VIC, Australia) on a MJ Research Bio-Rad Chromo 4 cycler (Gladesville, NSW, Australia), according to published *TAp63* and *ΔNp63* primer sequences and published PCR conditions [30]. Relative expression will be analyzed by the Pfaffl methodology with β-actin or GAPDH as the endogenous control [31].

Luciferase Assay

ΔNp63 promoter sequences with indel polymorphisms were cloned into pGL3 luciferase reporter vector (KpnI/XhoI sites Promega, Madison, WI, USA). Human embryonic kidney (HEK-293) (Sigma-Aldrich Pty. Ltd, Castle Hill, NSW, Australia) was used for luciferase assays. Cells were cultured with DMEM (Invitrogen, Mulgrave, VIC, Australia), 10% fetal bovine serum at 37°C with 5% CO_2. Cells were plated on a 96-well plate (BD Biosciences, North Ryde, NSW, Australia). The following day cells were co-transfected with 200 ng of pGL3-promoter constructs containing *ΔNp63* promoters and TK *Renilla* plasmid DNA with 5 µl Lipofectamine Transfection Reagent (Invitrogen, Mulgrave, VIC, Australia). Cells were harvested 48 hours later and cell lysates assayed for luciferase activity with the Dual Luciferase Reporter Assay system (Promega, Madison, WI, USA). Luminescent signals from the firefly and *Renilla* luciferase reactions were measured in a FLUOstar Optima luminometer (BMG Labtech Pty. Ltd., Mornington, VIC, Australia). Luciferase assay reagent was added, firefly luciferase luminescence measured, followed by Stop & Glo Reagent (Promega, Madison, WI, USA). The signal intensity from *Renilla* luciferase was used to normalize the signal from the firefly luciferase. One-way ANOVA with

Table 8. The associated anomalies of BEEC patients compared with the *p63*[−/−] mouse phenotype.

Associated anomalies of BEEC patients	*p63*[−/−] knock-out mice phenotype
Exomphalos (88–100%)	Exomophalos [9]
Ano-rectal malformations	Ano-rectal malformations [32]
Spina bifida, sacral hypoplasia, myelomeningocele	Stunted tail [12]
Cleft palate, median cleft face syndrome [27]	Cleft palate [33]
Epidermolysis bullosa	Non-stratified skin epidermis [33]
Absence of feet, tibial deformity [27]	Absent or stunted limb buds [12]

doi:10.1371/journal.pgen.1003070.t008

post-hoc Bonferroni's multiple comparisons test was applied to compare the relative expressions (GraphPad Prism, GraphPad Software, Inc.).

Accession Numbers

The following submitter National Center for Biotechnology Information (NCBI) SNP (ss) accession numbers were assigned to the SNPs and indel observed in this study: promoter position −2657, ss#541026548; −2651, ss#541027004; −2431, ss#541027120; −2293-2282, ss#5411027737; −2260, ss#541027867; −1944, ss#541027977; −1287, ss#541028071; −1209, ss#541028198; −1059, ss#541028317; −71, ss#541028452 (2 bp); and −71, ss#5410228600 (4 bp).

Author Contributions

Conceived and designed the experiments: WC SW. Performed the experiments: SW ND KJ PK-SN KWZ FN-YF SK-WT WY-WF IM RQ. Analyzed the data: SW KWZ WC PK-SN FN-YF SK-WT WY-WF IM RQ. Contributed reagents/materials/analysis tools: JH ME DB KA AT TB EJ LL WC. Wrote the paper: SW WC.

References

1. Martinez-Frias ML, Bermejo E, Rodriguez-Pinilla E, Frias JL (2001) Exstrophy of the cloaca and exstrophy of the bladder: two different expressions of a primary developmental field defect. Am J Med Genet 99: 261–269.
2. Gearhart JP, Jeffs RD (1989) State-of-the-art reconstructive surgery for bladder exstrophy at the Johns Hopkins Hospital. Am J Dis Child 143: 1475–1478.
3. Boyadjiev SA, Dodson JL, Radford CL, Ashrafi GH, Beaty TH, et al. (2004) Clinical and molecular characterization of the bladder exstrophy-epispadias complex: analysis of 232 families. BJU Int 94: 1337–1343.
4. Nelson CP, Dunn RL, Wei JT (2005) Contemporary epidemiology of bladder exstrophy in the United States. J Urol 173: 1728–1731.
5. Froster UG, Heinritz W, Bennek J, Horn LC, Faber R (2004) Another case of autosomal dominant exstrophy of the bladder. Prenat Diagn 24: 375–377.
6. Ludwig M, Ruschendorf F, Saar K, Hubner N, Siekmann L, et al. (2009) Genome-wide linkage scan for bladder exstrophy-epispadias complex. Birth Defects Res A Clin Mol Teratol 85: 174–178.
7. Shapiro E, Lepor H, Jeffs RD (1984) The inheritance of the exstrophy-epispadias complex. J Urol 132: 308–310.
8. Reutter H, Qi L, Gearhart JP, Boemers T, Ebert AK, et al. (2007) Concordance analyses of twins with bladder exstrophy-epispadias complex suggest genetic etiology. Am J Med Genet A 143A: 2751–2756.
9. Cheng W, Jacobs WB, Zhang JJ, Moro A, Park JH, et al. (2006) DeltaNp63 plays an anti-apoptotic role in ventral bladder development. Development 133: 4783–4792.
10. Koster MI, Kim S, Mills AA, DeMayo FJ, Roop DR (2004) p63 is the molecular switch for initiation of an epithelial stratification program. Genes Dev 18: 126–131.
11. Yang A, Kaghad M, Wang Y, Gillett E, Fleming MD, et al. (1998) p63, a p53 homolog at 3q27-29, encodes multiple products with transactivating, death-inducing, and dominant-negative activities. Mol Cell 2: 305–316.
12. Mills AA, Zheng B, Wang XJ, Vogel H, Roop DR, et al. (1999) p63 is a p53 homologue required for limb and epidermal morphogenesis. Nature 398: 708–713.
13. Baskin LS, Hayward SW, Young P, Cunha GR (1996) Role of mesenchymal-epithelial interactions in normal bladder development. J Urol 156: 1820–1827.
14. Ching BJ, Wittler L, Proske J, Yagnik G, Qi L, et al. (2010) p63 (TP73L) a key player in embryonic urogenital development with significant dysregulation in human bladder exstrophy tissue. Int J Mol Med 26: 861–867.
15. Rodriguez S, Gaunt TR, Day IN (2009) Hardy-Weinberg equilibrium testing of biological ascertainment for Mendelian randomization studies. Am J Epidemiol 169: 505–514.
16. Gaunt TR, Rodriguez S, Day IN (2007) Cubic exact solutions for the estimation of pairwise haplotype frequencies: implications for linkage disequilibrium analyses and a web tool 'CubeX'. BMC Bioinformatics 8: 428.
17. Ni C, Ye Y, Wang M, Qian H, Song Z, et al. (2009) A six-nucleotide insertion-deletion polymorphism in the CASP8 promoter is associated with risk of coal workers' pneumoconiosis. J Toxicol Environ Health A 72: 712–716.
18. Miceli-Richard C, Gestermann N, Ittah M, Comets E, Loiseau P, et al. (2009) The CGGGG insertion/deletion polymorphism of the IRF5 promoter is a strong risk factor for primary Sjogren's syndrome. Arthritis Rheum 60: 1991–1997.
19. Burchard EG, Silverman EK, Rosenwasser LJ, Borish L, Yandava C, et al. (1999) Association between a sequence variant in the IL-4 gene promoter and FEV(1) in asthma. Am J Respir Crit Care Med 160: 919–922.
20. Kulozik AE, Bellan-Koch A, Bail S, Kohne E, Kleihauer E (1991) Thalassemia intermedia: moderate reduction of beta globin gene transcriptional activity by a novel mutation of the proximal CACCC promoter element. Blood 77: 2054–2058.
21. Sun T, Gao Y, Tan W, Ma S, Shi Y, et al. (2007) A six-nucleotide insertion-deletion polymorphism in the CASP8 promoter is associated with susceptibility to multiple cancers. Nat Genet 39: 605–613.
22. Frank B, Rigas SH, Bermejo JL, Wiestler M, Wagner K, et al. (2008) The CASP8 −652 6N del promoter polymorphism and breast cancer risk: a multicenter study. Breast Cancer Res Treat 111: 139–144.
23. Haiman CA, Garcia RR, Kolonel LN, Henderson BE, Wu AH, et al. (2008) A promoter polymorphism in the CASP8 gene is not associated with cancer risk. Nat Genet 40: 259–260; author reply 260–251.
24. Pittman AM, Broderick P, Sullivan K, Fielding S, Webb E, et al. (2008) CASP8 variants D302H and −652 6N ins/del do not influence the risk of colorectal cancer in the United Kingdom population. Br J Cancer 98: 1434–1436.
25. Ihrie RA, Marques MR, Nguyen BT, Horner JS, Papazoglu C, et al. (2005) Perp is a p63-regulated gene essential for epithelial integrity. Cell 120: 843–856.
26. Qi L, Chen K, Hur DJ, Yagnik G, Lakshmanan Y, et al. (2011) Genome-wide expression profiling of urinary bladder implicates desmosomal and cytoskeletal dysregulation in the bladder exstrophy-epispadias complex. Int J Mol Med 27: 755–765.
27. Celli J, Duijf P, Hamel BC, Bamshad M, Kramer B, et al. (1999) Heterozygous germline mutations in the p53 homolog p63 are the cause of EEC syndrome. Cell 99: 143–153.
28. Maas SM, de Jong TP, Buss P, Hennekam RC (1996) EEC syndrome and genitourinary anomalies: an update. Am J Med Genet 63: 472–478.
29. Chuangsuwanich T, Sunsaneevithayakul P, Muangsomboon K, Limwongse C (2005) Ectrodactyly-ectodermal dysplasia-clefting (EEC) syndrome presenting with a large nephrogenic cyst, severe oligohydramnios and hydrops fetalis: a case report and review of the literature. Prenat Diagn 25: 210–215.
30. Chilosi M, Zamo A, Brighenti A, Malpeli G, Montagna L, et al. (2003) Constitutive expression of DeltaN-p63alpha isoform in human thymus and thymic epithelial tumours. Virchows Arch 443: 175–183.
31. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29: e45.
32. Ince TA, Cviko AP, Quade BJ, Yang A, McKeon FD, et al. (2002) p63 Coordinates anogenital modeling and epithelial cell differentiation in the developing female urogenital tract. Am J Pathol 161: 1111–1117.
33. Yang A, Schweitzer R, Sun D, Kaghad M, Walker N, et al. (1999) p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. Nature 398: 714–718.

