USING BALANCED TRANSLOCATION TO DISCOVER NOVEL DISEASE CAUSING GENES: POTENTIAL INVOLVEMENT OF *BASONUCLIN 2* IN CONGENITAL HEART DISEASE AND *PALMITOYLATED MEMBRANE PROTEIN 7* IN MATURITY-ONSET DIABETES OF THE YOUNG

APPROVED BY SUPERVISORY COMMITTEE

_____
Andrew Zinn, MD., PhD.

_____
Roger Schultz, PhD.

_____
Helen Hobbs, MD.

_____
Keith Parker, MD., PhD.

To my family and friends, especially my husband, Vijay, who is the best of both

USING BALANCED TRANSLOCATION TO DISCOVER NOVEL DISEASE CAUSING GENES: POTENTIAL INVOLVEMENT OF *BASONUCLIN 2* IN CONGENITAL HEART DISEASE AND *PALMITOYLATED MEMBRANE PROTEIN 7* IN MATURITY-ONSET DIABETES OF THE YOUNG

By

ELIZABETH JOYCE BHOJ

DISSERATAION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, TX

May, 2008

Acknowledgments

First, I would like to thank Dr. Andrew Zinn for being an extraordinarily supportive mentor, and for giving me the opportunity to do research in his lab on such an interesting project. It was extremely gratifying to work with a mentor who makes students his priority and will do anything in his power to make graduate school a productive and rewarding experience for them. I would also like to thank my previous mentor, Dr. Daniel Garry, for his support and guidance; having the chance to work with him and his lab members made me a far better graduate student and scientist.

I would also like to thank the members of my dissertation committee, Drs. Roger Schultz, Helen Hobbs, and Keith Parker, for their ever-helpful discussions and contributions of individual areas of expertise to my project. My committee meetings were always a source of inspired ideas and new directions for my projects.

Additionally, I would like to show my appreciation for all the members of the Zinn and Kublouai labs, who were so generous with their advice and support. Purita Ramos, Lane Jaeckle Santos, Kristen Tolson, Bassil Kublouai, Terry Gemelli, Joyce Wang have made countless contributions to this work, and I am grateful to them all. I would also like to thank Guy Bartov for answering my endless barrage of questions about FISH preparation, and his patience through

months of my trial and error. Much additional help with interpreting FISH was provided by the technicians in the cytogenetics laboratory in Veripath laboratories, for which I am thankful.

I would also like to acknowledge our very cooperative collaborators, Dr. Vidu Garg, Dr. Linda Baker, and Dr. Stefano Romero, for bringing these unique patients to our attention and providing their expertise during my studies.

I have been supported, both financially and intellectually, by the Medical Scientist Training Program at UT Southwestern, and continue to be very grateful to be a student in this program.

I would very much like to thank the family I started out with when I began this program, my mother Barbara, father Charlie, and brother Richard, for their unconditional support and understanding. I have always felt so fortunate to have the good luck to be born into such a caring and wonderful family. I would also like to thank the family I gained in the summer of 2006, when I married my husband, Vijay, who had been an invaluable source of advice, compassion, encouragement, and chocolate. His daily contributions were innumerable and invaluable, and I am thankful for the calm and comfort he continues to provide.

May, 2008

Using Balanced Translocation to Discover Novel Disease Causing Genes: Potential involvement of *Basonuclin 2* in Congenital Heart Disease *and* *Palmitoylated Membrane Protein 7* in Maturity-Onset Diabetes of the Young

Publication No. _____

Elizabeth Joyce Bhoj, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2008

Supervising Professor: Andrew Zinn, M.D., Ph.D.

Patients with balanced chromosomal translocations and disease have been used to identify novel disease-causing genes by examining the genes affected by the chromosomal breakpoints. We have identified a man with unexplained maturity-onset diabetes of the young (MODY) and a balanced translocation (7q22;10p12), and a man carrying a balanced translocation associated with partial anomalous pulmonary venous return (PAPVR) and hypospadias.

In our first proband, the translocation breakpoints were mapped to high resolution using FISH and somatic cell hybrids and the junctions amplified by PCR and sequenced. No annotated genes were disrupted by the translocation. Three of the four genes bordering the translocation breakpoints showed biallelic transcription in lymphoblastoid cells; the other gene did not have any informative SNPs. The 10p12 breakpoint was 220 kb 5' to the *Membrane Protein, Palmitoylated* (*MPP7)* gene, which encodes a protein required for proper cell

polarity. As this biological function is shared by *HNF4A*, a known MODY gene, the translocation may cause islet cell dysfunction by altering *MPP7* expression in a tissue- or a cell-specific fashion.

In our second proband, tiling array comparative genomic hybridization (aCGH) for chromosomes 9 and 13 was performed on DNA from cultured cells from the productions of conception of his offspring with karyotype (46,XX,der(9)t(9;13)(p22;q22). The results mapped both breakpoints to kilobase resolution. The chromosome 13 breakpoint lies in an intergenic region. The nearest gene, *Lim-domain Only 7* (*LMO7)*, is ~400 kb away. The chromosome 9 breakpoint lies within an intron of *basonuclin 2* (*BNC2*), a gene of unknown function with a complex pattern of alternative transcripts. Expression of *LMO7* and *BNC2* is enriched in fetal heart and penis. Transcripts from both alleles of *LMO7* and *BNC2* were present in immortalized lymphocytes with the balanced translocation. One *BNC2* missense mutation of unclear significance was identified among 10 patients with sporadic total or PAPVR.

Table of Contents

Acknowledgments

Abstract……i

Table of Contents……..iv

Publications……...……vi

List of Tables……vii

List of Figures……..……viii

List of Abbreviations……...……ix

Chapter 1. General Background …….1

A. Introduction

B. Mechanisms of Translocations

C. Effects of Balanced Translocations

D. Transmission of Balanced Translocations

E. Balanced Translocations and Disease-causing Genes

F. Objectives of the Dissertation Research

Chapter 2. Materials and Methods……11

Chapter 3. Identification of Palmitoylated *Membrane Protein 7* as a Candidate Gene for Maturity Onset Diabetes of the Young……..19

A. Introduction

B. Results

C. Discussion

Chapter 4. Identification of *Basonuclin 2* as a Candidate Gene for Partial Anomalous Pulmonary Venous Return and Hypospadias……..55

A. Introduction

B. Results

C. Discussion

Chapter 5. Conclusions and Future Directions……86

Bibliography

Vita

Publications

Publications that resulted from this thesis:

Elizabeth J. Bhoj, Linda A. Baker, Frederick F. Elder, Vidu Garg, Andrew R. Zinn (2008) *Array CGH mapping of a translocation breakpoint identifies the basonuclin 2 gene (BNC2) as a candidate for partial anomalous pulmonary venous return and hypospadias*, submitted

Elizabeth J. Bhoj, Stefano Romeo, Roger A. Schultz, Andrew R. Zinn (2008) *MODY-like diabetes associated with a balanced 7;10 translocation: potential involvement of MPP7,* submitted

Other publications:

Vijay G. Bhoj, Qinmiao Sun, **Elizabeth J. Bhoj**, Cynthia Somers, Xiang Chen, Asuncion Mejias, Ana M. Gomez, Hasan Jafri, Octavio Ramilo, and Zhijian J. Chen (2008) *MAVS and MyD88 Are essential for innate immunity but not cytotoxic T lymphocyte response against respiratory syncytial virus,* submitted

List of Tables

Table 1. Examples of breakpoint mechanisms ……..3

Table 2. Primer sequences used for screening for mutations in *BNC2*……14

Table 3. Previously described forms of MODY……22

Table 4. BACs localized by iterative FISH on Chromosome 10……...35

Table 5. BACs localized by iterative FISH on Chromosome 7…….36

Table 6. Patients screened for mutations in *BNC2*……67

Table 7. Sequencing results of *BNC2* from 10 patients with PAPVR……...68

List of Figures

Figure 1. Incidence of diabetes in the United States ……16
Figure 2. Glucose-sensitive insulin release……20
Figure 3. Pedigree and karyotype of MODY proband……...32
Figure 4. FISH of *PTF1a* region……34
Figure 5. Sequence trace across Der10 and Der7……..37
Figure 6. Alignment of normal and derivative chromosomes……...37
Figure 7. FISH of RP11-1141O6, and cen10 on chromosome 10……39
Figure 8. FISH of RP11-243L13 and RP11-1039P22 on chromosome 7……..39
Figure 9. Chromosome 10 breakpoint region, UCSC browser……..40
Figure 10. Chromosome 7 breakpoint region, UCSC browser……..40
Figure 11. Genes closest to breakpoint region on der10 and der7……40
Figure 12. Allelic expression of *MMP7*, *WAC*, and *mTERF*……42
Figure 13. Domain structure of MMP7……..44
Figure 14. Incidence of hypospadias……53
Figure 15. Pedigree and karyotype of PAPVR proband……56
Figure 16. aCGH results and genome browser view of breakpoint region……57
Figure 17. FISH results and BAC localization……..58
Figure 18. Allelic expression of *BNC2* and *LMO7*……60
Figure 19. q-PCR data for *LMO7* and *BNC2* in proband and normal controls…..61
Figure 20. q-PCR expression data for *LMO7* and *BNC2* in fetal and adult tissues……63
Figure 21. ISH of *BNC2* in mouse embryos……..66
Figure 22. Conservation of novel non-synonymous SNP in *BNC2*……...68
Figure 23. Exon/intron structure of *BNC2*……70

List of Abbreviations

aCGH	array comparative genomic hybridization
ADP	Adenosine diphosphate
ALP	*Actin-associated LIM protein*
ANKRD1/CARP	*Cardiac ankyrin protein*
APVR	Anomalous pulmonary venous return
ASD	Atrial septal defect
ATP	Adenosine triphosphate
BAC	Bacterial artificial chromosome
BMI	Body mass index
BNC1	*Basonuclin 1*
BNC2	*Basonuclin 2*
bp	base pairs
Ca	Calcium
CAPN10	Calpain-10
CCD	Charge coupled device
cDNA	complementary deoxyribonucleic acid
CEL	Carboxyl ester lipase
Cen	Centromere
CHD	Congenital heart disease
Chr	Chromosome
Csk	c-src tyrosine kinase
Ct	Cycle threshold
CTLA4	Cytotoxic T-Lymphocyte Antigen 4
CXORF6	Chromosome X open reading frame 1
DAPI	4',6-diamidino-2-phenylindole
Der	Derivative chromosome
dl	deciliter
DLG1	*Discs large 1*
DM	Diabetes mellitus
DNA	Deoxyribonucleic acid
EBV	Epstein-Barr virus
ENPP1	Encoding Ectonucleotide Pyrophosphatase Phosphodiesterase 1
F	Female

FAM	6-carboxy-fluorescein phosphoramidite
FISH	Fluorescent in-situ hybridization
FZD1	*Frizzled 1*
GAD2	*Glutamic Acid Decarboxylase 2*
GAD65	*Glutamic Acid Decarboxylase 65*
GCK	*Glucokinase*
GLUT2	*Glucose transporter 2*
Hes3	*Hairy and enhancer of split 3*
HLA	Human leukocyte antigen
HNF1	*Hepatocyte nuclear factor 1*
HNF1B	*Hepatocyte nuclear 1B*
HNF4a	*Hepatocyte nuclear factor 4A*
htRNA	heteronuclear ribonucleic acid
IDDM	Insulin-dependent diabetes mellitus
IPF1	*Insulin promoter factor 1*
K	Potassium
kb	kilobase
KCNJ11	*Potassium inwardly-rectifying channel, subfamily J, member 11*
KDR	*Kinase insert domain receptor*
kg	kilogram
LCR	Low-copy repeats
LIM	Lin11, Isl-1, Mec-3
Lin7	*Lin-7 homolog A*
LMO7	*Lim-domain only 7*
Lyp	*Lymphoid phosphatase*
M	Male
m	meter
MAGUK	Membrane-associated guanylate kinase
MAMLD1	*Mastermind-like domain containing 1*
MAVS	*Mitochondrial anti-viral signaling protein*
Mb	Megabase
MELAS	Mitochondrial myopathy, Encephalopathy, Lactic Acidosis, and Stroke-like episodes syndrome
mg	milligram
MODY	Maturity-onset diabetes of the young
MPP7	*Membrane protein, palmitoylated 7*
mtDNA	mitochondrial deoxyribonucleic acid
mTERF	*Mitochondrial transcription termination factor*
MyD88	*Myeloid differentiation primary response gene 88*

NEUROD1	*Neurogenic differentiation factor 1*
NIDDM	Non-insulin-dependent diabetes mellitus
NKX2-5	*NK2 transcription factor related, locus 5*
NLS	Nuclear localization signal
P2	Promoter region 2
PAPVR	Partial anomalous pulmonary venous return
PCR	Polymerase chain reaction
PDZ	PSD95, DlgA, zo-1
PH	Pulmonary hypertension
PHA	Phytohemagglutinin
POC	Products of conception
PPARγ	*Peroxisome proliferators-activated receptor-γ*
PTF1a	*Pancreas specific transcription factor 1a*
PTPN1	*Protein tyrosine phosphatase, non-receptor type 1*
PTPN22	*Protein tyrosine phosphatase, non-receptor type 22*
q-PCR	quantitative polymerase chain reaction
RNA	Ribonucleic acid
RT-PCR	Real time polymerase chain reaction
SIFT	Sorting intolerance from tolerance
SNP	Single nucleotide polymorphism
STS	Sequence-tagged site
TAPVR	Total anomalous pulmonary venous return
TCF1	*Transcription factor 1*
TCF2	*Transcription factor 2*
TCF7L2	*Transcription factor 7-like 2*
TFAP2B	*Transcription factor 2, AP-2*
t-SNARE	target soluble NSF attachment receptor
UCHL3	*Ubiquitin C-terminal hydrolase L3*
UCSC	University of California – Santa Cruz
UTR	Untranslated region
VNTR	Variable-number tandem repeat
VSD	Ventral septal defect
v-SNARE	vesicle soluble NSF attachment receptor
WAC	*WW domain-containing Adaptor with a Coiled-Coil Region*
Wnt	*Wingless-type*
X-EDMD	X-linked Emery-Dreifuss muscular dystrophy
ZASP	*Z-band alternatively spliced PDZ motif protein*

Chapter 1

General Background

Introduction

Chromosomal translocations occur when a portion of two or more chromosomes become relocated to a portion of a different chromosome. If no genomic information is lost in this transfer then a balanced translocation results; if there is a significant loss or gain of material then an unbalanced translocation forms. If only two chromosomes are involved then it is termed a balanced reciprocal translocation. This type of translocation has a population frequency of 1/1000 to 1/673, making it one of the most common cytogenetic variants (Van Dyke, Weiss et al. 1983). There are a virtually infinite number of translocations possible within the human genome, as any single nucleotide on either chromosome can be affected.

Mechanisms of Translocation

There are some areas of chromosomes that are more susceptible to creating breakpoints; these have been useful in the study of the mechanisms of translocations. Three of the main contributors to the formation of breakpoints are

sequences that are found on multiple chromosomes, fragile sites, and certain DNA secondary structures (Cohen, Cans et al. 1996; Shaffer and Lupski 2000; Stankiewicz and Lupski 2002) (Table 1).

It has been shown that when a sequence is found on multiple chromosomes it allows recombination between nonallelic regions of homology on those chromosomes. Usually, genetic material is exchanged between sister chromatids to allow greater variety in the gene pool, however, when exchange occurs in these homologous regions on different chromosomes, a translocation can result. The homologous regions are often high-copy number repeats like *Alu* and satellite DNA, or low-copy repeats (LCR). An interesting case of female-to-male sex reversal caused by an X;Y translocation was shown to be mediated by *Alu* repeats in the sex chromosomes (Rouyer, Simmler et al. 1987). LCRs are found throughout the genome, but there is a higher concentration at pericentromeric regions. A common (4;8) translocation resulting in Wolf-Hirschhorn syndrome is mediated through LCR leading to recombination between 4p16 and 8p23 (Wieczorek, Krause et al. 2000).

The most common recurring translocation, between chromosomes 11 and 22, has not been associated with regions of homology, but with AT-rich palindromic sequences. These palindromic sequences are comprised of two regions that are complementary to each other, and are predicted to form hairpin

secondary structures. The hairpin structures are targets for nucleases that cause double-stranded breaks, which can result in recombination and translocation (Kurahashi and Emanuel 2001). The role of these sequences has been confirmed by an additional palindromic-mediated translocation found between 17;22 that causes neurofibromatosis (Kehrer-Sawatzki, Haussler et al. 1997).