No TAP63 promoter mutation is detected in bladder exstrophy–epispadias complex patients

Tom Darling[a], Istiak Mahfuz[a], Stefan J. White[a], Wei Cheng[a,b,*]

[a]*Monash Institute of Medical Research, Monash University, Melbourne, Australia*
[b]*Department of Paediatrics and Department of Surgery, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia*

Abstract

Background/Purpose: Bladder exstrophy–epispadias complex (BEEC) is thought to have a genetic component in its pathogenesis. Previously we found that p63$^{-/-}$ mice show increased ventral apoptosis and develop a BEEC phenotype. Down-regulation of the anti-apoptotic ΔNP63 and an up-regulation of pro-apoptotic TAP63 isoforms have been demonstrated in BEEC patient bladder tissues. We have previously shown that insertion/deletion polymorphisms of the ΔNp63 promoter are associated with an increased risk of BEEC. In this study, we specifically examined the *TAP63* promoter to see if any sequence changes might lead to up-regulation of TAP63 and exaggerated apoptosis in BEEC patients.
Methods: i) Bioinformatic analysis of the TAP63 promoter was performed to identify putative regulatory regions. ii) High-resolution Melt and Sanger sequencing was used to screen targeted regions in 112 BEEC patient DNA samples for potential sequence variants. iii) Sequence variation was analysed for significance against normal population frequency data.
Results: i) We identified multiple epigenetic markers of transcriptional regulation within highly conserved areas of the TAP63 promoter sequence. ii) Of the 112 buccal swab DNA samples, adequate and successful screening ranged between 48 and 67 for each region. iii) No novel sequence variation or mutation was uncovered. iv) Two known SNPs were identified. However, allele frequency analysis was not statistically significant.
Conclusion: Our data do not associate genetic variation within the TAP63 promoter region with an increased risk of BEEC. Our data so far suggests that only ΔNP63 promoter aberration is involved in BEEC pathogenesis.
© 2013 Elsevier Inc. All rights reserved.

Bladder exstrophy–epispadias complex (BEEC) represents a spectrum of congenital urogenital anomalies of increasing severity which are believed to share a common pathogenesis. Traditionally, BEEC is thought to be due to

defects that result from partial or incomplete descent of the urogenital sinus septum and rupture of the cloacal membrane during weeks 4–6 of embryogenesis, whereby the arrested process of epithelial-mesenchymal induction and migration

---

* Corresponding author. Department of Paediatrics and Department of Surgery, Southern Medical School, Block E, Level 5, 245 Clayton Road, Clayton, Victoria 3168, Australia. ▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪

towards the midline results in a BEEC phenotype. Severity is presumed to be related to early or late timing of this occurrence, which is supported by the presence of mixed or intermediate phenotypes [1].

BEEC repair involves a series of major reconstructive surgeries and imposes a significant medical, financial and psychosocial burden on both patients and their families. Affected individuals continue to suffer lifelong problems including incontinence, genital disfigurement, renal impairment, and an increased risk of bladder cancer, depression and low self esteem [2]. Antenatal diagnosis is poor, with one UK study finding only a quarter of babies born with bladder or cloacal exstrophy had received antenatal diagnosis [3].

Several factors suggest a genetic aetiology of BEEC. Males are twice as likely to be affected, and there is a significantly higher incidence in Caucasian vs. Asian ethnicities [1,4]. Whilst it seems most cases are sporadic, BEEC affected individuals carry a 400-fold increased risk of having affected offspring [5]. A twin concordance study on 56 twin pairs found monozygotic twins are 5.6 times more likely than dizygotic twins to be affected if their co-twin is affected [6]. Published familial reports where multiple members are affected include two siblings, two third degree cousins, and two uncle–nephew pairs, suggestive of a complex polygenetic or multifactorial inheritance rather than a classical Mendelian inheritance pattern [7]. Studies into environmental factors in BEEC development have been largely inconclusive, with one study finding maternal smoking and radiation exposure in the first trimester to be weakly associated with severe cloacal exstrophy only [1,8]. Assisted reproductive techniques have been found to be associated with an eight-fold increase in BEEC risk [9]. A small number of BEEC cases have occurred in patients with chromosomal abnormalities located across chromosomes 1, 6, 9, 11, 19 and 22 [1,10–14]; however, these cases are in the minority and the genes affected in these chromosomal anomalies remain unclear.

The *TP63* gene has been one of the few genes studied in BEEC pathogenesis. *TP63* codes for the pro-apoptotic TAP63 and anti-apoptotic ΔNP63 isoforms through alternate promoter usage. These transcriptional regulatory proteins act in opposition to activate or repress the expression of downstream target genes, of which over five thousand target sites have been identified [15].

It is the anti-apoptotic ΔNP63 isoform that is the predominant isoform expressed in normal murine and human ventral bladder urothelium. We have previously demonstrated p63−/− null mice (without the anti-apoptotic ΔNp63) show exaggerated apoptosis in the ventral urothelium. The poorly developed ventral bladder urothelium affects the induction of the adjacent mesenchyme. The mesenchyme fails to develop into detrusor muscle and this leads to the development of a bladder exstrophy phenotype [16]. In human BEEC patients, down-regulation of ΔNP63 and up-regulation of TAP63 isoform expression is observed in bladder tissues [17,18]. Gene sequencing in BEEC patients

to date has not revealed any genetic variation in *TP63* exons [17,18]. However, we have recently demonstrated that three insertion/deletion (In/Del) polymorphisms within the ΔNP63 promoter region are associated with a statistically increased risk of BEEC [18]. These polymorphisms were shown to down-regulate ΔNp63 expression *in vitro*, mimicking ΔNP63 down-regulation in human BEEC bladder tissue.

As the TAP63 isoform antagonizes ΔNP63 and up-regulates apoptosis, it is imperative to investigate if there is any concomitant variation within the TAP63 promoter that may contribute to the pathogenesis of BEEC. We hypothesized that the possible exaggerated ventral bladder urothelium apoptosis may be due to either down-regulation of the ΔNP63 isoform and/or up-regulation of the TAP63 isoform. The aim of this new study is to identify TAP63 promoter sequence variations, which may contribute to up-regulation of TAP63 isoforms in patients with BEEC. This may lead to exaggerated apoptosis and a BEEC phenotype in humans.

1. Materials and methods

1.1. Patient DNA samples

The study cohort consisted of 112 BEEC patients, including the sample previously identified with abnormal TAP63 expression. Buccal swab DNA samples (n = 109) originated from India, Bangladesh, China, Australia, Spain, Canada and USA, and were obtained with informed consent from the parents or guardians and ethics approval from the respective institutions. By rubbing the inside of the mouth on each side with a buccal swab, DNA from the exfoliated cells are collected and placed in a storage medium. The swab DNA is stable in the medium at room temperature for 2 months. These were posted to our institute via normal airmail. Buccal DNA isolation was carried out using the BuccalAMP™ DNA Storage and Extraction kit obtained from Epicentre® as per manufacturer's protocol. The DNA concentration of each sample was determined by spectrophotometry using a Nanodrop® ND-1000 spectrophotometer. A small number DNA samples extracted from bladder tissue obtained during surgery were also used (n = 3).

1.2. TAP63 promoter region analysis and primer design

An area 2.5 kb upstream of the TAP63 transcription start site (TSS) was analysed using the University of California, Santa Cruz (UCSC) genome browser (http://genome.ucsc.edu, accessed September 2012) [19]. The current standard human reference genome sequence (February 2009 GRCh37/hg19) [20,21] was displayed and interpreted using information from the ENCODE [22], TRANSFAC® 7.0 [23] and NCBI dbSNP135 [24] databases. From this information primer pairs were designed to encompass

transcription factor binding sites in the promoter where a mutation would be most likely to alter TAp63 production. PCR amplification was done using a Labnet multi-gene PCR Machine (Model: Multigene Gradient, Catalogue number TC 9600-G-230v).