Table 1 – Examples of breakpoint mechanisms

Type of Breakpoint	**Example(s)**
A-T rich palindrome	11;22 translocation resulting in DiGeorge syndrome, 17;22 translocation resulting in neurofibromotosis
Alu DNA	X;Y translocation resulting in sex reversal
Low copy repeats	4;8 translocation resulting in Wolf-Hirschhorn syndrome

Effects of Balanced Translocations

Typically carriers of balanced reciprocal translocations may have reproductive difficulties due to unbalanced segregation, but often have no other phenotypic effects; however if either of the breakpoints is in or near a gene then the function of that gene may be affected (Kim, Herrick et al. 2005; Ligon, Moore et al. 2005).

In addition to directly perturbing transcription units, balanced translocations can alter expression of neighboring genes via position effects. Effects on gene expression can extend as far as 1 MB away from the breakpoints, and perturb *cis*-acting distant genomic elements, such as enhancers, repressors, and insulators. These elements are located upstream of the transcription unit, in the introns of the gene, or downstream from the poly-adenylation site. Also, there may be multiple elements that interact with multiple alternate promoters and exons. These elements are often very well-conserved among species because of their importance in coordinating transcription (Jamieson, Perveen et al. 2002).

Chromatin structure affected gene transcription, Chromatin can exist in states that range from being closed (heterochromatic) to open (euchromatic). Repetitive sequences, which usually are not transcribed are located in the heterochromatin. Facultative heterochromatin is found where there is transiently transcriptionaly-silenced. In regions that are highly transcribed the chromatin is in a euchromatic form. Gene transcription is also affected by acetylation, methlylation, and ubiqutination of histone tails (Egger, Liang et al. 2004).

Balanced translocation can interfere with chromatin configuration in various ways, including inactivation of neighboring genes, which is often seen with X;autosome translocations where, through interference of methlylation patterns, the inactivation of the X chromosome spreads to the autosomal chromosome (Rouyer, Simmler et al. 1987). There are known chromosomal

rearrangements that alter chromatin states that cause disease, such as facioacopulohumeral dystrophy, which causes neuromuscular degeneration of the facial and shoulder muscles (Gabellini, Green et al. 2002).

Interestingly, mutations in remote elements can produce a myriad of different phenotypes, since they can have effects on gene expression that can be partial, tissue-specific, or temporally-specific. This theory is supported by the findings that these elements are often found in tissue-specific and developmentally-regulated genes (Kleinjan and van Heyningen 2005).

Higher order transcriptional regulation may exist at the level of nuclear organization (Cremer, Cremer et al. 2006) . Microscopic studies have revealed that each chromosome occupies a characteristic territory within the nucleus (Cremer, Cremer et al. 2006). Within the nuclei of human lymphocytes, the chromosomes and chromosomal regions that are more gene-dense and more highly expressed tend to be located in the center of the nucleus, whereas chromosomes (and regions) that are gene-poor tend to reside closer to the nuclear envelope. This pattern is also seen in species other than humans, including chicken, *hydra vulgaris*, and *stylonychia lemmae* (Mateos-Langerak, Goetze et al. 2007).. It is possible that this nuclear organization is disrupted in the presence of a balanced chromosomal translocation, thus effecting in differences in gene transcription. As more is discovered about the physiology and importance of

nuclear architecture, it will be easier to hypothesize about the results of balanced translocations on this process.

If the breakpoint interrupts genes on both chromosomes, then chimeric genes or aberrant transcripts can be generated. The chimeric genes can result in dysregulated transcription or protein activity, cause a gain of novel function, or act through a dominant-negative mechanism (Nothwang, Kim et al. 2001).

It is also possible that translocations can result in genetic material being gained or lost at the junction; this is termed a cryptic unbalanced translocation (Menten, Maas et al. 2006). Fetuses with balanced translocations have up to three times as many congenital abnormalities, and patients with mental retardation have a seven-fold higher incidence of balanced translocations than the general population (Warburton 1984; Warburton 1991).

The Developmental Genome Anatomy Project (DGAP) was developed to characterize apparent balanced chromosomal rearrangements in patients with multiple congenital anomalies; the ultimate goal of this project is to map and identify new genes involved in human development (Higgins, Alkuraya et al. 2008). In a recent DGAP study of 40 patients with cytogenetically balanced translocations and clinical abnormalities, 15 (37%) had deletions identified through aCGH that ranged in length from 500 kb to 12 Mb (Higgins, Alkuraya et al. 2008). Additionally, in 18 of those 40 patients whose breakpoints were also molecularly characterized, they described 11 microdeletions between 1-16 bps,

six insertions of 2-17 bp, two duplications of 3 and 13 bps, three complex deletions/insertions, four balanced rearrangements, and only one case with no rearrangements. Therefore, it is clear that although many translocations may be initially categorized as balanced based on their cytogenetics appearance, the majority have insertions and deletions of various sizes.

Transmission of Balanced Translocations

Often unaffected or minimally affected carriers of balanced translocations are identified serendipitously through karyotypes of their progeny. Transmitting an unbalanced translocation usually has deleterious effects on reproduction, e.g., infertility due to recurrent spontaneous abortion. During meiosis in a normal cell, 23 bivalent pairs form between the homologous chromosomes; however, only 21 bivalent pairs form in cells with a balanced translocation. The two derivative and two normal chromosomes involved in the translocation form a quadrivalent structure to maximize homologous binding. After cell division these quadrivalent structures can result in balanced or unbalanced daughter cells. If there is a 2:2 alternate segregation, where two of each of the chromosomes is passed down, then those cells will either carry the balanced translocation or the normal chromosomes. When there is a 2:2 adjacent, 3:1, or 4:0 segregation of

chromosomes either a trisomy, monosomy, tetrasomy, or nullisomy results. Studies have shown that both types of 2:2 segregation are the most common, and all four possible segregation patterns are documented in both spermatocytes and oocytes (Munne, Bahce et al. 1998; Escudero, Lee et al. 2000).

As opposed to the finding that greater than 95% of numerical chromosome abnormalities are maternally-derived, about 75% of translocations are paternally-derived. It has been hypothesized that the spermatogonial cells' ability to undergo mitosis over the life-span may allow mutations to accumulate in these cells. This, coupled with the sensitivity of spermatogenesis to mutagens, may explain this bias (Chandley 1991). An interesting and unexplained exception to this ratio is *de novo* Robertsonian translocations, where the q arms of two acrocentric chromosomes produce a single chromosome, of which 90% are maternally-derived (Page and Shaffer 1997).

Balanced Translocations and Disease-causing Genes

Balanced translocations have been used to localize genes responsible for a variety of conditions. The breakpoints of the translocation are likely to mediate the disease process through disruption of the expression of genes on the breakpoint or in the vicinity. The first disease-causing gene identified by mapping the breakpoint of a balanced chromosomal translocation was in 1986 when the

gene defective in chronic granulomatous disease identified to be interrupted in a patient (Royer-Pokora, Kunkel et al. 1986). Genes responsible for a variety of conditions, including obesity, cleft palate, blepharophimosis syndrome, DiGeorge syndrome, Duchene muscular dystrophy, and congenital cataracts have been identified using this strategy (Ray, Belfall et al. 1985; Budarf, Collins et al. 1995; Yoshiura, Machida et al. 1998; De Baere, Fukushima et al. 2000; Holder, Butte et al. 2000; Jamieson, Farrar et al. 2007).

Objectives of dissertation research

Balanced translocations in patients with multifactorial diseases have been used to identify novel causative genes at the breakpoints. Diabetes, congenital heart diseases, and hypospadias are all complex diseases with evidence of a genetic contribution. The objectives of this dissertation research are to utilize two patients that have balanced translocations, one with a rare form of diabetes and the other with congenital heart disease and hypospadias, to identify and characterize novel gene defects that cause these disorders.

Specific Aim 1: To identify and characterize the causative gene defect responsible for MODY-X in a patient with a balanced 7;10 translocation

Seven genes are known to cause MODY. In 15-80% of the families with this disorder the disease-causing gene remain unknown. We have identified a

unique patient that has a balanced 7;10 translocation and MODY-X. The breakpoints of this translocation will be molecularly mapped, and genes involved will be studied.

Specific Aim 2: Identify the genetic defect causing congenital heart disease and hypospadias in a patient with a balanced 9;13 translocation

Previous research has described a number of genes involved in APVR malformation and in hypospadias, although these account for only a small portion of known cases, suggesting that other genes are also responsible. We identified a patient who has a balanced 9;13 translocation coincident with these two conditions. The breakpoints of this translocation will be mapped, and genes implicated will be interrogated through expression analysis in fetal and adult tissues.

To confirm causality between a candidate causative gene, we will screen other patients with either surgically-repaired PAPVR or TAPVR for mutations in genes involved in the translocation breakpoints. We will determine the relative contribution of the candidate genes to the incidence of APVR in the population.

Chapter 2

Materials and Methods

Oligonucleotide array Comparative Genomic Hybridization

Genomic DNA was purified from cultured cells or peripheral blood leukocytes by standard methods.(Dubos, Pannetier et al. 2008) CGH was performed by Nimblegen Systems Inc. (Madison, WI) using a whole genome array or chromosome-specific tiling arrays containing 385,000 isothermal 50 – 75mer oligonucleotide probes. The aCGH performed on leukocyte genomic DNA from SW1119 had a mean probe density of one probe per 6270 bp for the

whole genome array (Catalog No. B4366-00-01). The aCGH performed on leukocyte genomic DNA from SW1222 had a mean probe density of one probe per 255 bp for the chromosome 9 array (Catalog No. B3740001-00-01), and one probe per 225 bp for the chromosome 13 array (Catalog No. B3744001-00-01). Pooled human reference DNA from a phenotypically normal male was obtained from Promega (Madison, Wisconsin).

Fluorescence in situ hybridization (FISH)

Metaphase chromosomes from either phytohemagglutinin (PHA)-stimulated whole blood lymphocytes or Epstein Barr Virus (EBV)-immortalized lymphoblasts were dropped onto washed glass slides in fixation buffer (3:1 Methanol:Acetic Acid) and used for fluorescent in situ hybridization (FISH). Bacterial artificial chromosome (BAC) clones (BACPAC Resources, Oakland, CA) were cultured, and BAC DNA was isolated using the BACMAX DNA isolation kit (Epicentre, Madison, Wisconsin). A total of 500 μg of genomic DNA was labeled with Spectrum Orange (Vysis, Downers Grove, IL) using a nick-transcription method according to the manufacturer's instructions. After addition of 10 μg of unlabeled human Cot1 DNA (Invitrogen, Carlsbad, California) to block hybridization of repetitive sequences, the DNA was precipitated, resuspended in 10 μL LSI/Whole Chromosome Painting (WCP) hybridization

buffer (Vysis), and hybridized to dehydrated slides overnight at 37°C. Washed and dehydrated slides were mounted by using DAPI II (Vysis) and visualized with an Olympus BX-61 fluorescent microscope equipped with a CCD camera and Applied Imaging Cytovision digital image acquisition system.

Somatic cell hybrids

Epstein-Barr virus-immortalized lymphoblastoid cells lines were generated from the subject by standard methods (Dubos, Pannetier et al. 2008). They were then fused to thymidine kinase (TK)-deficient Chinese hamster cells using polyethylene glycol 4000 (Invitrogen). Unfused lymphoblastoid cells were washed away and TK^+ human-hamster colonies were selected by culturing the cells in the presence of hypoxanthine-aminopterin-thymidine. After >10 serial passages, DNA was extracted from individual clones and analyzed by PCR with sequence tag sites (STS) and microsatellite markers for chromosome segregation.

Intronic allelic expression

The intron of *MPP7* was amplified with the following primers: CATTGCACGCTACGGAGTAA and TGCTTCACACACCTGCATCT. The intron of *WW domain-containing Adaptor with a Coiled-Coil Region (WAC)* was amplified with the following primers: TCCAAATCATTGTTTCTCAAACC and AATATTAGTTGGGCGTCGTG. The intron of *mitochondrial Transcription*

Termination Factor (mTERF) was amplified with the following primers: CCCACAACTGGCCTGTTAAA and CGAGGCCGGAAGTTAGTCTT. PCR products from the proband's genomic DNA were sequenced and a heterozygous SNP identified in each gene. Heteronuclear RNA was isolated from the immortalized lymphoblasts by adding 1 ml NP-40 lysis buffer (10 mM TrisHCl pH 4.0, 10 mM NaCl, 3 mM $MgCl_2$, 0.5% NP40) and then centrifuged at 5,000xg for 1 min , then the nuclear pellet was extracted and resuspended in 1 ml NP-40 lysis buffer and centrifuged under the same conditions, after one more spin an collection cycle the nuclear pellet was resuspended in 1 ml of Tripure (Roche, Indianapolis, Indiana), and the iScript cDNA synthesis kit (BioRad, Hercules, California) was used to make cDNA. RT-PCR products were amplified from DNAse-treated cDNA isolated from lymphoblastoid cells carrying the balanced translocation as described above and sequenced. Control reactions omitting reverse transcriptase were performed to rule out amplification of contaminating genomic DNA. Electropherograms of genomic versus cDNA sequences were compared to determine whether both alleles of *Mpp7*, *WAC*, and *mTERF* were expressed.

Exonic allelic expression

The 5' untranslated region (UTR) of *LMO7* was amplified with the following primers: TCAGATAATCTTAAGGCTGTTGG and

TTTCCTCATGAGCAGTGTGC. The 5' UTR of *BNC2* was amplified with the following primers: CTTTTCAAAGTTGCTGAAATAAAA and TTGCATTTAATGGCCTCAGA. PCR products from the proband's genomic DNA were sequenced and a heterozygous base identified in each gene. RT-PCR products were amplified from DNAse-treated cDNA isolated from lymphoblastoid cells carrying the balanced translocation as described above and sequenced. Control reactions omitting reverse transcriptase were performed to rule out amplification of contaminating genomic DNA. Electropherograms of genomic versus cDNA sequences were compared to determine whether both alleles of *LMO7* and *BNC2* were expressed.

Quantitative RT-PCR

Multiple human fetal tissue RNAs were purchased (Cat # 636747 Clontech, Mountain View, CA). Total RNA was isolated from immortalized lymphoblasts using Tripure (Roche, Indianapolis, Indiana), and then reverse transcribed using the iScript cDNA synthesis kit (BioRad, Hercules, California). Expressions of messenger RNA (mRNA) for 2 exons of *BNC2*, *LMO7*, and a control gene (*cyclophilin A*), were measured by real-time PCR with the TaqMan Gene Expression Assays products Hs00214187_m1 (*BNC2* exons 5 and 6), Hs00417700_m1 (*BNC2* exons 3 and 4), Hs00245600_m1 (*LMO7*), and *cyclophilin A* (4310883E, Applied Biosystems, Foster City, CA). The *BNC2* and

LMO7 probes contain a 6-carboxy-fluorescein phosphoramidite (FAM dye) label at the 5' end of the gene and a minor groove binder and nonfluorescent quencher at the 3' end and are designed to hybridize across exon junctions. The *cyclophilin A* control probe was labeled with VIC dye and used in multiplex with the *BNC2* or *LMO7* probes. For calibration and generation of standard curves, we used cDNA derived from human skeletal muscle RNA (Clontech). Reactions were performed in triplicate, and any samples with coefficient of variation among replicates >10% were repeated.

Mutation Screening

Ten patients with surgically-repaired PAPVR or TAPVR were recruited and genomic DNA was prepared from whole blood. Primers were designed by primer3 to PCR amplify the exons 1, 2, 2a, 3, 4, 5, and 6; products were treated with ExoSAP (USB Corp, Cleveland, OH) and sequenced with the same primers (Table 2). Sequencing data were analyzed using Seqman (DNAStar, Madison, Wisconsin).

Table 2 – Primer sequences used for screening for mutations in BNC2

Exon	Forward Primer	Reverse Primer
1	GCGCGAGTGTAAATCAGTCA	CCTAGAGCTCGCCTCAGAAA
2	TGAAGTAACCCTGCACATTCC	CAAGCTTGGGGAGTCACAAT
2a	CCTTCCCAAATGAAATGCAC	CCAGAAACAAAAGTGCCCATA
3	TGGGCAATATTTATTGAAATGG	CACACACACACACGAACATGA

4	CCAGCAATATTTCTAACTCCTGTTC	GCGAGGTTTGTCAAGGACTC
5 (1/4)	GAGCACACAGAGCACCTTCA	GGTTTTGGGCTCCACATTAG
5 (2/4)	CCAATAGAAATGCCCTGACC	AGGGCTTGTGAGTGCCAGAT
5 (3/4)	AGATTTAATTCGGGCCACCT	CCGGGTCCTGTTATGCTTAG
5 (4/4)	CCAGGCATGTCTGTGAAGG	GGATTTCTTTTAGTGTGAAGTCCAA
6 (1/2)	GATACCTAGTGAAAGCTGAATCAAT	GCAAATGTTGCACATGATCC
6 (2/2)	AGGGGCTGAAGTTTCAGGAT	CCCAAGTACATAAGCGCACA

Electronic Databases and Other Internet Resources

URLs for resources referenced in this work are as follows:

UCSC Genome Browser, http://genome.ucsc.edu/cgi-bin/hgGateway

Database of Genomic Variants, http://projects.tcag.ca/variation/

Novartis Gene Expression Atlas, http://symatlas.gnf.org/SymAtlas/

Sorting Intolerance From Tolerance (SIFT) program,

http://blocks.fhcrc.org/sift/SIFT.html

Developmental Genome Anatomy Project, http://dgap.harvard.edu

primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3)

Chapter 3

Identification of *Palmitoylated Membrane Protein 7* as a Candidate Gene for Maturity Onset Diabetes of the Young

Introduction

Diabetes mellitus (DM) is a disease that causes impairment in glucose utilization and affects more than 150 million people worldwide (Hattersley and Pearson 2006). In the United States, prevalence is rising dramatically (Figure 1).

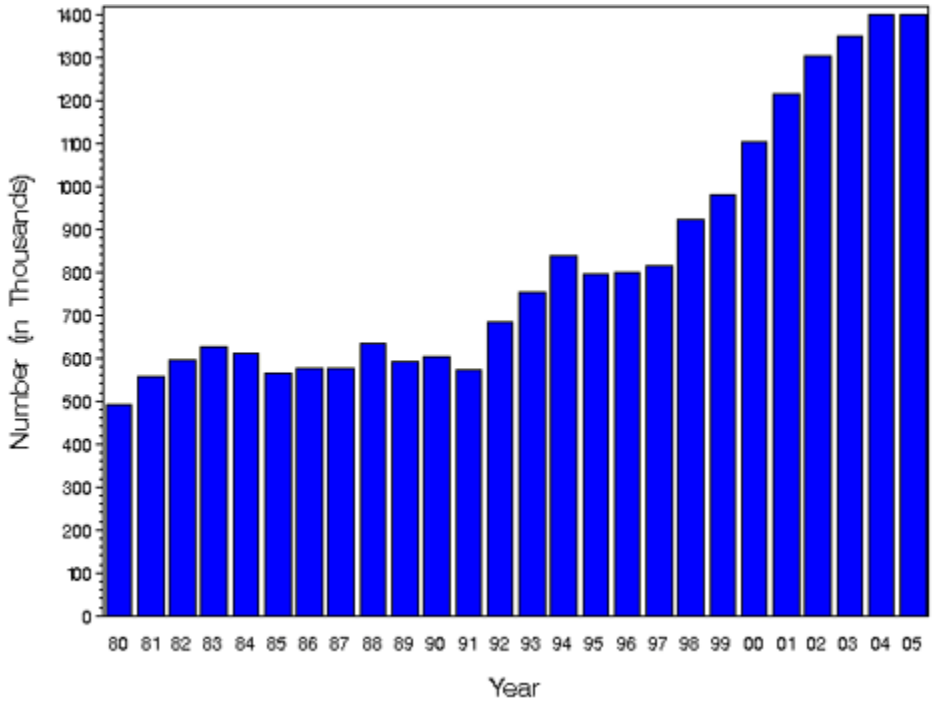


Figure 1 – Incidence of total diabetes cases in the United States is increasing significantly (http://www.cdc.gov/diabetes/statistics/incidence/fig1.htm).

Both environmental and genetic components have been described that lead to susceptibility across populations (Florez, Hirschhorn et al. 2003). These relative contributions differ between the different forms of DM, which fall into three main categories: Insulin-Dependent Diabetes Mellitus (IDDM), also called type 1 DM, Non-Insulin-Dependent Diabetes Mellitus (NIDDM), or type 2 DM, and Maturity-Onset Diabetes of the Young (MODY). These forms are all characterized by chronic hyperglycemia that perturbs carbohydrate, protein, and fat metabolism. In less severe forms, only glucose intolerance may be evident, but fulminant DM produces hyperglycemia even after fasting. In its early stages DM is often diagnosed incidentally before producing symptoms. The symptoms of short-term hyperglycemia include thirst, polyuria, blurred vision, weight loss, and excessive appetite, and more severely, ketoacidosis or nonketotic hyperosmolarity, which can lead to stupor, coma, and death if left untreated. The consequences of long-term hyperglycemia are equally devastating; damage and failure of the eyes, kidneys, blood vessels, and heart can result (Joslin, Kahn et al. 2005).

Forms of Diabetes Mellitus

Insulin-Dependent Diabetes Mellitus (IDDM), also called type 1 DM, is characterized by an autoimmune-mediated loss of beta cells in the islets cells of the pancreas. Patients with IDDM usually present in childhood with near-complete loss of insulin production secondary to the beta cell failure (Bartsocas and Gerasimidi-Vazeou 2006).

The most common type of DM is Non-Insulin-Dependent Diabetes Mellitus (NIDDM), or type 2, which is characterized by insulin resistance in peripheral tissues. It is often seen in conjunction with metabolic syndrome, which includes obesity, dyslipidemia, and hypertension (Candib 2007). It has long been realized that type 2 DM shows a familial predisposition, but few causative genes have been identified. (Winckler, Weedon et al. 2007)

There are also cases of DM that result from mitochondrial DNA mutations, some are associated with deafness, and others with marked lipodystrophy (Maassen, t Hart et al. 2007). Additionally, in maturity-onset diabetes of the young (MODY), a group of autosomal dominant forms of DM result from rare mutations that affect insulin production or action (Steiner, Tager et al. 1990). These MODY-causing defects include the inability to convert pro-insulin to insulin, formation of structurally abnormal insulin, or impaired receptor binding. These patients have mild hyperglycemia and abnormally high levels of

insulin or C-peptide, which is formed during the conversion of insulin from pro-insulin (Steiner, Tager et al. 1990).

There have also been cases of DM reported that result from insult to the pancreas as a whole. Fibrocalculous pancreatopathy, pancreatitis, pancreatectomy, pancreatic cancer, cystic fibrosis, hemochromatosis, and various endocrinopathies have all been shown to contribute to DM (Joslin, Kahn et al. 2005). Administration of certain drugs may also precipitate DM, such as glucocorticoids, nicotinic acid, diazoxide, phenytoin, and pentamidine (Ferner 1992). A number of congenital infections, such as rubella and cytomegalovirus, have also been implicated in DM (Jaeckel, Manns et al. 2002).

Role of Insulin Release in DM

There are several steps that must be coordinated to ensure beta cells can appropriately release insulin in a glucose-dependent manner (Figure 2). All of the known genetic defects that result in MODY affect steps in this glucose-sensitive insulin secretion. First, the high-capacity, low-affinity glucose transporter GLUT2 moves glucose across the beta cell membrane. As it enters glycolysis, glucose is phosphorylated by glucokinase into glucose-6-phosphate, in the rate-determining step in the pathway. The pyruvate that results is used primarily by the mitochondria to increase the ATP:ADP ratio in the cell. The beta cell is a unique

environment that allows the ratios of ATP:ADP to fluctuate in response to glucose level; most other cells adapt to maintain a constant ratio. The excess ATP produced acts as a second messenger that links glucose levels to insulin secretion through both the triggering and amplifying pathways.

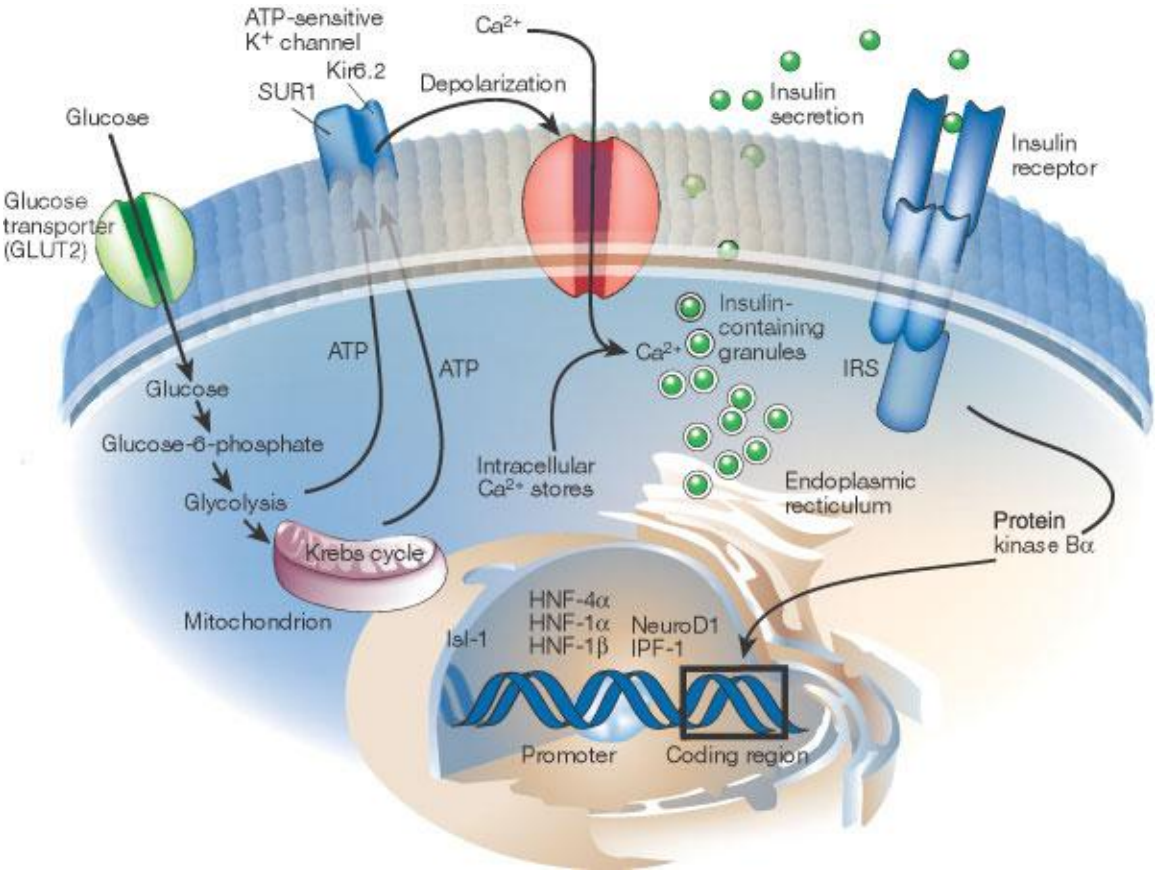


Figure 2 – Diagram of the signaling cascade needed to release insulin from a pancreatic beta cell in glucose-dependent manner (Bell and Polonsky 2001)

The triggering pathway of glucose-induced insulin secretion is mediated through the potassium and calcium channels in the electrically-excitable beta cells. Influx

of glucose level results in an increase in the ATP/ADP ratio. The Kir6.2 subunits of the ATP-sensitive potassium channels closes in response to the excess ATP (Ashcroft, Harrison et al. 1984). This lowers the potassium conductance of the plasma membrane, which allows the membrane to depolarize. This initial depolarization allows the voltage-operated calcium channels to open, calcium influxes into the cell, and further depolarization occurs. The depolarization plateaus and activates calcium action potentials. The cell then repolarizes and the membrane potential begins to oscillate with periodic openings of the calcium channels. The time the calcium channels spend open gets longer as the glucose concentration rises until the cell is depolarized at the plateau potential (Gomis, Sanchez-Andres et al. 1996).

The high concentration of intracellular calcium activates a number of pathways that lead to insulin release. First, insulin granules, stored as large dense core vesicles are recruited via the cytoskeleton to the plasma membrane. Calcium-calmodulin protein kinase II assists to mobilize the granules from the Golgi to the exocytotic site (Easom 1999). Next the granules dock to the plasma membrane via v-SNARE binding to t-SNARE. While docked, the proteins are reorganized and the high calcium concentration triggers fusion of the membranes. Insulin is then released into the extracellular space (Lang 1999). To reinforce the calcium signal, glucose-induced insulin secretion stimulates the exocytotic pathway to increase insulin secretion (Henquin 2000).

Maturity Onset Diabetes of the Young

Maturity Onset Diabetes of the Young (MODY) is an uncommon collection of monogenic insulin-secretion pathologies. It was first described in 1960 in young lean patients who had only mild diabetes, with little progression after years of follow up (Fajans and Conn 1960). As opposed to type 1 diabetes, where insulin secretion is eliminated through autoimmune insult to the pancreas, or type 2 diabetes, where the target tissues become resistant to insulin, MODY patients experience high blood glucose levels that result from a defect in glucose-sensitive insulin secretion.

The clinical criteria for MODY include the following:

1) Autosomal dominant inheritance pattern

2) Onset before age 30

3) Correction of fasting hyperglycemia without insulin for at least two years post-diagnosis

4) Absence of ketosis

MODY is often found incidentally on routine testing and may require only minimal treatment throughout the course of the disease (Fajans, Bell et al. 2001).

There have been reported cases of MODY world-wide, and the estimated contribution to the total DM population ranges from 2-5% (Velho and Froguel

1998). There are seven known genes that cause autosomal dominant MODY through haploinsufficiency (Table 3).

Table 3 – Known genetic defects causing MODY

Condition	Gene	Chromosome	Phenotype	Reference
MODY1	*HNF4A (hepatocyte nuclear factor 4 alpha*	20q	diabetes, microvascular disease	(Yamagata, Furuta et al. 1996)
MODY2	*GCK (glucokinase)*	7p	mild diabetes, impaired glucose regulation	(Froguel, Zouali et al. 1993)
MODY3	*TCF1 (HNF1 alpha, hepatocyte nuclear factor 1 alpha)*	12q	diabetes, microvascular disease	(Yamagata, Oda et al. 1996)
MODY4	*IPF1 (insulin promoter factor 1)*	13q	diabetes	(Stoffers, Ferrer et al. 1997)
MODY5	*TCF2 (HNF1B, hepatocyte nuclear factor 1 beta*	17q	diabetes, renal cysts, urogenital malformations	(Horikawa, Iwasaki et al. 1997)
MODY6	*NEUROD1 (neurogenic differentiation)*	2q	diabetes	(Malecki, Jhala et al. 1999)
MODY7	*CEL (carboxyl ester lipase)*	9q	diabetes, exocrine pancreatic dysfunction	(Raeder, Johansson et al. 2006)

MODY1 is caused by mutations in *Hepatocyte Nuclear Factor 4 alpha (HNF4a)*, an orphan nuclear receptor gene (Giuffrida and Reis 2005). Although it accounts for only about 4% of MODY cases, it was the first molecular defect

described that causes the disease. A large cohort with a nonsense mutation in *HNF4a* has been followed for over 40 years (Hattersley and Pearson 2006). Disease-causing mutations in *HNF4a* interfere with the ability of HNF4a to bind to the *HNF1a* promoter, and disrupt a regulatory loop between the two transcription factors (Gragnoli, Lindner et al. 1997; Hansen, Parrizas et al. 2002; Servitja and Ferrer 2004). These two key transcription factors control the expression of hundreds of genes in the islets, including the genes encoding insulin, glucokinase, GLUT2, and aldolase (Kulkarni and Kahn 2004; Odom, Zizlsperger et al. 2004). The disruptions in the normal transcriptional control of these proteins cause less insulin to be secreted in response to a rise in blood glucose, resulting in these patients generally having an increased blood glucose level. Approximately 30% of the patients require external insulin administration, and microvascular complications are common (Pearson, Boj et al. 2007).

Another function of HNF4a is in the formation and maintenance of tight junctions and establishment of polarized epithelial morphology. When HNF4a is overexpressed in embryonal carcinoma cells, it causes the formation of tight junctions in a dose-dependent manner (Spath and Weiss 1998; Chiba, Gotoh et al. 2003). Tight junction-associated proteins are upregulated in islets during maturation and are present in mature beta cells. Tight junctions between the b-cells may be necessary for proper insulin secretion (Orci, Unger et al. 1973; Orci, Malaisse-Lagae et al. 1975; Collares-Buzato, Carvalho et al. 2004). Glucose also

upregulates tight junctions in a dose-dependent manner in cultured rat islet cells in culture (Semino, de Gagliardino et al. 1987). Tight junctions may be required to separate the high concentrations of glucagon, insulin, and somatostain in the apical surface from their receptors on the basal surface and ensure there is no pathological inhibition of secretion through autoregulation (Unger, Dobbs et al. 1978; Charollais, Gjinovci et al. 2000; Calabrese, Caton et al. 2004).

Glucokinase (GCK) is a hexokinase-like enzyme that catalyzes the first and rate-limiting step in the glycolytic pathway, the phosphorylation of glucose to glucose-6-phosphate. GCK is therefore the regulatory enzyme that controls the amount of glucose available to the glycolytic pathway, and plays the role of a glucose sensor (Giuffrida and Reis 2005). When this glucose-sensing system is lost, the cell cannot secrete insulin in a glucose-responsive manner. As a consequence, these patients become hyperglycemic. Mutations in *GCK* cause MODY2, which accounts for about 20% of known MODY cases (Froguel, Zouali et al. 1993; Frayling, Evans et al. 2001). The mutations are spread over the entire gene, and have variable effects on the enzyme activity. Despite the variability in the functional effect of these mutations, the phenotype of MODY2 patients is remarkably consistent. The patients with mutations in glucokinase have only a slight elevation in fasting glucose, which does not require treatment (Hattersley and Pearson 2006).