1.3. High resolution melting (HRM) analysis

High resolution melting (HRM) analysis was used to identify any potential sequence variants within the chosen regions. HRM was performed as has been previously described [25] using a LightScanner HR96 machine with a melting range of 70–96 °C. Curve analysis was performed using the LightScanner HRM analysis software (Version 2.0.0.1331). A fluorescence plot (x = time, y = fluorescence) was used to assess amplification efficiency. Normalized and difference fluorescence curves in the curve region encompassing $+/-$ 1 °C from the melting start/stop site were used to identify any potential sequence variants. For difference curves, a baseline was determined from the most frequent curve, against which normalized fluorescence was plotted. Presence of a single PCR product of expected size was confirmed post-HRM with gel electrophoresis.

1.4. Sequencing

Potential sequence variants were confirmed through Sanger sequencing on an Applied Biosystems 3130xl Genetic Analyzer. Sequence chromatograph trace files were viewed using the 4Peaks sequence viewing software by Mekentosj, freely available at http://www.mekentosj.com/science/4peaks. SNP calling was done visually, and sequences were aligned to the genome using BLAT [26].

1.5. Statistical analysis

Identified SNPs were compared to the published data of normal population variation by review of the NCBI dbSNP

for known allele frequency and ethnic variation. Expected normal population SNP frequency within the sample size was calculated using the Hardy Weinberg equation, and compared to sample SNP frequency for significance using Pearson's chi-squared test with GraphPad Prism version 5.0d for Mac, available at http://www.graphpad.com/scientific-software/prism/.

2. Results

2.1. Determination of transcriptional factor binding sites from bioinformatic analysis

Bioinformatic analysis of the highly conserved TAP63 promoter region identified the presence of multiple epigenetic markers of transcriptional regulation within highly conserved areas of the genome. H3K27Ac histone markers, DNase1 hypersensitivity regions, CpG islands and RNA polymerase binding sites indicative of transcriptional regulatory regions were identified. Putative binding sequences for the transcription factors COE1, *MEF2A*, *NFkB*, *PU1 and* FOXA1 were also found within 2 kb of the TAp63 transcription start site.

Eight TAP63 promoter regions, designated 1–8 were designed to encompass these elements, shown in Fig. 1. Primers for PCR amplification of these eight regions are detailed in Table 1. Five primers could be optimised successfully, whilst primers 2, 3 and 4 displayed poor melting properties and were unable to be analysed.

2.2. High resolution melting results

All 112 patient samples were subjected to the screening process, with success ranging between 48 and 67 samples screened across each of regions 1, 5, 6, 7 and 8 (Table 2). A targeted selection of variants within those samples which produced HRM curves of sufficient quality were run on gel
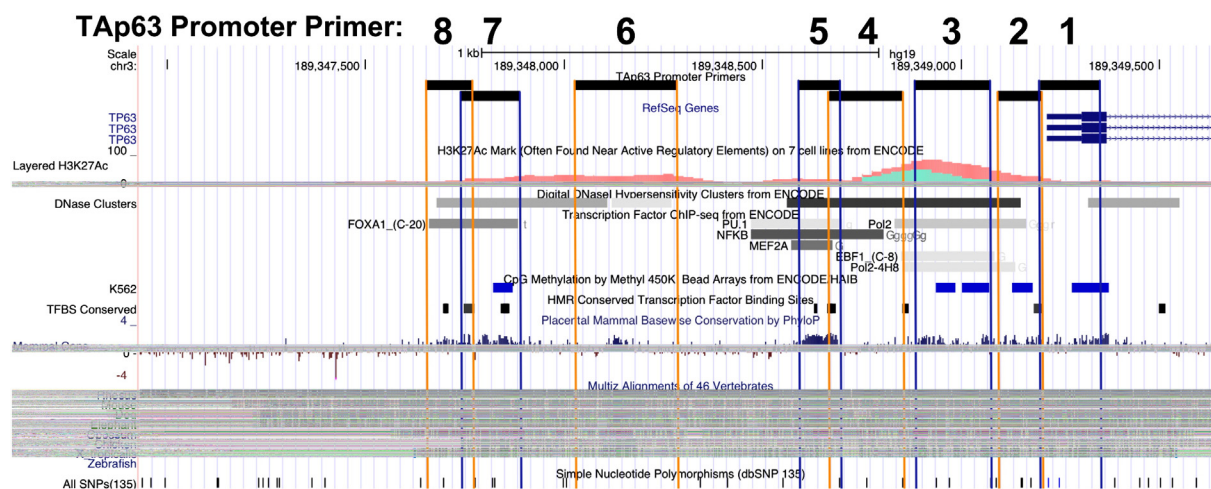

Fig. 1   Bioinformatic analysis and primer locations of 2.5 kb region upstream of TAP63 TSS.

Table 1   The TAP63 promoter primer details.

TAp63 Promoter Primer:	Sequence 5′ → 3′
1 FOR	CCCTATTGCTTTTAGCCTCCCGGC
1 REV	TGCCACCCTACAGTACTGCCCT
2 FOR	GTCAATTGATGAATCTCATTGGCTA
2 REV	TCTGATAGCATTTGACCCTATTGC
3 FOR	AGTCCAGGCTGCTGAAATTAAAC
3 REV	AGGATGCCCAGCTGGTAAGAA
4 FOR	AACAGGCGGTTGGCTGAAAG
4 REV	GTGTTATGTAGCTGGAGCAGGG
5 FOR	GCTGAGAGGAAGAGTGAGTTC
5 REV	GGTTGGCTGAAAGGGAAACTG
6 FOR	GGGTGAAGGGAAGATGTGAACTC
6 REV	TGGTTTCCTTTCGAAGTAGGTCA
7 FOR	TAGCAGGGAAGAAAGCCAAAC
7 REV	AAGCTGTTTGATTTGGCTGGG
8 FOR	TTGTGGAGCAAGCTGTTTGA
8 REV	ATCCTGGCTGAAGCAACAAA

electrophoresis and sequenced. No novel sequence variants were detected. Region 1 showed 48 samples spread across 3 distinct curves (Fig. 2). Two samples from each curve were selected for sequencing which confirmed the variation was the expected result of a known common A/T single nucleotide polymorphism (SNP) (dbSNP rs28673064) within the region. Regions 5 (Fig. 3) and 8 displayed uniform curve shapes across 60 and 56 patients respectively and selected sequencing of 3 patients from each region confirmed no sequence variation as aligned to the current standard UCSC human reference genome sequence (February 2009 GRCh37/hg19). Across region 6, 49 samples were successfully screened and HRM produced 8 distinct curves, with 42 samples belonging to group C (baseline), seen in Fig. 4. To deduce the origin of each curve profile, gel electrophoresis was first run to confirm expected product size (Fig. 5), from which the curves of products F and H were presumed to be a result of primer-dimer formation, suggested by the similarity of their curves combined with appearance of a product < 100 bp on gel. This is supported by their almost identical curve pattern. Sample E showed poor amplification likely distorting its melting curve, which appears to be a mix between the primer and sample curves. Based on clear identifiable bands on the gel, products A, B, C, D and G were selected and successfully sequenced in both forward and reverse directions. Despite their obvious curve variation, these sequences matched the reference sequence and no variation was found. We believe this curve variation is due to other sample impurities or experimental variables. Region 7 showed one clear variant curve among 66 uniform curves, which was sequenced and found to be a known SNP (rs138456383).

2.3. Single nucleotide polymorphism (SNP) frequency analysis

We analysed the risk association of two SNPs found in this set of experiments. The A/T heterozygous SNP

(rs28673064) within the region bound by primer 1 had a minor allele frequency of A = 0.485, as reported in the NCBI dbSNP from data from the 1000 Genomes Project [27]. In our working sample size of 48 BEEC patients, 19 were A/T heterozygous, 13 were homozygous A and 15 were homozygous T. Pearson's chi-squared analysis showed allele frequency trended towards a normal distribution however was not statistically significant due to limited sample size (p = 0.7142) (Fig. 6). A further known heterozygous A/G SNP (rs138456383) with minor allele frequency (MAF) G = 0.004 (1000 Genomes) was identified in one sample within region 7. This SNP had no disease association listed within the database and as an individual result was not considered significant.

3. Discussion

3.1. Mechanisms of TAP63 transcriptional regulation

In this study, we examined the roles transcription factors COE1, MEF2A, NFkB, PU1 and FOXA1, whose binding motifs were identified in the TAP63 promoter, play a role in BEEC pathogenesis. Forkhead box protein A1 (FOXA1) is a transcriptional activator showing spatiotemporal expression within normal bladder development [28]. The factors COE1 and NFkB have been found to play roles in B cell regulation of P53 mediated apoptosis, whilst MEF2A has been shown to be involved in myocyte cell growth control and PU1 may regulate the alternate splicing of target genes. Given their roles in the cell cycle, disruption of any these processes represent a potential upstream mechanism for TAP63 dysregulation and may contribute to BEEC pathogenesis.