MODY 3 is caused by mutations in *Hepatocyte Nuclear Factor 1 alpha (HNF1a)*, also known as *Transcription factor 1 (TCF1)* (Yamagata, Oda et al. 1996). About 63% of MODY in Caucasians is caused by mutations in *HNF1a* (Frayling, Evans et al. 2001). As described above, *HNF1a* forms a transcriptional activation loop with *HNF4a* to control the transcription of many islet cell genes. As with patients with HNF4a mutations, these patients experience a rise in blood sugar due to a defect in the transcriptional control of glucose-sensitive insulin secretion. Unlike the clinical course of MODY2, patients with MODY3 have progressive beta cell dysfunction, and by the third decade of life have usually developed overt diabetes and its attendant complications (Hattersley and Pearson 2006).

Insulin Promoter Factor 1 (IPF1) is a homeodomain transcription factor that plays a role in the embryonic development of the pancreas (Stoffers, Ferrer et al. 1997). Mutations in *IPF1* cause MODY 4, which can present with a wide range of severity of insulin secretion defective diabetes symptoms in heterozygotes, and pancreatic agenesis in homozygotes (Giuffrida and Reis 2005). It accounts for less than 1% of known MODY mutations (Hattersley and Pearson 2006).

Hepatocyte nuclear factor 1 beta (HNF1b), also known as *Transcription Factor 2 (TCF2)* causes MODY5, a unique syndrome that includes diabetes, renal cysts, and urogenital malformations (Horikawa, Iwasaki et al. 1997). These

mutations account for about 2% of MODY patients (Hattersley and Pearson 2006). There is heterogeneity of the severity of the symptoms, even among carriers of the same mutation (Edghill, Bingham et al. 2006). Generally, the beta-cell dysfunction in these patients is more severe than in those with mutations in *HNF1a,* and they also show signs of hepatic insulin resistance (Brackenridge, Pearson et al. 2006). Additionally, it has been shown that these patients have reduced beta-cell mass secondary to defects in fetal pancreatic development (Maestro, Boj et al. 2003).

MODY6 is caused by mutations in *Neurogenic Differentiation 1 (NEUROD1),* a helix-loop-helix transcription factor essential for normal islet cell development. MODY6, a very rare form, is characterized by variable-severity diabetes in isolation (Malecki, Jhala et al. 1999).

Finally, MODY7, caused by mutations in *Carboxyl Ester Lipase (CEL)* is somewhat controversial as a true form of MODY because of the involvement of the exocrine pancreas. By either definition, it accounts for only a small proportion (<2%) of the patients with MODY (Hattersley and Pearson 2006). CEL is an enzyme found in high concentrations in pancreatic secretions and is responsible for the hydrolysis of cholesterol esters in the duodenum. These patients present in the third decade of life with diabetes due to impaired insulin secretion, and they also have pancreatic exocrine defects. Interestingly, CEL is not expressed in

islets, and its exact role in diabetes remains unclear (Raeder, Johansson et al. 2006).

In addition to these known mutations, there is also a population of MODY patients who have no identifiable mutations in any of the known causative genes; they are considered to have MODY-X. The rate of MODY-X is variable among ethnicities, ranging from 20% in Caucasians to 80% in Japanese (Fajans, Bell et al. 2001). It is thought that these patients harbor mutations in yet-to-be described MODY genes.

Genetics of Type 1 DM

There are a number of genes that contribute to Type 1 diabetes; as would be expected with an autoimmune disease, many are involved in modulating the immune system (Trucco and Dorman 1989; Pugliese 2004). The first to be discovered was the HLA region on chromosome 6p21 (Onengut-Gumuscu and Concannon 2005). The HLA region has been known to modulate disease risk since 1974, and may account for up to half of the familial clustering seen with Type 1 DM. The correlation with certain HLA haplotypes seems to be strongest within European populations, with the average sibling risk ratio between 2.5 and 3.6 (Nerup, Platz et al. 1974).

The second region associated with Type 1 DM is the 5' region of the *insulin* gene locus, where a variable-number tandem repeat (VNTR) and SNP markers were discovered. A shorter VNTR allele is linked to a relative risk of sibling risk ratio of 1.29. When examining the explanation for this finding, it was discovered that the number of repeats in the VNTR affects the variability in the steady state level of *insulin* gene transcripts among individuals. These *insulin* gene transcripts are found in both the pancreas and the thymus, where they are thought to allow immunological tolerance to occur. It is hypothesized that the risk allele alters the levels of *insulin* gene transcripts in the thymus, therefore the immune system processes there are perturbed, and tolerance to insulin is more easily disrupted (Lucassen, Julier et al. 1993).

The third region implicated in raising the risk of Type 1 DM contains *CTLA4*, which is involved in the T cell co-stimulatory process. It carries a modest sibling risk ratio of 1.01, but is often significant in linkage studies (Ueda, Howson et al. 2003). The most recent gene associated with Type 1 DM is *PTPN22*, which encodes the lymphoid-specific tyrosine phosphatase, Lyp. *PTPN22* appears to affect many autoimmune disorders because it has also been associated with rheumatoid arthritis, systemic lupus erythematosus, and thyroiditis. A coding SNP found in Type 1 DM patients causes *PTPN22* to interact less with a kinase, *Csk*, which may lower the activation threshold for T cells (Bottini, Musumeci et al. 2004).

Genetics of Type 2 DM

There is some controversy about the genes that contribute to the risk of developing Type 2 DM (Frayling 2007; Owen and McCarthy 2007; Watanabe, Black et al. 2007). The first definitively-causative gene was found by the discovery of an association between both BMI and insulin sensitivity, and a polymorphism in *peroxisome proliferators-activated receptor-gamma (PPAR-γ).* PPAR-γ acts as a nuclear receptor that plays a key role in adipocyte differentiation (ref). It is also the point of action for the thiazolidinedione drugs used to raise insulin-sensitivity in Type 2 DM patients. A lower frequency allele has a protective effect against developing Type 2 DM, presumably by lower transactivation of downstream genes, which raises insulin sensitivity (Deeb, Fajas et al. 1998). The relative risk in the general population for homozygous carriers of the deleterious allele is 1.25; in European populations 77% of people are carriers of this allele (Altshuler, Hirschhorn et al. 2000).

Calpain-10 (CAPN10), on Chromosome 2q, is a calpain-like cysteine protease that has also been linked to Type 2 DM (McCarthy 2003). Three SNPs were found in a Mexican-American population that created a high risk haplotype, and another SNP in Europeans has also been shown to play a role in DM. Within the Mexican-American group the odds ratio is 2.13, but falls to 1.17 within the

European population (Hanis, Boerwinkle et al. 1996). Due to the low odds ratio this is not a universally accepted causative gene for Type 2 DM.

The gene with the largest impact on the heritability of Type 2 DM was uncovered by DeCode in 2006; it is *transcription factor 7-like 2 (TCF7L2)* on chromosome 10. The odds ratio associated with the inheritance of each risk allele is 1.4. This gene has also found to be associated with progression from impaired glucose tolerance to diabetes (Florez, Jablonski et al. 2006). TCF7L2 is thought to play a role in the regulation of proglucagon gene expression in enteroendocrine cells, and mediate risk through this process (Grant, Thorleifsson et al. 2006).

Two genes on chromosome 20 have been associated with Type 2 DM: *protein tyrosine phosphatase, non-receptor-type 1 (PTPN1)* and *HNF4A.* Sequence variations in PTPN1 were initially reported to be associated with insulin-resistance in Hispanic Americans, but this observation failed to be replicated in other populations (Bento, Palmer et al. 2004). *HNF4A*, the gene responsible for MODY1, was considered a candidate gene due to it's associated with diabetes. Initial studies failed to find an association (Frayling, McCarthy et al. 2000), but a SNP was found in the promoter region in the Ashkenazi Jewish cohort that appeared to relate to acute insulin response (Love-Gregory, Wasson et al. 2004). This association has been replicated in other populations, but the association remains controversial (Damcott, Hoppman et al. 2004; Weedon, Owen et al. 2004; Bagwell, Bento et al. 2005; Hansen, Rose et al. 2005).

A common variant of *HNF1A*, found only in the Oji-Cree Indians that causes MODY3, also confers risk for type 2 diabetes (Triggs-Raine, Kirkpatrick et al. 2002). A gene that causes severe syndromic neonatal diabetes, transient neonatal diabetes, and raises the risk of Type 2 DM, encodes Kir6.2, a component of the beta cell potassium ATP channel. Disruption of this gene, *KCNJ11*, upsets normal glucose-stimulated insulin release when the ATP-mediated signal cannot be transduced into channel closure (Gloyn, Weedon et al. 2003). Interestingly, since sulphonylurea drugs act on channel closure, it was found that patients with these mutations respond strongly with improved glycemic control when they are administered (Pearson, Flechtner et al. 2006).

Another known MODY gene, *GCK*, has been shown to have a common promoter variant that affects birth weight and fasting glucose (Weedon, Clark et al. 2006). A recently-identified candidate gene *is Encoding Ectonucleotide Pyrophosphatase Phosphodiesterase 1 (ENPP1),* which inhibits insulin signaling (Meyre, Bouatia-Naji et al. 2005). It is likely that additional causative genes will be discovered in the near future, as a recent study has suggested four new risk loci to be explored, and many genome-wide association studies are currently underway (Sladek, Rocheleau et al. 2007).

It was hypothesized that the genes that cause MODY could also contribute to the genetic susceptibility towards type 1 and 2 diabetes, which comprise the vast majority of DM in the population. However, multiple studies have failed to

conclusively show this connection beyond a few individual examples, and large-scale non-biased genome-wide linkage analysis has identified several alternate candidate genes for type 1 and 2 DM not implicated in MODY (Florez, Hirschhorn et al. 2003; Saxena, Voight et al. 2007). However, the identification of MODY genes has provided important insights into molecular mechanisms of glucose homeostasis. Interestingly, there is some evidence that a genetic overlap between type 1 and type 2 diabetes exists that is mediated by common pathophysiological insults (Leiter and Lee 2005).

Results

The proband was found at age 32 to have incidental hyperglycemia (serum glucose 220 mg/dL) during an evaluation for a minor gastrointestinal ailment. He was given a probable diagnosis of MODY because of his relatively low BMI (26.8 kg/m^2), lack of GAD65 antibodies that would suggest autoimmune-mediated diabetes, and evidence of low levels glucose-stimulated insulin secretion and a normal level of insulin sensitivity obtained through extensive clinical testing. Over the subsequent eight years he was treated with diet alone, and based on surrogate indices derived from insulin and fasting blood glucose measurements, his estimated beta cell insulin secretory function secretion decreased by only 8%, which is consistent with a diagnosis of MODY.

A prenatal karyotype for advanced maternal age showed his daughter

carries a balanced 7q22;10p12 translocation. Testing of the parents showed that she inherited the translocation from the proband (Figure 3). The daughter is 8 years old and in good health, with no significant past medical history. Her parents have not allowed any clinical or research testing to be performed on her.

Our proband did not meet the strict criteria for the diagnosis of MODY. His age at diagnosis was 32, and there was no clear pattern of autosomal dominant inheritance. The lack of inheritance would be expected if the translocation caused MODY and occurred *de novo*. His family history is notable for type 2 diabetes mellitus in his mother, associated with obesity (BMI 30.8 kg/m^2), hypertension, hypertriglyceridemia, and macrovascular complications. Her beta cell insulin secretory function decreased by 50% over the nine years since her diagnosis at age 55. Since, based on his insulin secretion and sensitivity clinical studies, our patient clearly has an insulin secretion defect, and not decreased insulin sensitivity, and since his disease has not markedly progressed in the eight years since his diagnosis despite lack of therapy, we concluded that he and his mother have different causes of their diabetes. The mother had normal chromosomes by report. The father, who had no reported evidence of diabetes, was deceased so his chromosomal status is unknown. No other immediate family members have agreed to have their chromosomes examined, and none have reported a history of diabetes.

Since our patient has an insulin secretion defect, no evidence of insulin resistance, and an indolent disease course, he and his mother most likely have distinct etiologies of their diabetes. The daughter who carries the translocation is not diabetic, but it is possible she will develop MODY-like diabetes later in life.

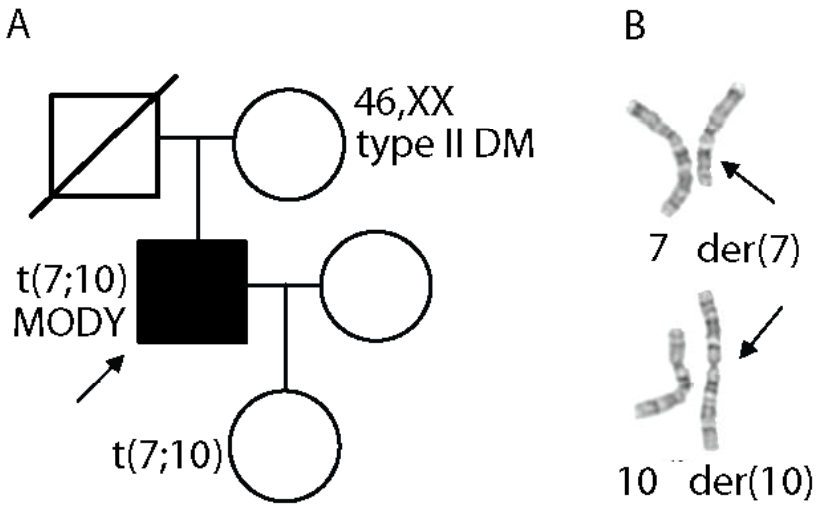


Figure 3- **A**. Pedigree of proband with MODY-like diabetes and a balanced 7;10 translocation. **B**. Partial G-banded karyotype of proband showing normal and derivative chromosomes 7 and 10. Arrows indicate breakpoints.

Coding sequences of the genes causing MODY1-4 and MODY6 (*GCK*, *HNF1A*, *IPF1*, *NEUROD1*, and *HNF4A*) were sequenced in the proband to rule out known causes of MODY with compatible clinical presentations. No mutations were found in any of these genes. The *HNF1B* gene was not sequenced since the

patient had none of the pathogneumonic features of this form of MODY. Two candidate genes in the cytogenetic vicinity of the translocation breakpoint at 10p12, *PTF1A* and *GAD2 (GAD65)*, were investigated. BAC clones spanning each gene were hybridized to metaphase spreads from the proband's lymphocytes. Neither gene was deleted or disrupted by the translocation. (Figure 4).

We performed additional in situ hybridization studies on the patient's lymphocytes using BAC clones from chromosomes 7 and 10 to narrow the location of the breakpoints. (Tables 4 and 5) Concurrently, we generated somatic cell hybrids of the proband's lymphocytes and hamster cells. We obtained one hybrid clone containing the derivative 10 chromosome but not the normal chromosome 10 or the derivative 7 chromosome. We then used this hybrid clone to map the breakpoints by STS content mapping using PCRs for sequences near the 7q and 10p breakpoints delineated by FISH.