3.2. No mutation is detected in the TAp63 promoter

Our data fail to associate genetic variation of the TAP63 promoter region with increased risk of bladder exstrophy–epispadias complex. This indicates that the dysregulated TAP63 expression seen in the few patients reported may not be explained by TAP63 promoter mutation in the regions specified [17,18]. Possible mechanisms that may explain

Table 2   The number of patients screened across each primer set.

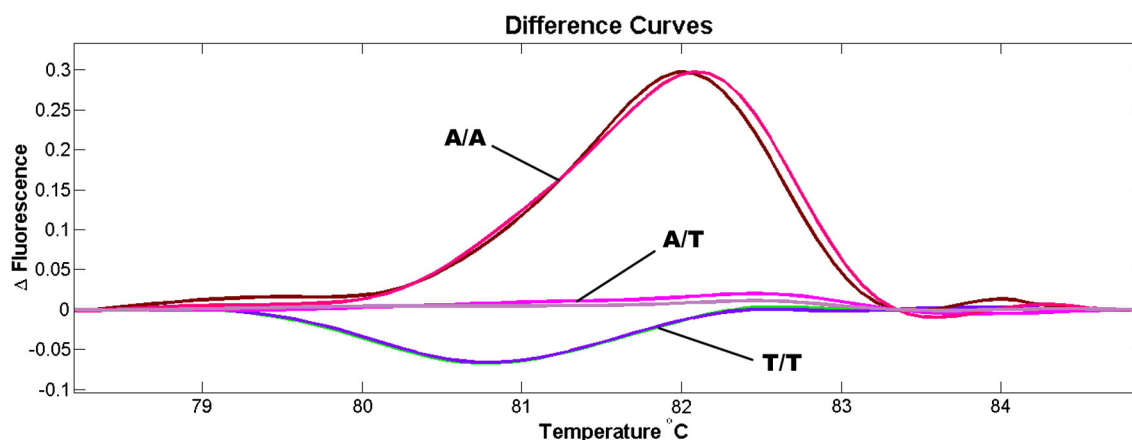| TAp63 promoter primer pair	Number of patients screened
1	48
2	N/A
3	N/A
4	N/A
5	60
6	49
7	67
8	56

Fig. 2   Region 1 SNP variant HRM with different curves chosen for sequencing (6 patient samples).

published results of both up- and down-regulation of TAP63 isoforms in differing BEEC patients include: i) differing patient ages with potentially different bladder TAP63 expression, ii) altered TAP63 expression resulting from exposure of bladder urothelium to the exterior, iii) impact from various surgical interventions and subsequent scarring, iv) differing sites of ventral bladder tissue specimen collection and v) differing sub-epithelial tissue in each sample, diluting the expression of TAp63.

It should be emphasized that fresh RNA tissue samples from neonates with bladder expression is very difficult to obtain, and larger studies will help define whether the extent of TAp63 dysregulation in BEEC patients is more prevalent than currently thought. Our study was similarly limited in scope due to a small sample size.

Our inability to successfully screen all available samples in each of our selectively targeted regions may be the result of sub-optimal buccal swab DNA quantity and quality, hampering the high resolution melt (HRM) screening process which is typically dependent on high quality sample DNA. Difficulty in collection from infants and small children may result in low yield and swab contamination whilst international transport may degrade samples. Spectrophotometry of selected non-working samples showed elevated levels of protein contamination, which may also explain the

HRM curve variation in samples later confirmed to be of identical DNA sequence.

3.3. Alternative explanations

Our result does not support increased TAP63 expression as a possible aetiology of increased apoptosis in bladder exstrophy. It cannot however exclude TAP63 involvement in the disease, including potential disruption of yet unknown alternative TAP63 regulatory sites. Whilst evidence so far supports decreased ΔNP63 expression leading to decreased anti-apoptotic activity in the ventral bladder urothelium as one of the mechanisms of BEEC pathogenesis, the inheritance pattern of BEEC suggests that the condition may be multifactorial and there are likely to be other factors involved. These may include aberrations of P63 transcriptional factors affecting the expression of p63 and/or downstream P63 target genes. There is also the possibility of other p63-independent pathways being involved.

Qi et al. have identified a number of BEEC candidate genes in patient tissue samples relating to desmosomal and cytoskeletal assembly, including the P63 downstream target genes P53 effector related to PMP22 (PERP) and desmoplakin (DSP) [29]. These intracellular junction proteins adhere cells within tissues providing the multicellular
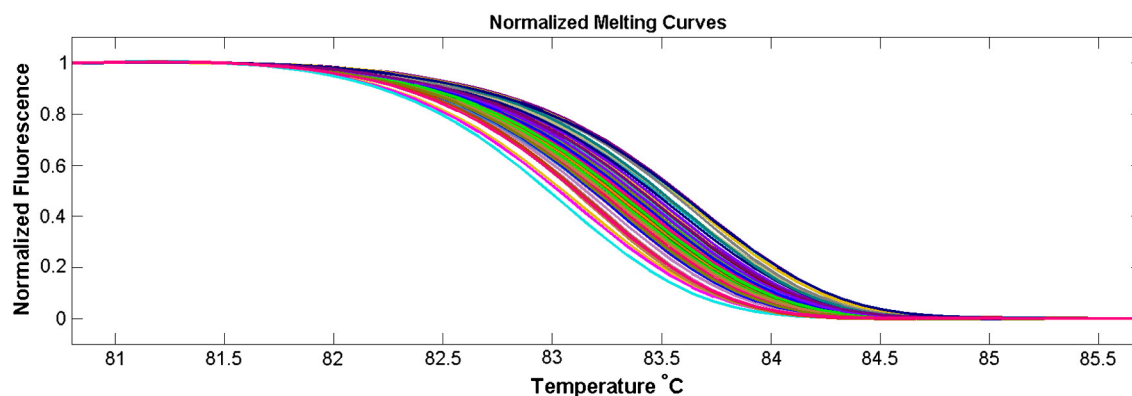

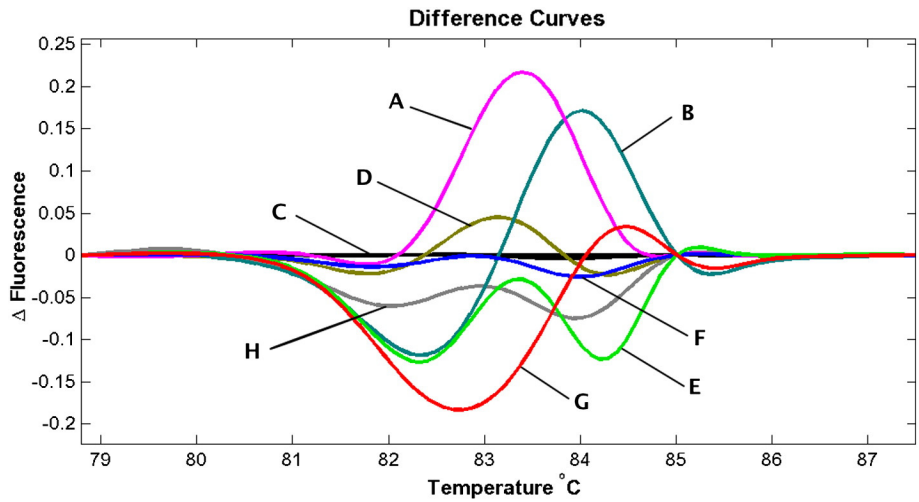**Fig. 3**   Normalised HRM curves of region 5 (60 patient samples).

Fig. 4 Grouped HRM difference curves from region 6.

structure and strength essential for withstanding mechanical stress such as in the bladder and skin [30]. Like P63, PERP and DSP are expressed in stratified epithelia during embryogenesis [31,32]. Perp$^{-/-}$ mice die of dehydration shortly after birth through blister formation in the stratified epithelium [31] similar to p63$^{-/-}$ mice [16] whilst Dsp alteration can lead to epidermolysis bullosa, a blistering skin condition which has previously been reported in BEEC patients [33,34]. As such, PERP and DSP are promising candidates for involvement in BEEC pathogenesis. The role of PERP and DSP in p63 regulated development is to maintain epithelial integrity. Disruption of the P63 > PERP > DSP pathway resulting in a loss of epithelial integrity during bladder development may represent the next promising step in our understanding of BEEC pathogenesis.

3.4. The published evidence of BEEC pathogenesis

Published data has so far described ΔNP63 as the main P63 isoform implicated in BEEC [16–18]. We have previously found that ΔNp63 is preferentially expressed in ventral bladder urothelium during development and in its absence p63$^{-/-}$ mice develop a bladder exstrophy phenotype [16]. Two studies showing reduced ΔNP63 expression in human bladder tissue have established its involvement in human BEEC pathogenesis [17,18]. In search of a genetic aetiology for this altered expression, P63 sequencing in two separate studies of 15 and 22 BEEC patients has not revealed any exon mutations [17,18]. Our sequencing of the ΔNP63 promoter in a study of 163 BEEC DNA patients has however found four insertion/deletion (In/Del) polymorphisms which confer an increased risk of BEEC and decrease the transcriptional efficiency of ΔNP63 in human embryonic kidney (HEK-293) cells [18]. These polymorphisms show a significant bias between Caucasian and non-Caucasian groups which is highly significant as there is an increased risk of BEEC in Caucasian populations [35].