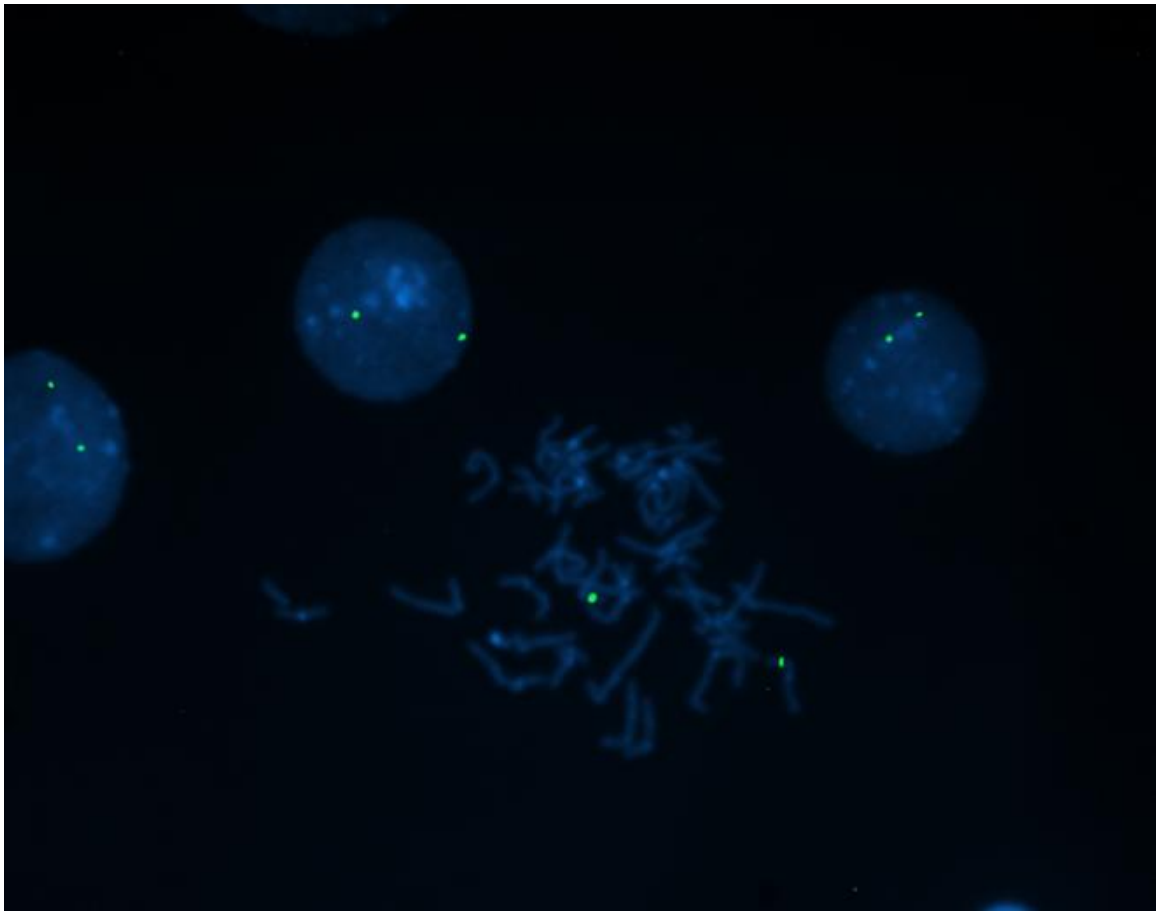


Figure 4 – FISH performed on proband's lymphocytes that shows two signals of equal strength on the PTF1A gene

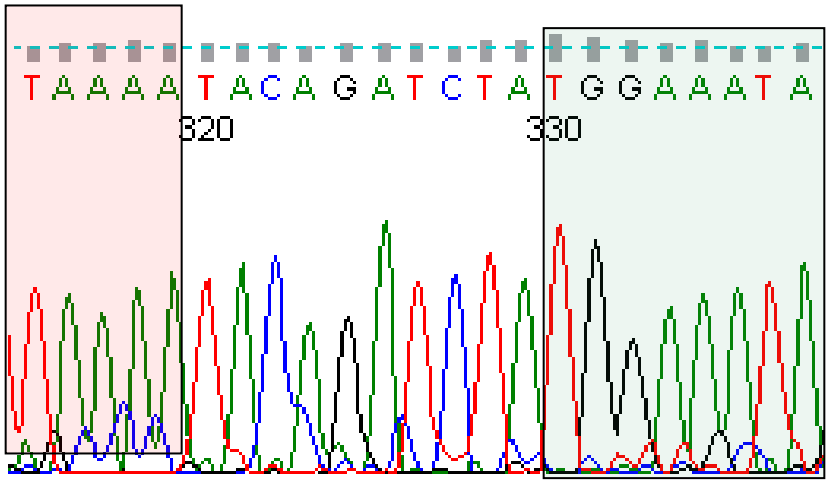
Probe name	Location	Position
Chr10 Centromere	Chr10 Centromere	Telomeric
RP11-842L14	chr10:23,437,119-23,631,605	Telomeric
RP11-93807	chr10:23,455,030-23,651,932	Telomeric
RP11-8M18	chr10:24,468,137-24,627,073	Telomeric
RP11-145L14	chr10:24,492,615-24,668,907	Telomeric
RP11-20I6	chr10:25,558,649-25,686,425	Telomeric
RP11-29I13	chr10:25,629,405-25,795,512	Telomeric
RP11-66P13	chr10:26,528,583-26,697,629	Telomeric
RP11-151P2	chr10:27,551,855-27,699,969	Telomeric
RP11-691I13	chr10:28,119,597-28,280,539	Telomeric
RP11-218D6	chr10:28,204,715-28,410,912	Telomeric
RP11-1001A15	chr10:28,585,303-28,798,773	Telomeric
RP11-1141O6	chr10:28,743,300-28,897,829	Breakpoint
RP11-49D4	chr10:28,937,787-29,106,019	Centromeric
RP11-899L22	chr10:29,097,773-29,268,197	Centromeric
RP11-670A13	chr10:29,273,767-29,484,306	Centromeric
RP11-91C12	chr10:29,523,209-29,688,334	Centromeric
RP11-89H9	chr10:30,639,662-30,790,236	Centromeric
RP11-79K19	chr10:33,372,797-33,522,133	Centromeric
RP11-453N3	chr10:38,908,842-39,115,033	Centromeric

Table 4 –BACs localized by Iterative FISH on Chromosome 10

Probe name	Location	Position
Chr7 Centromere	Chr7 Centromere	Telomeric
RP11-146I22	chr7:44,087,312-44,251,557	Telomeric
RP11-52M17	chr7:44,476,234-44,657,334	Telomeric
RP11-105B9	chr7:44,657,350-44,836,117	Telomeric
RP11-302H15	chr7:44,695,197-44,868,875	Telomeric
RP11-451M14	chr7:77,325,125-77,498,804	Telomeric
RP11-91E1	chr7:78,386,257-78,555,541	Telomeric
RP11-115O3	chr7:80,346,611-80,494,540	Telomeric
RP11-90N9	chr7:83,058,388-83,205,508	Telomeric
RP11-533L22	chr7:84,729,613-84,904,491	Telomeric
RP11-46O13	chr7:88,215,574-88,393,463	Telomeric
RP11-67H10	chr7:90,550,561-90,707,273	Telomeric
RP11-243L13	chr7:90,766,932-90,924,661	Breakpoint
RP11-1039P22	chr7:90,818,304-90,998,927	Breakpoint
RP11-165L2	chr7:91,126,014-91,304,379	Centromeric
RP11-888H2	chr7:92,050,250-92,231,711	Centromeric
RP11-25N17	chr7:92,643,302-92,807,137	Centromeric
RP11-7B9	chr7:93,552,399-93,722,001	Centromeric
RP11-127F3	chr7:94,960,563-94,960,933	Centromeric
RP11-79N11	chr7:95,000,414-95,163,199	Centromeric
RP11-94N6	chr7:95,563,312-95,716,651	Centromeric
RP11-611N5	chr7:97,362,119-97,362,639	Centromeric
RP11-10D8	chr7:98,261,079-98,410,011	Centromeric
RP11-645N11	chr7:102,351,684-102,507,525	Centromeric
RP11-46J20	chr7:106,974,619-107,152,110	Centromeric

Table 5 –BACs localized by Iterative FISH on Chromosome 7

Ultimately we identified sequences abutting both breakpoints and used these sequences to design PCRs that amplified both junction fragments. Alignment of the sequences of the PCR products to the reference human genome sequence revealed that the breakpoints were at chr10:28,832,302 with a four nucleotide deletion (chr10:28,832,303-28,832,306) on the derivative 7 chromosome, and at chr7:90,883,582 with a ten nucleotide insertion of TACAGATCTA on the derivative 10 chromosome. Thus, to a first approximation, the translocation was molecularly as well as cytogenetically balanced (Figures 5 and 6).



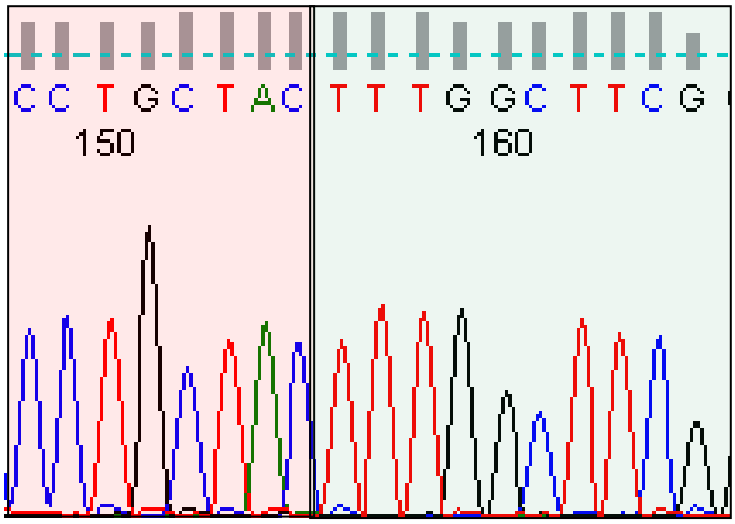


Figure 5 – a) Sequence Trace Data across the breakpoint on Der10 demonstrates a 10bp insertion b) Sequence Trace Data across the breakpoint on Der7 demonstrates a 4 bp deletion

```
chr10    CTCCCGTCTACGTGC---------------TTTTAATGCAG
der(10)  aacagtaattactatttccaTAGATCTGTATTTTAATGCAG
der(7)   CTCCCGTCTAC---------tttggcttcggtcaccattta
chr7     aacagtaattactatttccatttggcttcggtcaccattta
```

Figure 6 – Alignment of normal and derivative chromosome regions surrounding the breakpoint demonstrate a 10bp insertion on der(10) and a 4 bp deletion in der(7)

To confirm these breakpoints we performed FISH on the patient's transformed lymphoblasts with BAC clones RP11-1141O6 (chromosome 10) and a combination of RP11-243L13 and RP11-1039P22 (chromosome 7 breakpoint). Both hybridizations showed three signals, as expected (Figures 7 and 8).

Although RP11-243L13 was predicted to cross the chromosome 7 breakpoint delineated by PCR, this BAC clone gave only two signals, probably because of the abundance on one side of the breakpoint of repetitive sequences whose hybridization was suppressed by Cot1. Interestingly, a commercial chromosome 10 centromere probe (Vysis) used as a control hybridized strongly to the derivative 10 chromosome but showed barely detectable hybridization to the normal chromosome 10. G-banded karyotyping of the same metaphase cells confirmed the presence of the derivative 10 and the normal chromosome 10. The 10cen FISH result most likely represents the presence of a rare polymorphism in the centromere repeat sequence of the normal chromosome 10. This phenomenon has not been reported for this chromosome 10 probe (Vysis, personal communication) but has been seen previously with other centromere probes (Schultz, personal communication).

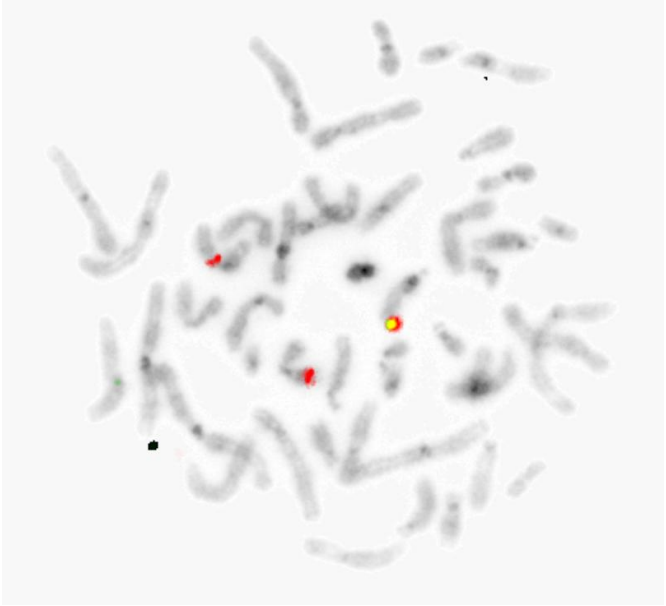


Figure 7 - RP11-1141O6, and cen10, confirms the breakpoint location on Chromosome 10

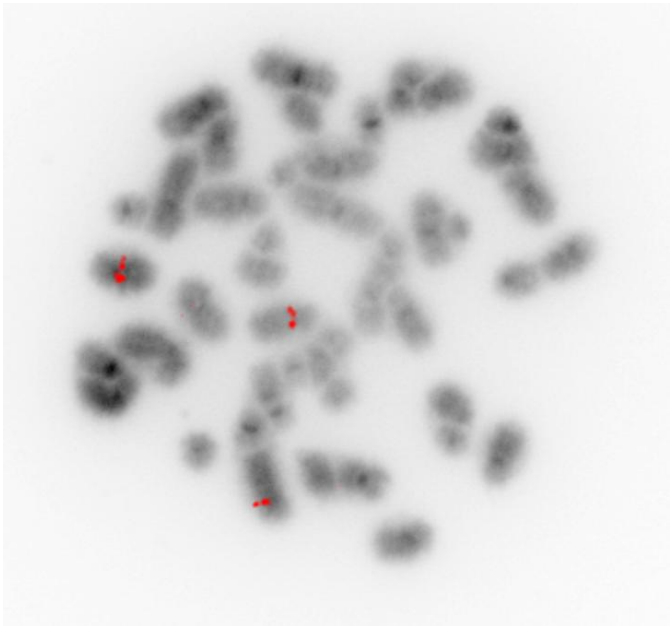


Figure 8 - RP11-243L13 and RP11-1039P22, confirms the breakpoint location on Chromosome 7

Neither breakpoint directly disrupted any known protein coding gene, microRNA, or other annotated functional genomic element. We therefore investigated genes neighboring the breakpoints that could be subject to position effects (Figures 9, 10,11).

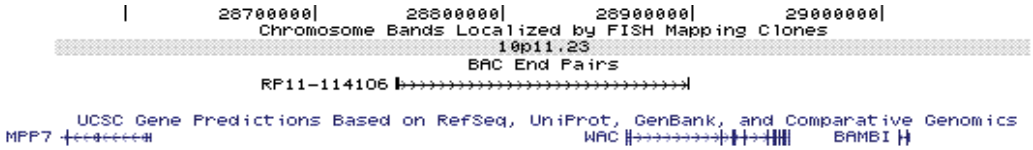


Figure 9 – Chromosome 10 breakpoint region, UCSC browser

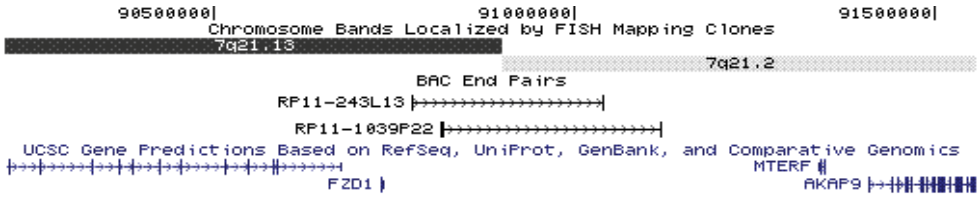


Figure 10 – Chromosome 7 breakpoint region, UCSC browser

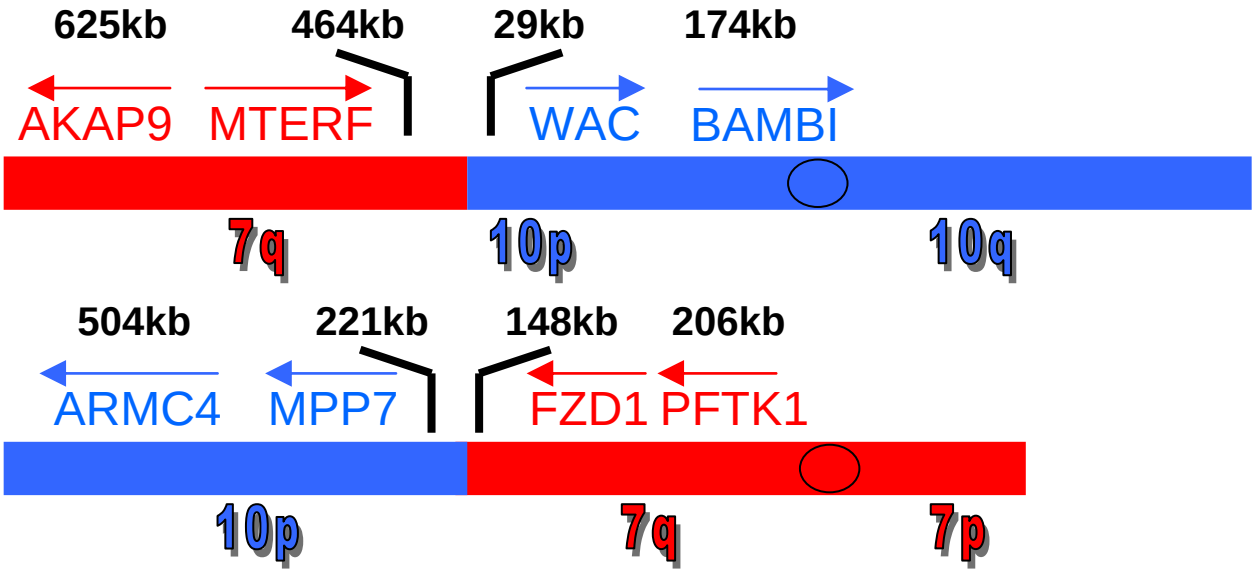
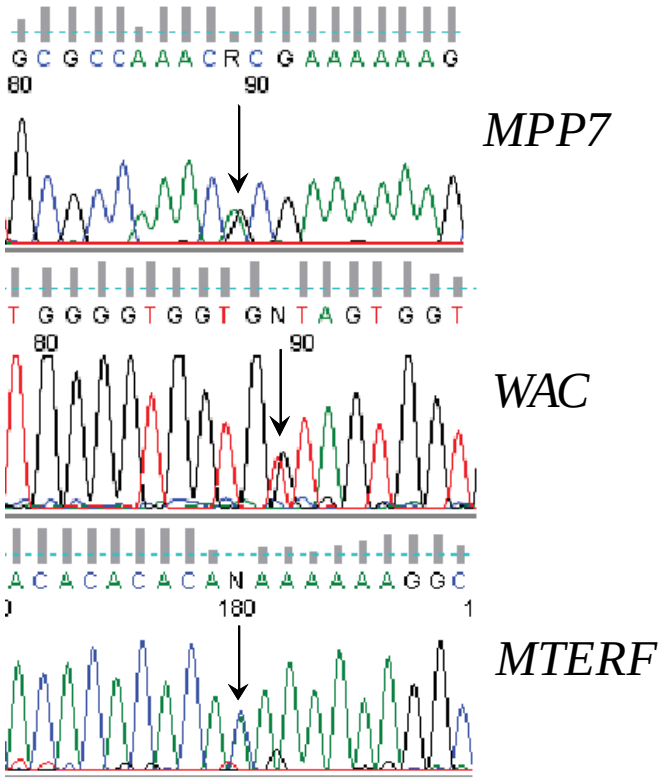


Figure 11 – Genes closest to breakpoint region on der10 and der7

None of these genes are known to play a role in pancreatic islet cell function. To examine whether the translocation affected expression of these genes, we attempted to identify expressed polymorphisms. We sequenced the exonic regions but did not find any heterozygous variations in either coding or untranslated regions. We then sequenced intronic regions of *Membrane Protein, Palmitoylated (MPP7)*, *mitochondrial Transcription Termination Factor (mTERF)*, and *WW domain-containing Adaptor with a Coiled-Coil Region (WAC)* to identify heterozygous SNPs that could be used to examine allelic expression in hnRNA, as recently described (Gimelbrant, Hutchinson et al. 2007). *Frizzled (FZD1)* comprises only one exon and thus could not be assayed in this fashion. The other



three genes, *MMP7*, *WAC*, and *mTERF*, all showed biallelic transcription, with approximately equal abundance of the two alleles (Figures 12a, 12b, and 12c), with a peak height ratio similar to that of genomic DNA sequences. Thus there did not appear to be any major position effects on expression of these genes, at least in lymphoblastoid cells.

Figure 12a,b,and c – Electropherograms showing heterozygous nucleotides (arrows) in RT-PCR products from *MPP7*, *WAC*, and *MTERF* hnRNA indicating biallelic transcription.

Discussion

An Italian patient was identified with a clinical phenotype consistent with MODY who also had a balanced translocation. Since the most common MODY mutations in the Italian population are in *GCK* (22%) and *HNF1a* (63%), initially

we screened these genes for mutations (Frayling, Evans et al. 2001). We also sequenced *IPF1*, *NEUROD1*, and *HNF4a*, since mutations in these genes cause a similar phenotype. We then hypothesized that the MODY-like diabetes in the proband resulted from disruption of a gene by his balanced translocation. Somatic cell hybridization was used to clone the breakpoint, and we confirmed our findings using aCGH. No known MODY genes were discovered to be disrupted by the balanced translocation, as had been reported for another patient with MODY1 who had a balanced translocation (3p21.2;20q12) that disrupted regulatory sequences upstream of *HNF4A* (Gloyn, Ellard et al. 2002).

The cytogenetic breakpoints of our patient's translocation suggested two potential candidate genes, *Pancreas Specific Transcription Factor 1 A (PTF1A)*, located on 10p12.2, and *Glutamate Decarboxylase 2 (GAD2)*, located on 10p12.1. *PTF1A* homozygous mutation carriers have severe neonatal DM and cerebellar agenesis, and heterozygous carriers of the mutation have a much higher incidence of adult-onset DM than the general population (Hoveyda, Shield et al. 1999; Sellick, Barker et al. 2004). *GAD2*, which catalyzes the production of gamma-aminobutyric acid from L-glutamic acid, contains a SNP in its promoter associated with increased BMI and plasma glucose levels, and is a target for islet cell auto-antibodies in type 1 diabetes (Boesgaard, Castella et al. 2007). However, when we used BAC clones containing these genes to perform FISH, we were ultimately able to show that the chromosome 10 breakpoint was at least 5

megabases away from the coding regions of both these gene. This distance is much greater than the maximum distance (~1 Mb) over which position effects have been described for other human genetic diseases (Kleinjan and van Heyningen 2005). GOOD

Somatic cell hybrid technique allowed us to identify the molecular characteristics of the breakpoint, which were confirmed by FISH. No significant loss or gain of genetic material was identified at the breakpoint: a 4 bp deletion from ch10 and a ten bp insertion found on der10. No motifs were identified at the breakpoint that would be expected to predispose the region to translocations, including low copy number sequences, fragile sites, and certain DNA secondary structures, such as AT-rich regions that lead to the formation of hairpins. (Cohen, Cans et al. 1996; Shaffer and Lupski 2000; Stankiewicz and Lupski 2002). Thus, is it likely that this, like most balanced translocations, was a sporadic event unique to this family. Consistent with this scenario was the finding that the mother did not have the deletion and there is no family history of miscarriages or infertility, although it remains possible that the *de novo* translocation was inherited from the paternal side of the family. Since the father is deceased we were unable to evaluate him.

No genes were identified at the breakpoint, but there were four genes located within 1 MB, the previously-described limit for position effects (Jamieson, Perveen et al. 2002; Kleinjan and van Heyningen 2005). We

examined these four genes -- *MPP7*, *mTERF*, *FZD1*, and *WAC*-- for position effects. No evidence was found to support aberrant expression of any of the four genes in lymphoblastoid cells. It is possible that these genes may be mis-expressed in other tissues, such as the pancreas, that would not be demonstrated by our data, which was performed only on lymphoblastoid cells.

One gene 221 kb downstream of the breakpoint on 10p, *MPP7*, merited further consideration as a candidate MODY gene. *MPP7* was discovered in 2004 through its similarity to the other members of the Membrane Protein Palmitoylated family (Katoh 2004). This family of genes encode membrane-associated guanylate kinases (MAGUK). This family of proteins are located at areas of cell-cell contact, where they are essential for multi-protein complex assembly (Figure 13).

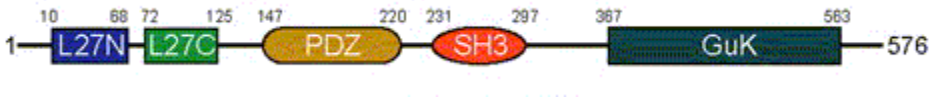


Figure 13 – Domain structure of MMP7 (Stucke, Timmerman et al. 2007)

MPP7 forms a tripartite complex with Discs Large 1(DLG1) and Lin7 at the epithelial adherans junctions (Bohl, Brimer et al. 2007). Absence of MPP7 produces profound defects in the ability of cells to form and maintain tight junctions, and therefore retain cell polarity (Stucke, Timmerman et al. 2007). The gene defective in MODY1, *HNF4a*, also plays a role in the maintenance of tight

junctions (Chiba, Gotoh et al. 2003). MPP7 is expressed at highest levels in the pancreas, although its distribution in islet versus exocrine cells is not known (Novartis Gene Expression Atlas) (Su, Wiltshire et al. 2004).

mTERF is essential for mitochondrial DNA (mtDNA) transcription and replication. It regulates how much read through transcription of the ribosomal DNA continues into the major coding strand of mtDNA. Recently, it also has been implicated in controlling the rate of replication pausing in the mitochondria (Hyvarinen, Pohjoismaki et al. 2007). Mutations in the mtDNA sequence that bind mTERF lead to MELAS (Mitochondrial myopathy, Encephalopathy, Lactic Acidosis, and Stroke-like episodes) Syndrome through defects in protein synthesis and respiratory activity (Chomyn, Martinuzzi et al. 1992). It is located 464 Kb bases from the breakpoint, on chromosome 7.

Frizzled-1 (FZD1) is a well-studied 7-transmembrane domain protein that is the receptor for Wingless-Type (Wnt) and is located on chromosome 7, 148 kb away from the breakpoint. Binding of Wnt by FZD1 leads to the activation of Wnt target genes through the beta-catenin canonical signaling pathway. Physiologically, this pathway plays a key role in the homeostasis of renewal and differentiation in hematopoetic stem cells (Reya, Duncan et al. 2003). Pathologically, derangement of the expression and function of *Fzd1* has been implicated in many forms of cancer (Katoh 2007).

The gene closest to the breakpoint, *WW domain-containing Adaptor with a Coiled-Coil Region (WAC)*, is 29 Kb away on chromosome 10. WW domains contain two conserved tryptophan residues flanking a 20-22 amino acid spacer, and mediate protein interactions by binding proline-rich motifs. It is thought that WAC, through its WW domain, promotes cross-intron bridging of 5' and 3' splicing sites during RNA processing (Xu and Arnaout 2002).

We examined RNA from the proband's transformed lymphoblasts to find heterozygous SNPs in the four genes: *MPP7*, *mTERF*, *FZD1*, and *WAC*. We were unable to find an informative SNP in *FZD1*, due to its lack of introns, but the SNPs in *MPP7*, *mTERF*, and *WAC* all retained their heterozygosity. This suggests that, at least in lymphoblasts, these breakpoints do not exert obvious effects on the transcription of these genes. It is possible that there are tissue-specific regulatory elements that are disrupted by these breakpoints that would be evident in transcriptional analysis of those tissues. Also, aberrant transcripts of these genes may be created as a result of the translocation that would not be detected by this technique.

MPP7, like HNF4a, has been implicated in cellular polarization. When *HNF4a* is overexpressed in embryonal carcinoma cells it causes the formation of tight junctions in a dose-dependent manner (Spath and Weiss 1998; Chiba, Gotoh et al. 2003). Tight junction-associated proteins are upregulated in islets during maturation and are present in mature beta cells. Furthermore, glucose upregulates

the formation of tight junctions in a dose-dependent manner in cultured rat islet cells (Semino, de Gagliardino et al. 1987). It has been suggested that tight junction formation in islets is necessary for proper insulin secretion (Orci, Unger et al. 1973; Orci, Malaisse-Lagae et al. 1975; Collares-Buzato, Carvalho et al. 2004) by separating the high concentrations of glucagon, insulin, and somatostatin in the apical surface from their receptors on the basal surface, thus preventing pathological inhibition of secretion through autoregulation (Unger, Dobbs et al. 1978; Charollais, Gjinovci et al. 2000; Calabrese, Caton et al. 2004). This tight junction pathology may also underlie the progression to islet-cell failure found in advanced Type 2 DM (Prentki and Nolan 2006). We hypothesize that MPP7 could be playing a similar role in maintaining functional tight junctions between the islet cells. Disruption in MMP7 may result in insulin secretion deficiency through defective autoregulation.

Chapter 4

Identification of *Basonuclin 2* as a Candidate Gene for Partial Anomalous Pulmonary Venous Return and Hypospadias

Introduction

The heart is the first organ to form in the embryo, and begins to pump while the embryo undergoes major developmental milestones. It is one of the most sensitive organs to malformation, and about 10% of early miscarriages have severe congenital heart disease. (CHD) (Srivastava 2006). The incidence of clinically-relevant CHD in live births is 1%, and results in 6,000 deaths in the United States every. This makes it the number one cause of birth defect-related deaths (Sander, Klinkner et al. 2006). CHD is a multifactorial disease and environmental factors also contribute to CHD, including maternal rubella, pregestational diabetes, and exposure to certain teratogens (Jenkins, Correa et al. 2007).

The contribution of genetics to CHD is an area of intense interest as more genes are found to be associated with a variety of cardiac defects. A number of molecularly characterized syndromes have associated cardiac malformations, including DiGeorge syndrome, Williams-Beuren syndrome, Alagille syndrome,

Noonan syndrome, and Holt-Oram syndrome (Pierpont, Basson et al. 2007). Recently, non-syndromic forms of CHD have also yielded disease-causing genes, including *NKX2-5* in atrial septal defects and *TFAP2B* in patent ductus arteriosus (Satoda, Zhao et al. 2000; Elliott, Kirk et al. 2003).

Partial Anomalous Pulmonary Venous Return

Partial anomalous pulmonary venous return (PAPVR) is a developmental malformation that leads to one or more of the right-sided pulmonary veins draining to the superior or inferior vena cava, azygous vein, the right atrium, or the coronary sinus, or the left-sided pulmonary veins draining to the brachiocephalic vein, the coronary sinus, the right atrium, the inferior caval vein, ductus venosus, or the portal system (DeLuca 1984). When the right pulmonary veins drain into the inferior vena cava, the chest radiograph image resembles a scimitar, or curved sword, and is thus called either scimitar syndrome or hypogenetic left heart complex (Wang, Wu et al. 2008). In most cases, only one lung is involved in the aberrant connections; however, both lungs may form a single pulmonary vein that drains to the right atrium, coronary sinus, or a systemic vein. These aberrant connections allow some blood that normally travels to oxygenate peripheral tissues to recirculate back to the lungs. To accommodate this abnormal blood flow, a sinus venosus septal defect, a hole in the upper atrial septum that is contiguous with the superior vena cava, is often found in the

patients. In rare cases, other forms of atrial septal defects and patent foramen ovale are also seen (Freedom; Nichols 2006). Diagnosis of PAPVR can be made by echocardiography or cardiac catheterization; it is found in about 0.4% of live births as an isolated finding (Healey 1952).

PAPVR can be found in conjunction with other congenital heart defects; the most common is scimitar syndrome. As mentioned above, in scimitar syndrome the pulmonary veins connect to the inferior vena cava either above or below the diaphragm. Usually the left pulmonary veins are anomalous, and the affected lung is hypoplastic and malformed. There are also aortic-pulmonary collateral arteries to the abnormal lung, and if these arteries are significant then pulmonary hypertension can develop (Wang, Wu et al. 2008). Other defects associated with PAPVR include left isomerism, a form of *situs inversus* where patients demonstrate bilateral left-sidedness, and divided left atrium.

When only one or two veins drain anomalously, there is usually no detrimental effect on the developing fetus. Many cases are never diagnosed, as about 0.7% of routine autopsies report PAPVR as an incidental finding (Weiman, Lee et al. 1985). Much less commonly (about 1 in 10,000), surgical repair of the abnormal veins is indicated if the right ventricle of the heart becomes pathologically stressed by the additional blood volume (Rudolph 2001). Mortality for surgical repair of uncomplicated PAPVR is virtually zero. Rare reported

complications include pulmonary venous obstruction and residual left-to-right blood shunting. When symptoms do present, they can include recurrent pulmonary infection, fatigue, chest pain, difficulty breathing, and, in severe cases, congestive heart failure and pulmonary hypertension (Weiman, Lee et al. 1985).

In the fetus, the splanchnic plexus is the common drainage site for the developing lungs and other foregut derivatives, through the cardinal, umbilical, and vitelline veins (Neill 1956). There is nothing connecting the splanchnic plexus and the heart until the common pulmonary vein forms from the left atrium and contacts the pulmonary section of the plexus. The four major pulmonary veins from the plexus drain into the common pulmonary vein and into the left atrium. Later in development this common pulmonary vein becomes part of the left atrium, and the four major pulmonary veins drain directly into the atrium (Edwards 1953). This results in the normal cardiopulmonary anatomy seen at birth, with the blood from the lungs draining through the four pulmonary veins into the left atrium. The fetal connections to the cardinal and umbilical veins are no longer present. If key steps in this development are disrupted, than either PAPVR or TAPVR can result. Specifically, if the common pulmonary vein fails to form, the fetal connections from the pulmonary plexus and cardinal or umbilical system continue to be maintained to ensure adequate blood flow (Neill 1956).

Genetics of APVR

Although the majority of APVR cases appear to be sporadic, there have been well-described familial cases of TAPVR (Solymar, Sabel et al. 1987; Raisher, Dowton et al. 1991). In one large family of 14 affected members over three generations, the abnormal venous drainage of the affected patients occurred had many different anatomical sites. Five connected to a vertical vein, two to an infradiaphragmatic vein, and one each to the right atrium, right superior vena cava, and multiple connections (Bleyl, Ruttenberg et al. 1994). Additionally, other chromosomal rearrangements have been found in patients with TAPVR, including an interstitial deletion in chromosome 2, a terminal deletion of chromosome 11, and a mosaic ring spanning chromosome 12p (Ramer, Mowrey et al. 1990; Wu, Hwu et al. 2001; Harris, Siu et al. 2004).