Whether an independent role for the pro-apoptotic TAP63 isoform in BEEC pathogenesis exists is unanswered. We have found increased TAP63 expression in two BEEC patient samples compared to controls, and Ching et al. found an absence of TAP63 expression in three BEEC bladder tissue samples. Admittedly, the sample size is small. Whether these are secondary to or contribute to BEEC pathogenesis is
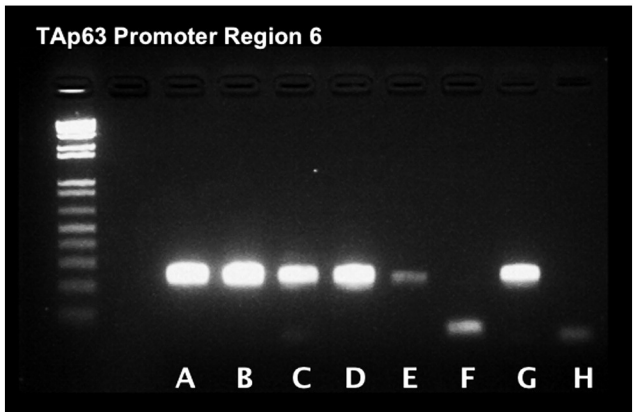

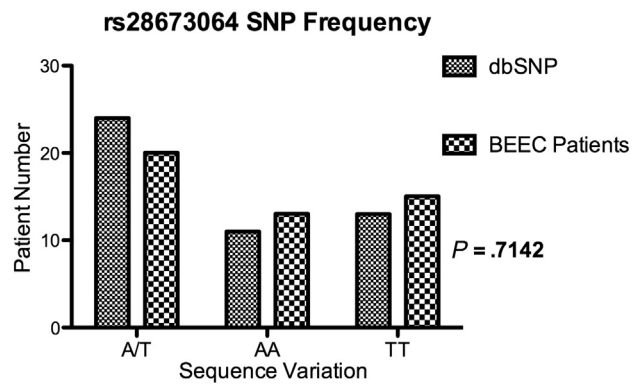**Fig. 5** Gel electrophoresis of grouped samples from region 6 as represented in Fig. 4.

Fig. 6 TAp63 region 1 SNP frequency analysis.

unknown. However, ΔNP63 and TAP63 isoforms have a dominant negative relationship whereby ΔNP63 negates the transactivating potential of TAP63, as it occupies the same binding site [15]. Thus an increase in TAp63-mediated ventral bladder apoptosis through a lack of ΔNp63 target site opposition might explain one mechanism by which reduced ΔNp63 expression leads to a bladder exstrophy phenotype. The results of this study further support the notion that TAP63 isoforms may play a secondary or absent role in this disease, whilst ΔNP63 plays a more prominent role.

In conclusion, we did not identify genetic variation within the TAP63 promoter region that could be associated with an increased risk of bladder exstrophy–epispadias complex. These data and our published findings in human and p63 knock-out mice suggest that ΔNP63 promoter mutation may be the more likely mechanism behind the P63 isoform dysregulation found in BEEC patients.

Acknowledgments

References

[1] Boyadjiev SA, Dodson JL, Radford CL, et al. Clinical and molecular characterization of the bladder exstrophy–epispadias complex: analysis of 232 families. BJU Int 2004;94(9):1337-43.

[2] Diseth THT, Emblem RR, Schultz AA. Mental health, psychosocial functioning, and quality of life in patients with bladder exstrophy and epispadias – an overview. World J Urol 1999;17(4):239-48.

[3] Goyal A, Fishwick J, Hurrell R, et al. Antenatal diagnosis of bladder/cloacal exstrophy: challenges and possible solutions. J Pediatr Urol 2012;8(2):140-4.

[4] Jayachandran DD, Bythell MM, Platt MWM, et al. Register based study of bladder exstrophy–epispadias complex: prevalence, associated anomalies, prenatal diagnosis and survival. J Urol 2011;186(5):6.

[5] Shapiro EE, Lepor HH, Jeffs RDR. The inheritance of the exstrophy–epispadias complex. J Urol 1984;132(2):308-10.

[6] Reutter H, Qi L, Gearhart JP, et al. Concordance analyses of twins with bladder exstrophy–epispadias complex suggest genetic etiology. Am J Med Genet 2007;143(22):2751-6 [Wiley Online Library].

[7] Reutter H, Shapiro E, Gruen JR. Seven new cases of familial isolated bladder exstrophy and epispadias complex (BEEC) and review of the literature. Am J Med Genet 2003;120A(2):215-21.

[8] Reutter H, Boyadjiev SA, Gambhir L, et al. Phenotype severity in the bladder exstrophy–epispadias complex: analysis of genetic and non-genetic contributing factors in 441 families from North America and Europe. J Pediatr 2011;159(5):825-831.e1.

[9] Zwink N, et al. Assisted reproductive techniques and risk of exstrophy–epispadias complex: a German case–control study. J Urol 2013;189(4):1524-9.

[10] Leonard NJN, Tomkins DJD. Diploid/tetraploid/t(1;6) mosaicism in a 17-year-old female with hypomelanosis of Ito, multiple congenital anomalies, and body asymmetry. Am J Med Genet 2002;112(1):86-90.

[11] Draaken M, Mughal SS, Pennimpede T, et al. Isolated bladder exstrophy associated with a de novo 0.9 Mb microduplication on chromosome 19p13.12. Birth Defects Res A Clin Mol Teratol 2013;97(3):133-9.

[12] Thauvin-Robinet C, Faivre L, Cusin VR. Khau Van Kien P, Callier P, Parker KL, et al. Cloacal exstrophy in an infant with 9q34.1-qter deletion resulting from a de novo unbalanced translocation between chromosome 9q and Yq. Am J Med Genet 2004;126A(3):303-7.

[13] Boyadjiev SA, South ST, Radford CL, et al. A reciprocal translocation 46,XY,t(8;9)(p11.2;q13) in a bladder exstrophy patient disrupts CNTNAP3 and presents evidence of a pericentromeric duplication on chromosome 9. Genomics 2005;85(5):8.

[14] Kosaki R, Fukuhara Y, Kosuga M, et al. OEIS complex with del(3)(q12.2q13.2). Am J Med Genet 2005;135(2):224-6.

[15] Yang A, Zhu Z, Kapranov P, et al. Relationships between p63 binding, DNA sequence, transcription activity, and biological function in human cells. Mol Cell 2006;24(4):593-602.

[16] Cheng W, Jacobs WB, Zhang JJR, et al. DeltaNp63 plays an anti-apoptotic role in ventral bladder development. Development 2006;133(23):4783-92.

[17] Ching BJ, Wittler L, Proske J, et al. p63 (TP73L) a key player in embryonic urogenital development with significant dysregulation in human bladder exstrophy tissue. Int J Mol Med 2010;26(6):861-7.

[18] Wilkins SS, Zhang KWK, Mahfuz II, et al. Insertion/deletion polymorphisms in the ΔNp63 promoter Are a risk factor for bladder exstrophy epispadias complex. PLoS Genet 2012;8(12):e1003070.

[19] Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. Genome Res 2002;12(6):996-1006.

[20] Lander ESE, Linton LML. 257. Initial sequencing and analysis of the human genome. Nature 2001;409(6822):860-921.

[21] International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature 2004;431(7011):931-45.

[22] Rosenbloom KR, Dreszer TR, Kent WJ. 27. ENCODE whole-genome data in the UCSC genome browser: update 2012. Nucleic Acids Res 2011;40(Database issue):D912-7.

[23] Matys V, Kel-Margoulis OV, Fricke E, et al. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res 2005;34(1):D108-10.

[24] Sherry ST, Ward MH, Sirotkin K. 7. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 2000;29(1):308-11.

[25] de Boer CM, Eini R, Gillis AM, et al. DICER1 RNase IIIb domain mutations are infrequent in testicular germ cell tumours. BMC Res Notes 2012;5:569.

[26] Kent WJW. BLAT – the BLAST-like alignment tool. Genome Res 2002;12(4):656-64.

[27] 1000 Genomes Project Consortium. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. Nature 2010;467(7319):1061-73.

[28] Oottamasathien S, et al. Directed differentiation of embryonic stem cells into bladder tissue. Dev Biol 2007;304(2):556-66.

[29] Qi L, Chen K, Hur DJ, et al. Genome-wide expression profiling of urinary bladder implicates desmosomal and cytoskeletal dysregulation in the bladder exstrophy–epispadias complex. Int J Mol Med 2011; 27(6):755-65.

[30] Ihrie RA, Attardi LD. A new Perp in the lineup: linking p63 and desmosomal adhesion. Cell Cycle 2005;4(7):873-6.

[31] Ihrie RAR, Marques MRM, Nguyen BTB, et al. Perp is a p63-regulated gene essential for epithelial integrity. Cell 2005;120(6):843-56.