In two unrelated families the disease has been mapped to a region on 4q12 using linkage analysis (Bleyl, Nelson et al. 1995; Bleyl, Botto et al. 2006). A gene in this region, *Kinase Insert Domain Receptor (KDR),* has been implicated as being defective in this family since it is a binding partner of vascular endothelial growth factor and is expressed in endothelial cell precursors early in development. However, no mutations have been found in affected patients. The mouse homolog, *fetal liver kinase-1*, is involved in the development of blood and blood vessels (Shalaby, Ho et al. 1997). Another candidate gene was recently

identified from studies of a patient with TAPVR and a balanced 10;21 translocation. The chromosome 10 breakpoint was 130 kb from *ANKRD1/CARP*, which was overexpressed in lymphocytes carrying the translocation (Cinquetti, Badi et al. 2008). *ANKRD1/CARP* had been previously identified as a heart-specific transcriptional repressor, specifically of atrial natriuretic factor, one of the master regulators of gene expression in the heart. *ANKRD1/CARP* was also over-expressed in some sporadic TAPVR patients (Cinquetti, Badi et al. 2008).

Hypospadias

Hypospadias, when the urethral orifice is located on the frontal side of the penis, is found in 0.7% live births (Baskin 2000; Gallentine, Morey et al. 2001). In the United States there has been an unexplained doubling of the incidence of both severe and mild hypospadias since 1970 (Paulozzi, Erickson et al. 1997) (Figure 13).

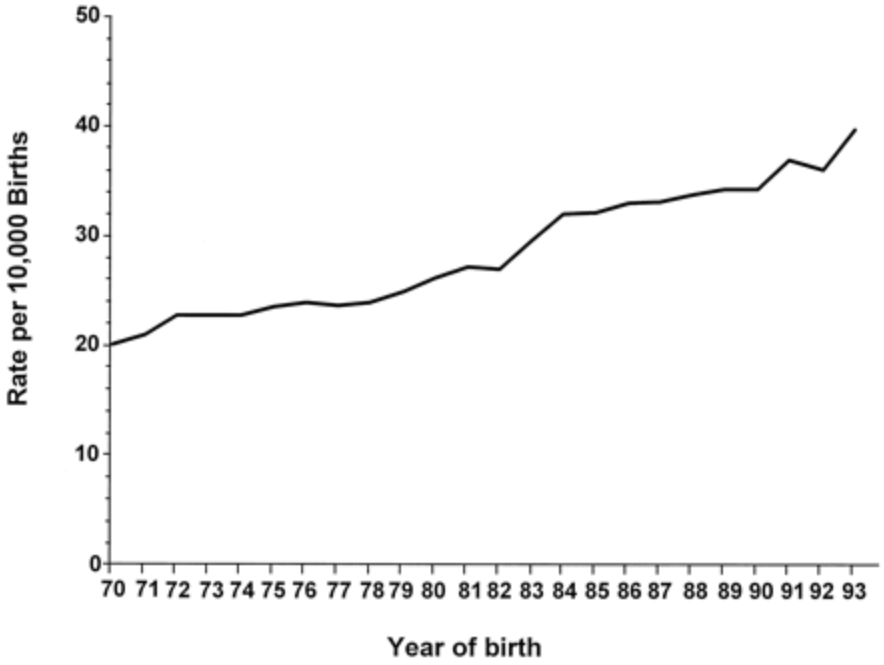


Figure 13 - Incidence of Hypospadias in the United States has doubled between 1970 and 1990 (Paulozzi, Erickson et al. 1997)

Hypospadias can be diagnosed in the fetus by ultrasound, but the majority of cases are diagnosed during the newborn physical exam. The majority of cases are surgically repaired within the first year of life on an outpatient basis. Success of these surgeries approaches 90%; reported complications include urethrocutanous fistula, diverticulum, and stricture (Duckett 1995).

The development of the penis begins by the end of the first month, when the hindgut and developing urogenital system travel to the cloacal membrane. The urorectal septum divides the cloacal membrane into the posterior anal membrane and anterior urogenital membrane. Around the urogenital membrane three protuberances form: the genital tubercle and two genital swellings. At this point in

development the surge of testosterone in male fetuses causes the distance between the anus and genital structures to increase, the phallus to elongate, the penile urethra to develop, and the prepuce to form (Jirasek, Raboch et al. 1968).

At the beginning of the third month of gestation the urethral groove is flanked on either side by the urethral fold on the ventral surface of the phallus. The medial edges of the endodermal urethral folds fuse to form the median raphe and begin the formation of the penile urethra. The foreskin begins to develop at this stage from two low preputial folds on either side of the penis. These folds fuse on the dorsal side of the penis and form a flat sheet at the proximal coronal edge. The sheet continues to grow by mesenchymal expansion until it reaches the entire glans. By the third month the glans is separated from the penile shaft by the coronal sulcus and the urethral folds and fused completely. The glandular urethra appears after 16 weeks, which is hypothesized to either be the result of endodermal cellular differentiation or intrusion of ectodermal tissue from the glans (Baskin 2000).

Previous studies have suggested that the majority of cases of hypospadias are sporadic, and that both environmental and genetic causes contribute (Fredell, Iselius et al. 2002). There have been multiple families described that appear to have an autosomal dominant form of isolated hypospadias (Page 1979). Mutations in the 5-alpha reductase gene and androgen receptor gene have been found in patients with severe hypospadias, usually in conjunction with other

genital malformations (Allera, Herbst et al. 1995; Sutherland, Wiener et al. 1996; Silver and Russell 1999).

Additionally, a rare X-linked form of the disorder has been mapped to the *CXORF6* gene. In a study of 166 cases of hypospadias, three carriers of nonsense mutations in *CXORF6* were described (Fukami, Wada et al. 2006). *CXORF6* is highly expressed in Sertoli and Leydig cells of the testes, and its expression rises sharply in the second trimester of gestation (O'Shaughnessy, Baker et al. 2007). Recent studies of the function of *CXORF6* have illustrated its role in penile development, and lead to renaming the gene "*mastermind-like domain containing 1*" (*MAMLD1*). Research has shown that steroidogenic factor 1 binds to an enhancer of MAMLD1 and transactivates it, which in turn transactivates *Hairy and enhancer of split 3 (Hes3),* which augments testosterone activity. When MAMLD1 was removed from Leydig tumor cells they were unable to produce normal levels of testosterone, which suggests a pathway for its involvement in hypospadias formation (Fukami, Wada et al. 2007).

Results

The proband was a 33-year-old man with a past medical history notable for mild unrepaired coronal hypospadias and PAVR that required surgical repair

in the third decade of life. Additionally, he had mild lower limb length inequality that only required treatment with a shoe lift, but no other significant medical history. The proband was found to carry a balanced 9;13 translocation (46,XY,t(9;13)(p22.3;q22.1) after karyotyping of products of conception from his wife's second spontaneous abortion revealed an unbalanced 9;13 translocation (46,XX,der(9)t(9;13)(p22;q22).The balanced translocation occurred *de novo* and was not inherited by the proband's 9-year-old daughter, who had a normal cardiac ultrasound after an episode of exercise-induced perioral cyanosis (Figure 15).

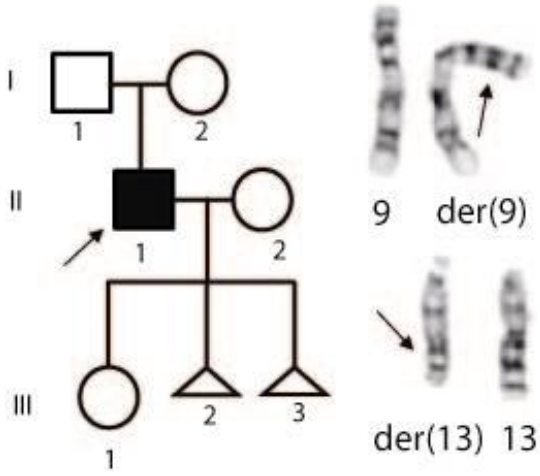
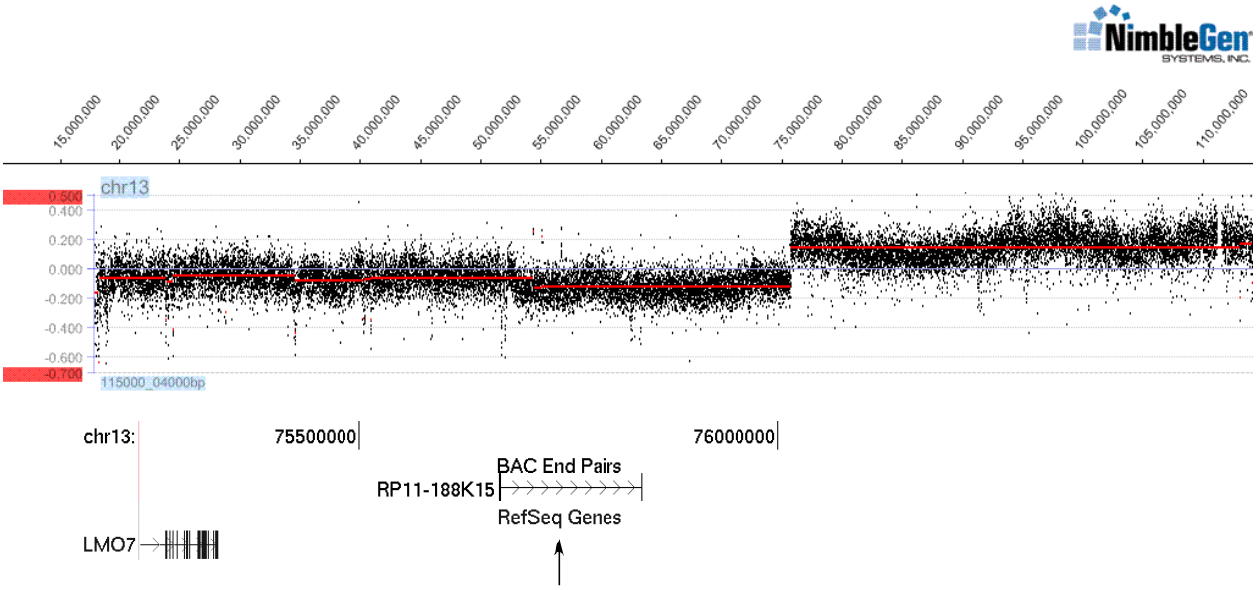


Figure 15 - A. Pedigree. The proband, II-1 carries the *de novo* balanced 9;13 translocation and the spontaneous abortion, III-3, carried an unbalanced segregant. Karyotypes on I-1, I-2, and III-1 were normal. III-2 was not karyotyped. B. Partial G-banded karyotype showing normal and derivative chromosomes 9 and 13.

We hypothesized that the molecular chromosomal breakpoints were the same in the proband with the balanced translocation and the products of conception with the unbalanced translocation. We performed arrayCGH on DNA from the products of conception using chromosome 9 and 13 tiling oligonucleotide arrays. The aCGH confirmed that the POC was trisomic for 13q22-qter (Figure 16a) and monosomic for 9pter-p22. From the CGH data, the breakpoints were mapped to a 3 kb interval of chromosome 9 and a 4 kb interval of chromosome 13 (chr9:16,498,500-16,501,500, chr13:75,738,000-75,742,000, UCSC Genome Browser, March 2006 assembly) (Figures 16a and 16b).

A


NimbleGen
SYSTEMS, INC.
15,000,000
20,000,000
25,000,000
30,000,000
35,000,000
40,000,000
45,000,000
50,000,000
55,000,000
60,000,000
65,000,000
70,000,000
75,000,000
80,000,000
85,000,000
90,000,000
95,000,000
100,000,000
105,000,000
110,000,000
chr13
0.500
0.400
0.200
0.000
-0.200
-0.400
-0.600
-0.700
115000_04000bp
chr13:
75500000
76000000
BAC End Pairs
RP11-188K15
RefSeq Genes
LMO7

B

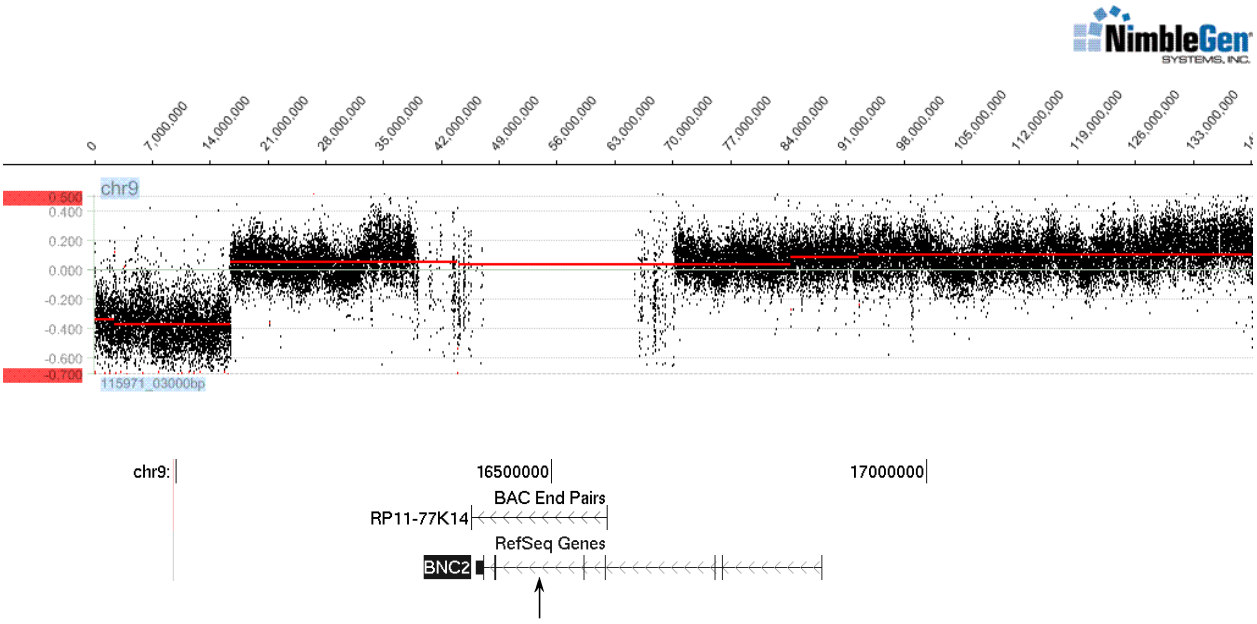

NimbleGen
SYSTEMS, INC.
0
7,000,000
14,000,000
21,000,000
28,000,000
35,000,000
42,000,000
49,000,000
56,000,000
63,000,000
70,000,000
77,000,000
84,000,000
91,000,000
98,000,000
105,000,000
112,000,000
119,000,000
126,000,000
133,000,000
chr9
0.500
0.400
0.200
0.000
-0.200
-0.400
-0.600
-0.700
115971_03000bp
chr9:
16500000
17000000
BAC End Pairs
RP11-77K14
RefSeq Genes
BNC2

Figure 16. Translocation breakpoints for chromosome 13 (A) and chromosome 9 (B). Top, tiling array CGH showing partial 13q trisomy (A) and partial 9p monosomy (B). Bottom, UCSC Genome Browser tracks showing RefSeq genes within 500 kb of each translocation breakpoint.

BAC clones containing the breakpoint regions on chromosome 9 (RP11-77K14) and chromosome 13 (RP11-188K15) were used as FISH probes to confirm the breakpoints in the proband. Both probes consistently showed two signals in the expected location when hybridized to control metaphases from a normal individual. By contrast, each probe consistently showed three signals, one with normal intensity and two with reduced intensity, when hybridized to the proband's metaphases (Figures 17a and 17b).

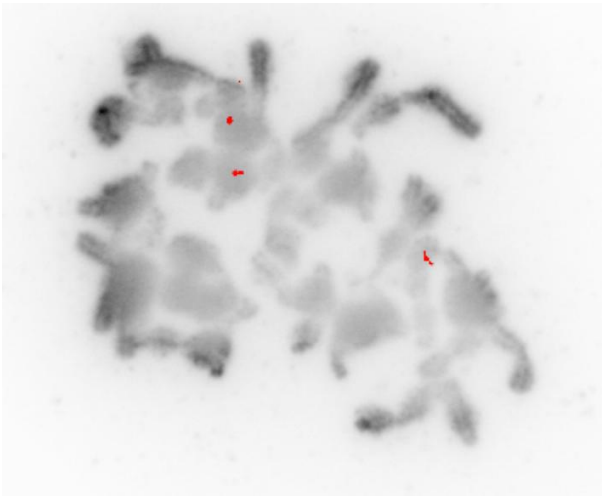


Figure 17a - FISH to proband metaphases showing three signals with BAC RP11-77k14 on 9p22.3, confirming that the breakpoints of the derivative chromosomes determined by CGH are the same for the balanced translocation and ruling out sizable deletion at either breakpoint.