[32] Marques MRM, Ihrie RAR, Horner JSJ, et al. The requirement for perp in postnatal viability and epithelial integrity reflects an intrinsic role in stratified epithelia. J Investig Dermatol 2005;126(1): 69-73.

[33] Moretti GG, Mazzaglia EE, D'Anieri AA, et al. Epidermolysis bullosa junctionalis associated with urinary bladder exstrophy: a case report. Pediatr Dermatol 1995;12(3):239-41.

[34] Jonkman MF, Pasmooij AMG, Pasmans SGMA, et al. Loss of desmoplakin tail causes lethal acantholytic epidermolysis bullosa. Am J Hum Genet 2005;77(4):653-60.

[35] Gambhir L, Höller T, Müller M, et al. Epidemiological survey of 214 families with bladder exstrophy–epispadias complex. J Urol 2008;179(4):1539-43.

BMC
Research Notes

SHORT REPORT   **Open Access**

Identification of *Streptococcus parasanguinis* DNA contamination in human buccal DNA samples

Istiak Mahfuz[1], Wei Cheng[2,3,4] and Stefan J White[1*]

Abstract

Background: The use of buccal swabs in clinical and scientific studies is a very popular method of collecting DNA, due to its non-invasive nature of collection. However, contamination of the DNA sample may interfere with analysis.

Findings: Here we report the finding of *Streptococcus parasanguinis* bacterial DNA contamination in human buccal DNA samples, which led to preferential amplification of bacterial sequence with PCR primers designed against human sequence.

Conclusion: Contamination of buccal-derived DNA with bacterial DNA can be significant, and may influence downstream genetic analysis. One needs to be aware of possible bacterial contamination when interpreting abnormal findings following PCR amplification of buccal swab DNA samples.

Findings

Buccal swabs are a popular, inexpensive and non-invasive method of collecting DNA samples. It is a convenient procedure for collecting DNA from geographically isolated populations for larger cohort studies [1], and has the advantage of avoiding the stressful process of venepuncture. When using buccal swab DNA, sampling or processing considerations may be important for obtaining optimal results [2]. If buccal swabs are not collected and/or handled properly, potential complications may occur during subsequent analysis [3]. Problems that can affect the interpretation include contamination, degradation, and insufficient yield [4-6]. Here we report a significant *Streptococcus parasanguinis* DNA contamination in human buccal-derived DNA. This bacteria is the most abundant microorganism in the mouth, and a primary colonizer of human tooth surfaces [7,8] that plays an important role in dental plaque formation [9,10].
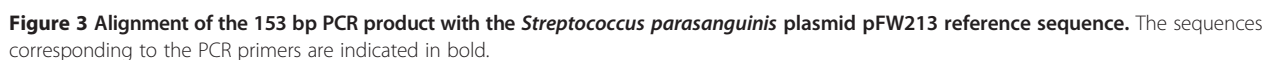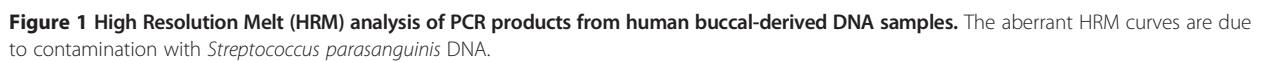
Contamination was detected during mutation screening of the Desmoplakin (*DSP*) gene in buccal DNA of Bladder Exstrophy Epispadias Complex (BEEC) patients. Use of samples for research was approved by the Research Ethics Committee of Royal Children's

Hospital (Approval Number# HREC28140A), Melbourne, Australia. Informed consent was also obtained from the patients or parents/guardians. The intention was to screen *DSP* exon 4 for potential sequence variants using High Resolution Melting (HRM) and Sanger Sequencing. For this, we designed exon-specific forward (5′ CTGTTTTCCTGCAGTGGTT 3′) and reverse (5′ TGGCCTGCACAGGTTTG 3′) primers, predicted to generate a 254 bp product. HRM was performed on 22 samples as previously described [11]. Five samples gave aberrant curves with HRM (Figure 1), and agarose gel analysis showed the presence of two bands in matching samples (Figure 2). Sanger sequence analysis showed that the 153 bp fragment did not align with any human sequence. Alignment with other organisms showed 98% identity with *Streptococcus parasanguinis* plasmid pFW213 [GenBank:EU685104.1] (Figure 3).

To our knowledge, there have been no previous reports of *Streptococcus parasanguinis* contamination affecting genetic analysis of buccal swab DNA. It is not unexpected to have Streptococcal DNA contamination in DNA from a human buccal swab, due to the high abundance in oral microflora [12,13]. In addition, when a buccal sample is collected, researchers generally do not enquire whether there is any oral disease. This might influence the quality of DNA isolated from buccal swabs. Under such

* Correspondence: ▮▮▮▮▮▮▮▮▮▮▮▮▮
[1]Centre for Genetic Diseases, Monash Institute of Medical Research, Monash University, 27-31 Wright Street, Clayton 3168, VIC, Australia
Full list of author information is available at the end of the article

Figure 1 High Resolution Melt (HRM) analysis of PCR products from human buccal-derived DNA samples. The aberrant HRM curves are due to contamination with *Streptococcus parasanguinis* DNA.

Figure 2 Gel electrophoresis of the PCR products shown in Figure 1. The expected product from the human DNA is 254 bp, with the 153 bp product matching *Streptococcus parasanguinis* DNA.

Figure 3 Alignment of the 153 bp PCR product with the *Streptococcus parasanguinis* plasmid pFW213 reference sequence. The sequences corresponding to the PCR primers are indicated in bold.

circumstances, a blood sample may be a preferable method for obtaining DNA. A comparison of blood- and buccal-derived DNA from a Danish nurse cohort has previously been reported. 100% of the blood-derived DNA samples could be genotyped or PCR amplified, whereas only 23% of the DNA samples from mouth swabs could be PCR amplified and none of the swab-derived DNA samples could be genotyped [14].

Our finding demonstrates that bacterial contamination may be a significant contributor to the total amount of DNA isolated from a buccal swab, and may explain why an apparently sufficient quantity of high quality DNA does not perform as expected in human-specific PCR amplifications.

Authors' contributions
IM performed the molecular analysis and drafted the manuscript. WC helped in coordination and drafting of the manuscript. SW conceived of the study, participated in its design and coordination and drafted the manuscript. All authors read and approved the final manuscript.

Authors' information
Wei Cheng and Stefan J White: Joint senior authors.

Author details
[1]Centre for Genetic Diseases, Monash Institute of Medical Research, Monash University, 27-31 Wright Street, Clayton 3168, VIC, Australia. [2]Department of Paediatrics, Southern Medical School, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia. [3]Department of Surgery, Southern Medical School, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia. [4]Department of Surgery, Beijing United Family Hospital, Beijing, China.

References
1. Freeman B, Powell J, Ball D, Hill L, Craig I, Plomin R: **DNA by mail: an inexpensive and noninvasive method for collecting DNA samples from widely dispersed populations.** *Behav Genet* 1997, **27**(3):251–257.
2. Walker AH, Najarian D, White DL, Jaffe JF, Kanetsky PA, Rebbeck TR: **Collection of genomic DNA by buccal swabs for polymerase chain reaction-based biomarker assays.** *Environ Health Perspect* 1999, **107**(7):517–520.
3. **Collecting a Buccal Swab – An Art or a Cinch?** http://www.hitdna.com/images/PDFs/publications/Taking a Buccal Swab.pdf.
4. Beckett SM, Laughton SJ, Pozza LD, McCowage GB, Marshall G, Cohn RJ, Milne E, Ashton LJ: **Buccal swabs and treated cards: methodological considerations for molecular epidemiologic studies examining pediatric populations.** *Am J Epidemiol* 2008, **167**(10):1260–1267.
5. Walsh DJ, Corey AC, Cotton RW, Forman L, Herrin GL Jr, Word CJ, Garner DD: **Isolation of deoxyribonucleic acid (DNA) from saliva and forensic science samples containing saliva.** *J Forensic Sci* 1992, **37**(2):387–395.
6. Zayats T, Young TL, Mackey DA, Malecaze F, Calvas P, Guggenheim JA: **Quality of DNA extracted from mouthwashes.** *PLoS One* 2009, **4**(7):e6165.
7. Tinanoff N, Gross A, Brady JM: **Development of plaque on enamel. Parallel investigations.** *J Periodontal Res* 1976, **11**(4):197–209.
8. Socransky SS, Manganiello AD, Propas D, Oram V, van Houte J: **Bacteriological studies of developing supragingival dental plaque.** *J Periodontal Res* 1977, **12**(2):90–106.
9. Carlsson J, Grahnen H, Jonsson G, Wikner S: **Establishment of Streptococcus sanguis in the mouths of infants.** *Arch Oral Biol* 1970, **15**(12):1143–1148.
10. Jenkinson HF, Lamont RJ: **Streptococcal adhesion and colonization.** *Crit Rev Oral Biol Med* 1997, **8**(2):175–200.
11. de Boer CM, Eini R, Gillis AM, Stoop H, Looijenga LH, White SJ: **DICER1 RNase IIIb domain mutations are infrequent in testicular germ cell tumours.** *BMC Res Notes* 2012, **5**:569.
12. The Human Microbiome Project Consortium: **Structure, function and diversity of the healthy human microbiome.** *Nature* 2012, **486**(7402):207–214.
13. Rudney JD, Chen R, Zhang G: **Streptococci dominate the diverse flora within buccal cells.** *J Dent Res* 2005, **84**(12):1165–1171.
14. Hansen TV, Simonsen MK, Nielsen FC, Hundrup YA: **Collection of blood, saliva, and buccal cell samples in a pilot study on the Danish nurse cohort: comparison of the response rate and quality of genomic DNA.** *Cancer Epidemiol Biomarkers Prev* 2007, **16**(10):2072–2076.

```
                  Minerva Biotecnologica
        ==========================================================


Title: Cloning and site-directed mutagenesis in NALP3 gene
Paper code: Minerva Biotec-1723
Submission Date: 2013-04-10 18:44:43
Article Type: Original Article


Files:
1): Manuscript
     Version: 1
     Description: Manuscript
     File format: application/msword

2): Tables 1
     Version: 2
     Description: Table
     File format: application/msword

3): Figures 2
     Version: 1
     Description: Figure
     File format: application/msword
```

Cloning and site-directed mutagenesis in *NALP3* gene

AHMAD HUMAYAN KABIR[1*], IMRAN AHMED[1], M BULBUL AHMED[2], ISTIAK MAHFUZ[3]

[1]Department of Botany, University of Rajshahi, Rajshahi 6205, Bangladesh

[2]Department of Botany, Government Azizul Haque College, Bogra, Bangladesh

[3]Monash Institute of Medical Research, Faculty of Medicine, Nursing, and Health Sciences, Monash University, Melbourne, Australia

[*]Corresponding author: ███████████████

Abstract

Cloning and site-directed mutagenesis in *NALP3* gene has been performed mediated by ligation independent cloning vector. Accurate overhang was created in plasmid carrying the *NALP3* gene by using *NALP3* T436P plasmid DNA in proof reading PCR and the mutated *NALP3* gene was successfully inserted to ligation independent vector, pTriEx-6 3C/LIC. Specific band was found in agarose gel electrophoresis followed by colony PCR. In addition to it, mutation was created at the 779 nucleotide position of the *NALP3* gene (c.779G>T) which was initiated by mutagenic primers and QuikChange site-directed mutagenesis kit. The mutated sequence was blasted to find out any homology within genome database which showed 92% sequence similarity with several *Homo sapiens* NLR family. Altogether, these results suggest that our methodological strategies could be used as a potential protocol for studying genetic disorders related to *NALP3*.

Key words: Ligation independent cloning, Mutagenesis, NALP3 gene, Colony PCR.

1

Introduction

The *NALP3* gene encodes for cryopyrin protein that is a member of nucleotide binding oligomerization domain (NOD) like receptors[1)]. The role of cryopyrin is to form inflammasomes that eventually participates in inflammation process. The *NLRP3* gene is located in the long (q) arm at position 44 of chromosome 1[10]. The *NALP3* protein is involved with apoptosis-associated protein PYCARD/ASC that is a member of the *NALP3* inflammasome complex. Mutation in *NALP3* gene has been reported to be associated with many disorders in immunology[11]. Among them, MWS (Muckle-Wells syndrome), FCAS (familial cold autoinflammatory Syndrome) and NOMID (neonatal-onset multisystem inflammatory disease) are most familiar. Cryopyrinopathies, mostly known as cryopyrin is also proven to be caused by mutation in the *NALP3*gene sequence[11]. Overlapping phenotypes is also found among MWS/NOMID or MWS/FCAS patients[7]. Several types of symptoms of inflammation are associated with the disorders caused for mutation in *NALP3* gene. In case of NOMID/CINCA, patients represents nonpruritic urticaria-like fever and rash in the early infancy period and later on, increase cerebral atrophy, intracranial pressure, sensorineural hearing loss and mental retardation are observed[7]. It was also evident that 25% of the MWS patients have been reported to develop malaise, progressive hearing loss, arthralgia[7]. FCAS disorder is involved with fever, headaches, arthralgia, and rashes in early few months[7].

Ligation Independent Cloning (LIC) is an efficient technique for cloning PCR products without restriction enzyme digestion or ligation reactions[2,4,5]. Vectors used in LIC cloning are created by treating a linear backbone with T4 DNA polymerase in the presence of only one dNTP. The T4 DNA polymerase removes nucleotides until it comes to a residue corresponding to the single dNTP present in the reaction mix[8]. In this reaction, the 5'-ends of the primers is used to generate the PCR fragment for cloning that contain an additional 12-14 nucleotide overhangs in the LIC vector. The 3'-terminal sequence is then removed by the exonuclease activity (3'-5') of T4 DNA polymerase in the presence of dGTP[12]. The LIC cloning is very efficient since it is formed by annealing and highly versatile. Moreover, the target protein carries only one nonnative amino acid at the N-terminus end after cleavage. As a result, methionine allows the removal of vector encoded sequences[8].

2

Mutagenesis study has an immense importance in human genetics research. QuikChange Site directed mutagenesis is a precious technique that permits site specific mutation in double-stranded plasmid to make point mutations which replace amino acids and delete or introduce single or multiple adjacent amino acids[6,3]. Site directed mutagenesis and mapping of mutated gene is of great importance to understand molecular mechanisms of diseases. Very few studies have been conducted on NALP3 gene so far and that is why this study was aimed.

The aim of this study was to create overhang in *NALP3* gene and clone the mutated *NALP3* gene by using ligation independent cloning vector. Further aim was to create mutation via site directed mutagenesis and verification of that mutation by sequencing with a view to show impact of the mutation and their molecular interactions in the inflammasome.

Materials and Methods

Ligation independent cloning

The mutated *NALP3* gene having 3085bp was used in the study. The plasmid carrying the *NALP3* gene was then taken to proof reading PCR to create overhangs in *NALP3* gene. In this proof reading PCR, old fast back *NALP3* T436P plasmid DNA and two primers (sense 5'GAC GACGACAAGAT-3' and antisense 5'GAGGAGAAGCCT-3) were used. The whole proof reading PCR method was conducted according to the Finnzymes Phusion™ High-Fidelity kit manual (New England Biolabs). The PCR product was then run in agarose gel electrophoresis to check the accurate size of the *NALP3* gene. The plasmid was then purified according to the Qiagen plasmid purification kit.

The checked PCR product was cloned into pTriEx-6 3C/LIC vector according to the Ek/LIC Cloning Kit manual (Novagen). After cloning, the recombinant plasmid was transformed into NovaBlue GigaSingles™ competent cells followed by Ek/LIC Cloning Kit manual (Novagen). The transformed cells were then cultured on ampicillin (100μg/ml) containing LB (Luria-Bertani) medium in order to to select the successful transformants. Three colonies for the transformants were taken by sterile loop for colony PCR. The primers used for colony PCR were pTriEx_FP 5'-CTGGTTATTGTGCTGTCTCATCA-3' and 157_RP 5'-CCACATGGTCTGCCTTCTCT-3'. Specific protocol was followed for the colony PCR (New England Biolabs). The aim of colony PCR was to screen for plasmid inserts directly from bacterial colonies. The PCR product was then loaded on agarose gel for checking.

3

Site directed mutagenesis

Methods for creation of mutation and sequencing of the mutated gene were carried out according to the standard protocols supplied by kit provider.

Mutagenic primer design

The mutagenic oligonucleotide primers were designed by QuikChange II XL tool for creation of mutation which had 35 nucleotides and 78°C annealing temperatures (Stratagene). Designed primers are shown in the table 2.

Mutagenesis

The QuikChange II site-directed mutagenesis method was performed to create mutation at specific position (779). After PCR, the amplified products were digested with *DpnI*, transformed into competent cell and incubated for the growth of bacteria. Ampicillin was used as antibiotic in modified LB broth and LB plates in transformation reaction. Whole methods have been followed by QuikChange II site-directed mutagenesis instruction manual.

Plasmid preparation and concentration measurement

After transformation, plasmid DNA was purified with the QIAprep Spin Miniprep Kit and concentration of the purified DNA was checked by NanoDrop ND-1000.

Sequencing

Purified DNA was then used for sequencing. Sense primer (left primer) 503 was used for sequencing PCR. After PCR, extension products were purified with DyeEx 2.0 Spin Kit followed by DyeEx 2.0 Spin protocol (Qiagen). Purified extension products were then re-suspended by adding TSR (template suppression reagent) followed by drying at 90°C up to 30 mins and placed for sequencing [7].

```
start len    tm      gc%   any   3' seq
LEFT_PRIMER 503   20   59.90   55.00  4.00  3.00 ACACACGACTGCGTCTCATC
```

Bioinformatics analysis

After sequencing, obtained sequences were then blasted to check out its resemblance with any wild type human sequences in NCBI database. Obtained sequences were aligned with wild type sequences by ClustalW2 software.

4

Results and Discussion

Cloning

Result from colony PCR revealed that *NALP3* gene was successfully inserted and cloned into LIC vector (Figure 1). LIC cloning vector was chosen for this experiment due to its efficiency to clone the PCR product without any enzyme digestion[2,4,5]. LIC overcomes some traditional cloning limitations, since clone can be obtained without digesting with restriction endonucleases and ligation. After transformation, the host cell enzyme ligates at the vector-insert site[2]. Successful cloning of target gene by using LIC vector was also reported by several workers[3]. Moreover, LIC cloning is able to check individual clone more quickly that express a target protein *in vitro* by amplifying by colony PCR using appropriate primers. This finding also pinpoints the potential use of ligation independent cloning technique for subsequent protein expression.

Site-directed mutagenesis

The mutation was occurred in the 779 nucleotide position due to the base substitution of G to T (c.779G>T) which is located at exon 3 position in the gene. In site directed mutagenesis, vector having a supercoiled double-stranded DNA (dsDNA) with a target insert, two synthetic oligonucleotide primers (QuikChange II Site Directed mutagenesis kit) were used. This experiment is very much comparable with another report where they used endonuclease *DpnI* for methylated and hemimethylated DNA sequences[7].

Sequencing of DNA which was successfully performed using BigDye Terminator v1.1 Cycle Sequencing Kit (Figure 2). Sequenced data was then blasted (Table 3) and maximum 92% identity was observed with several *Homo sapiens* NLR family, *NALP3* transcripts (Table 3).

Site directed mutagenesis could be utilized for determination of a specific gene activity. Moreover, these practical methodologies used in this study would be helpful for future researchers to carry out mutagenesis experiment. Protocol for creating mutation in *NALP3* could be used for further consequences of genetic disorders related to *NALP3*.

References

1. Aksentijevich I, Putnam CD, Remmers EF, Mueller JL, Le J, Kolodner RD, Moak Z, Chuang M, Austin F, Goldbach-Mansky R, Hoffman HM, Daniel LK, Kastner DL *et*

al. The clinical continuum of cryopyrinopathies: Novel CIAS1 mutations in North American patients and a new cryopyrin model. Arthritis & Rheumatism 2007;54(4): 1273-1285.

2. Aslanidis C, De Jong PJ, Schmitz G *et al*. Minimal length requirement of the single-stranded tails for ligation-independent cloning (LIC) of PCR products. PCR Methods Applications 1994;4:172-177.

3. Bardóczy V, Géczi V, Sawasaki T, Endo Y, Mészáros T et al. A set of ligation-independent *in vitro* translation vectors for eukaryotic protein production. BMC Biotechnology 2008;8:32.

4. Bonsor D, Butz SB, Solomons J, Grant S, Fairlamb IJS, Fogg MJ, Grogan G *et al*. Ligation independent cloning (LIC) as a rapid route to families of recombinant biocatalysts from sequenced prokaryotic genomes. Organic & Biomolecular Chemistry 2006;4:1252-1260.

5. Haun RS, Servanti IM, Moss J *et al*. Biotechniques 1992; 13: 515-518.

6. Lucy SMG, Lynda D, Rosemarie R, Frank RC, Mark ID *et al*. A New Vector for High-Throughput, Ligation-Independent Cloning Encoding a Tobacco Etch Virus Protease Cleavage Site. Protein Expression and Purification 2002;25(1):8-15.

7. Li J, Li C, Xiao W, Yuan D, Wan G, Ma L *et al*. Site-directed mutagenesis by combination of homologous recombination and DpnI digestion of the plasmid template in *Escherichia coli*. Analytical Biochemistry 2007;373:389–391.

8. Robert EN, Keith WY, Kristin MK *et al*. Ligation Independent Cloning: Efficient Directional Cloning of PCR Products. Methods in Molecular Biology 2002;192.

9. Sanjuan R, Daros J. One-step site-directed mutagenesis of viroid dimeric cDNA. Journal of Virological Methods 2007;145:1–75.

6

10. Agostini L, Martinon F, Burns K, McDermott MF, Hawkins PN, Tschopp J *et al*. NALP3 forms an IL-1beta-processing inflammasome with increased activity in Muckle-Wells autoinflammatory disorder. Immunity 2004;20(3):319-325).

11. Dinarello CA. Unraveling the NALP-3/IL-1β inflammasome: A big lesion from a small mutation. Immunity 2004;20:243–246.

12. Hoffman HM, Mueller JL, Broide DH, Wanderer AA, Kolodner RD *et al*. Mutation of a new gene encoding a putative pyrin-like protein causes familial cold autoinflammatory syndrome and Muckle-Wells syndrome. Nature Genetics 2001;29:301-305.

7

Table 1. Infavers Database information about R260L.

Location in the gene	exon 3
Usual protein name (complete nomenclature)	R260L (p.Arg260Leu)
Sequence change	c.779G>T
rs Number	-
Sequence	cDNA: TCTATATCCACTGTC<u>G</u>AGAGGTGAGCCTTGT
Alteration	Substitution
Base substituded	G>T
Consequence	Unknown
Functional tests	No
N Controls	74
Technique(s) used	Sequencing
Change/define RFLP	Unknown
Disease related symptoms	Symptomatic
Associated phenotype	Muckle-Wells Syndrome
Country of origin/*Ancestry*	United kingdom/*Unknown*
Reference	Neven, B Medline Abstract
Input date	2004-01-09
Contributed by	benedicte neven

Table 2. Mutagenic primer sequences.

Primer Name	Primer Sequence (5' to 3')
R260L	5'-gttctatatccactgtctagaggtgagccttgtga-3'
R260L_antisense	5'-tcacaaggctcacctctagacagtggatatagaac-3'

Table 3. Sequences producing significant alignments.

Accession	Description	Max score	Total score	Query coverage	E value	Max identity
<u>NM_0010 79821.2</u>	Homo sapiens NLR family, pyrin domain containing 3 (NLRP3), transcript variant 3, mRNA	292	292	38%	2e-76	92%
<u>NM_0011 27461.2</u>	Homo sapiens NLR family, pyrin domain containing 3 (NLRP3), transcript variant 4, mRNA	292	292	38%	2e-76	92%
<u>NM_0011 27462.2</u>	Homo sapiens NLR family, pyrin domain containing 3 (NLRP3), transcript variant 5, mRNA	292	292	38%	2e-76	92%
<u>NM_0048 95.4</u>	Homo sapiens NLR family, pyrin domain containing 3 (NLRP3), transcript variant 1, mRNA	292	292	38%	2e-76	92%
<u>NM_1833 95.2</u>	Homo sapiens NLR family, pyrin domain containing 3 (NLRP3), transcript variant	292	292	38%	2e-76	92%

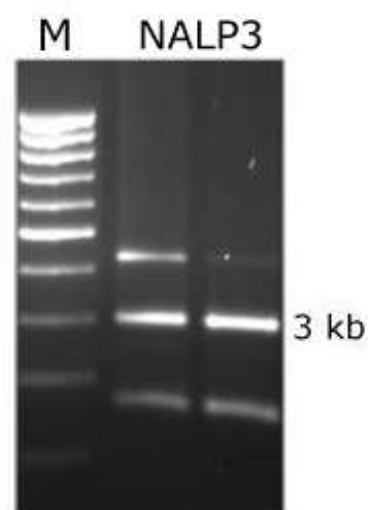|						
	2, mRNA					
NT_1671 86.1	Homo sapiens chromosome 1 genomic contig, GRCh37 reference primary assembly	292	292	38%	2e-76	92%
NW_0018 38554.1	Homo sapiens chromosome 1 genomic contig, alternate assembly (based on HuRef), whole genome shotgun sequence	291	291	38%	2e-76	92%

Figure 1. Agarose gel electrophoresis of the colony PCR product. Size of the marker gene and NALP3 gene was 10kb and 3kb, respectively.
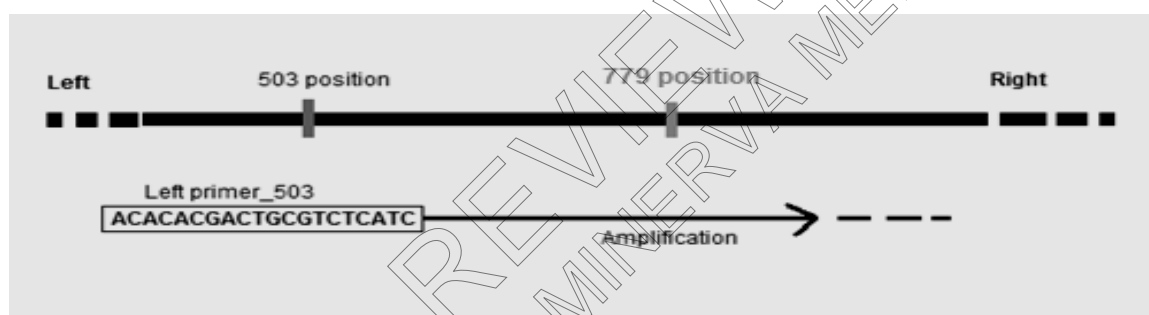

Figure 2. Position of sequence primer and target mutation position. Arrow shows the amplification movement of sequencing primer, left_503 that binds to 503 position, it will amplify and cover the target position (c.779G>T).

1