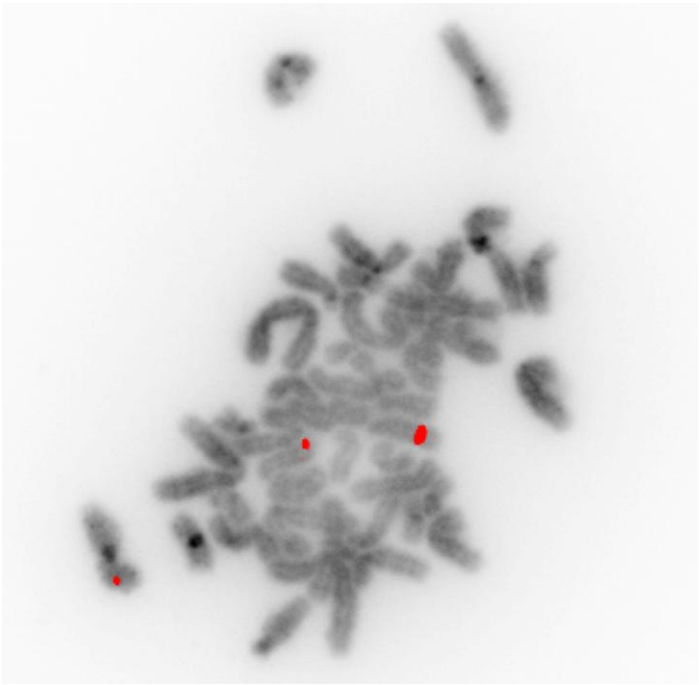
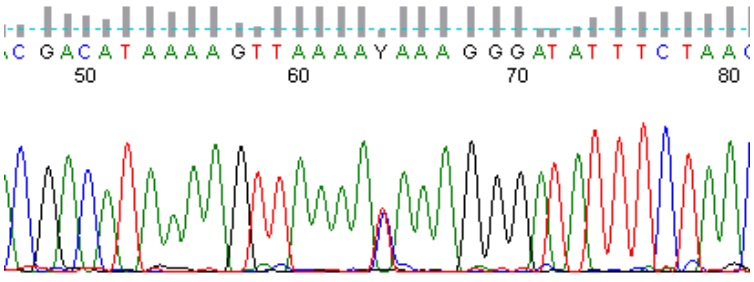


Figure 17b - FISH to proband metaphases showing three signals with RP11-188K15 on 13q22.2, confirming that the breakpoints of the derivative chromosomes determined by CGH are the same for the balanced translocation and ruling out sizable deletion at either breakpoints.

The breakpoint on chromosome 13 did not disrupt any known transcription unit. It was ~400 kb 3' of the nearest gene, *LIM-domain only 7 (LMO7)*. The breakpoint on chromosome 9 fell between exons 4a and 5 of *Basonuclin 2 (BNC2)*, a zinc finger protein.

To determine in the proband if both the translocated and untranslocated chromosomes were expressing *LMO7* and *BNC2*, we identified a heterozygous SNP in the 5' UTR of each gene in lymphoblast DNA. By sequencing the lymphoblast cDNA for those SNPs we were able to show that both genes maintained the heterozygosity of the SNPs, therefore mRNA from both chromosomes is being transcribed (Figure 18). This is expected for *LMO7* because of its distance from the breakpoint, and we hypothesize that because of *BNC2*'s transcriptional complexity that a breakpoint farther downstream gene causes little disruption at the 5' UTR.



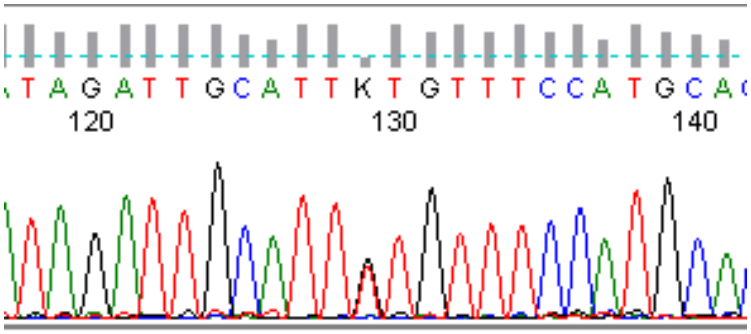


Figure 18 – Sequence traces of cDNA from proband *BNC2* (top) and *LMO7* (bottom) that retain a heterozygous SNP

Additionally the levels of mRNA of *BNC2* and *LMO7* in transformed lymphoblasts of the proband were compared with normal controls. *LMO7* seemed to be within the range demonstrated by the three normal controls; however *BNC2* wasn't expressed in the lymphoblasts of any of the subjects in high enough levels to detect (Figure 19). From these studies we doubted the contribution of a disruption in *LMO7* in contributing to the proband's phenotype

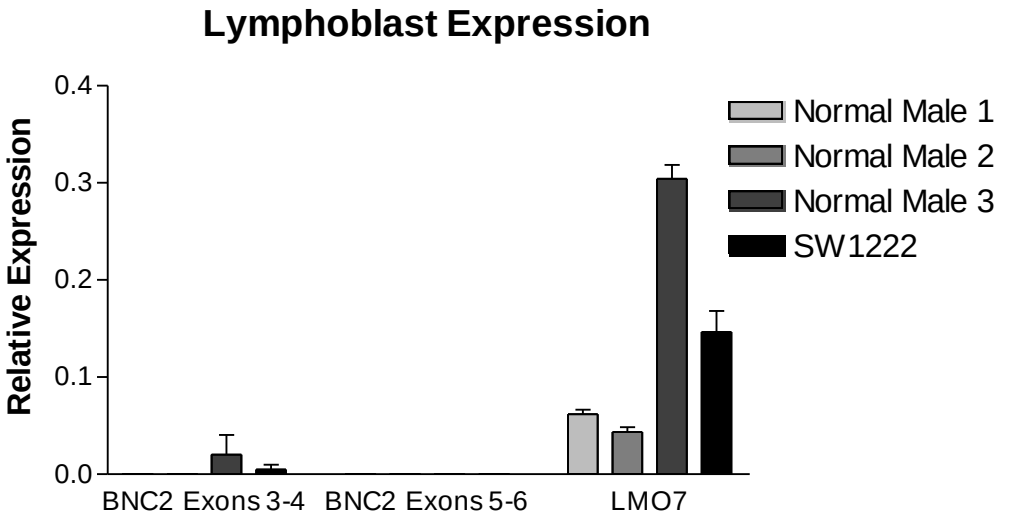
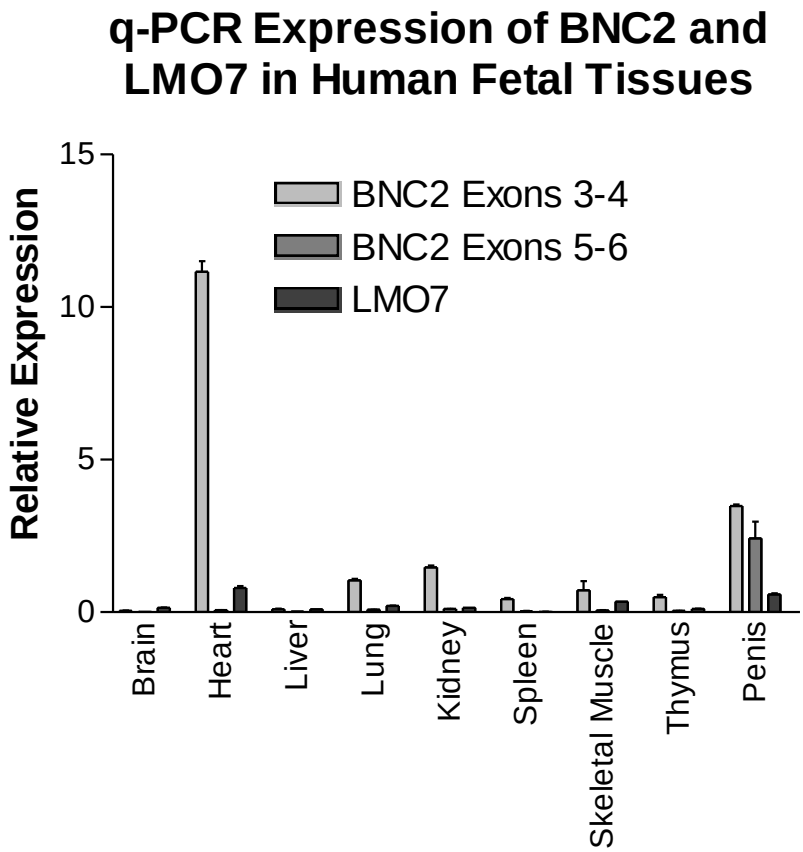
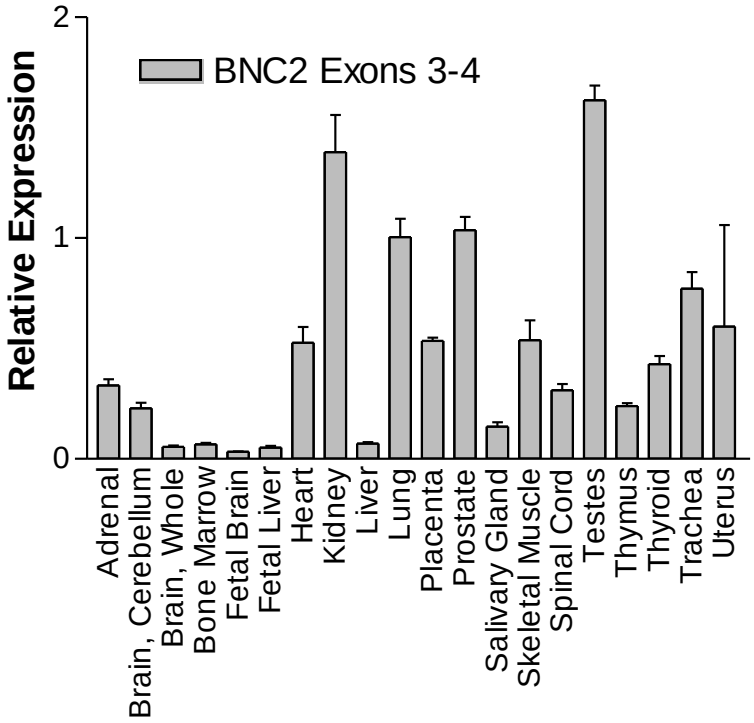


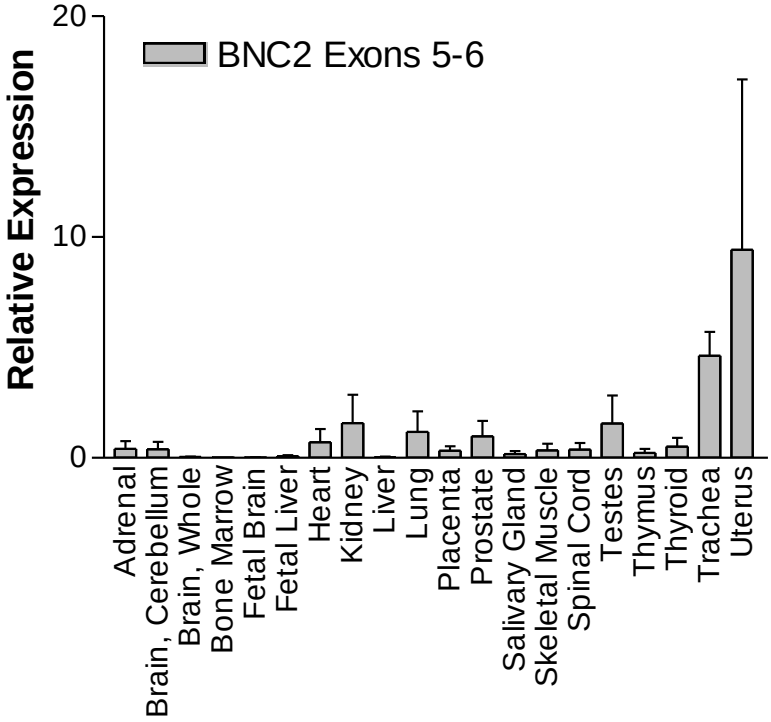
Figure 19 – quantitative RT-PCR expression data for *LMO7* and *BNC2* in transformed lymphoblasts from the proband and three normal male controls

The tissue distribution of *LMO7* and two exons of *BNC2* was assessed by tissue specific quantitative RT-PCR using RNA derived from an adult multiple tissue panel consisting of human adrenal gland, cerebellum, whole brain, bone marrow, fetal brain, fetal liver, heart, kidney, liver, lung, placenta, prostate, salivary gland, skeletal muscle, spinal cord, testis, thymus, trachea and uterus and a fetal tissues panel consisting of brain, heart, liver, lung, kidney, spleen, skeletal muscle, thymus and penis (Figure 20). For *LMO7* the highest levels in adult tissues were observed in the heart, lung, and skeletal muscle. *BNC2* was especially prevalent in the uterus, with expression also observed in the kidney, lung, prostate, testes and trachea.

We were especially interested in the fetal expression patterns. Isoforms of *BNC2* containing exons 3 and 4 were highly expressed in both the fetal heart and the fetal penis; isoforms containing exons 4 and 5 were also high in fetal penis. *LMO7* also showed the highest levels in the fetal heart and fetal penis. *Cyclophilin A* was used as an internal control and showed a similar C_t in all samples.


q-PCR Expression of BNC2 and LMO7 in Human Fetal Tissues
Relative Expression
15
10
5
0
BNC2 Exons 3-4
BNC2 Exons 5-6
LMO7
Brain
Heart
Liver
Lung
Kidney
Spleen
Skeletal Muscle
Thymus
Penis


Relative Expression
2
1
0
BNC2 Exons 3-4
Adrenal
Brain, Cerebellum
Brain, Whole
Bone Marrow
Fetal Brain
Fetal Liver
Heart
Kidney
Liver
Lung
Placenta
Prostate
Salivary Gland
Skeletal Muscle
Spinal Cord
Testes
Thymus
Thyroid
Trachea
Uterus


Relative Expression
20
10
0
BNC2 Exons 5-6
Adrenal
Brain, Cerebellum
Brain, Whole
Bone Marrow
Fetal Brain
Fetal Liver
Heart
Kidney
Liver
Lung
Placenta
Prostate
Salivary Gland
Skeletal Muscle
Spinal Cord
Testes
Thymus
Thyroid
Trachea
Uterus

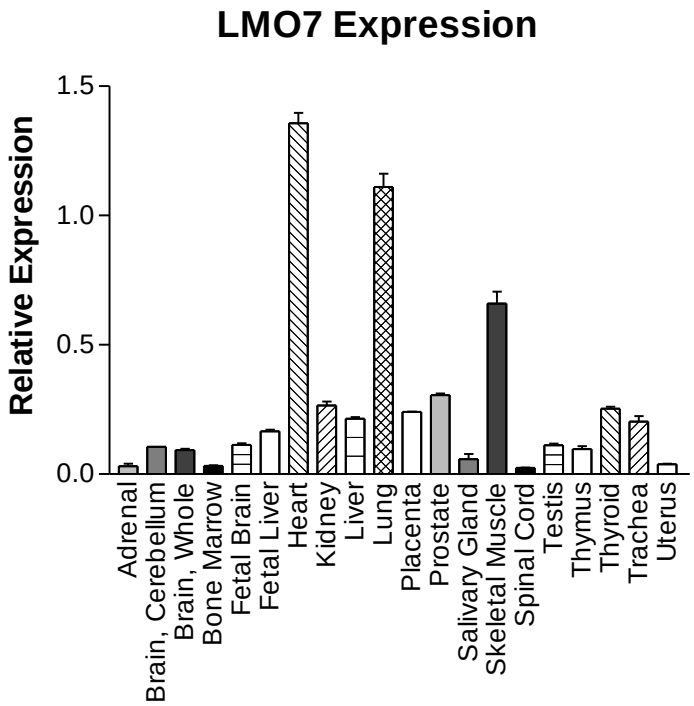


Figure 20 - Quantitative RT-PCR in normal human adult and fetal tissues panels consisting of human adrenal gland, cerebellum, whole brain, bone marrow, fetal brain, fetal liver, heart, kidney, liver, lung, placenta, prostate, salivary gland, skeletal muscle, spinal cord, testis, thymus, trachea and uterus and a fetal tissues panel consisting of brain, heart, liver, lung, kidney, spleen, skeletal muscle, thymus and penis.

In addition to q-PCR expression information, our collaborators, M. Maitra and V. Garg, performed ISH for *BNC2* on E13.5 mouse embryos. They found strong expression in the genital tubercle and in the cartilage of the developing limb buds consistent with the phenotype found in the proband. Interestingly, strong expression of *BNC2* mRNA was not seen in the heart but it was expressed in regions of developing lung, from which the pulmonary veins originate (Figure 21).

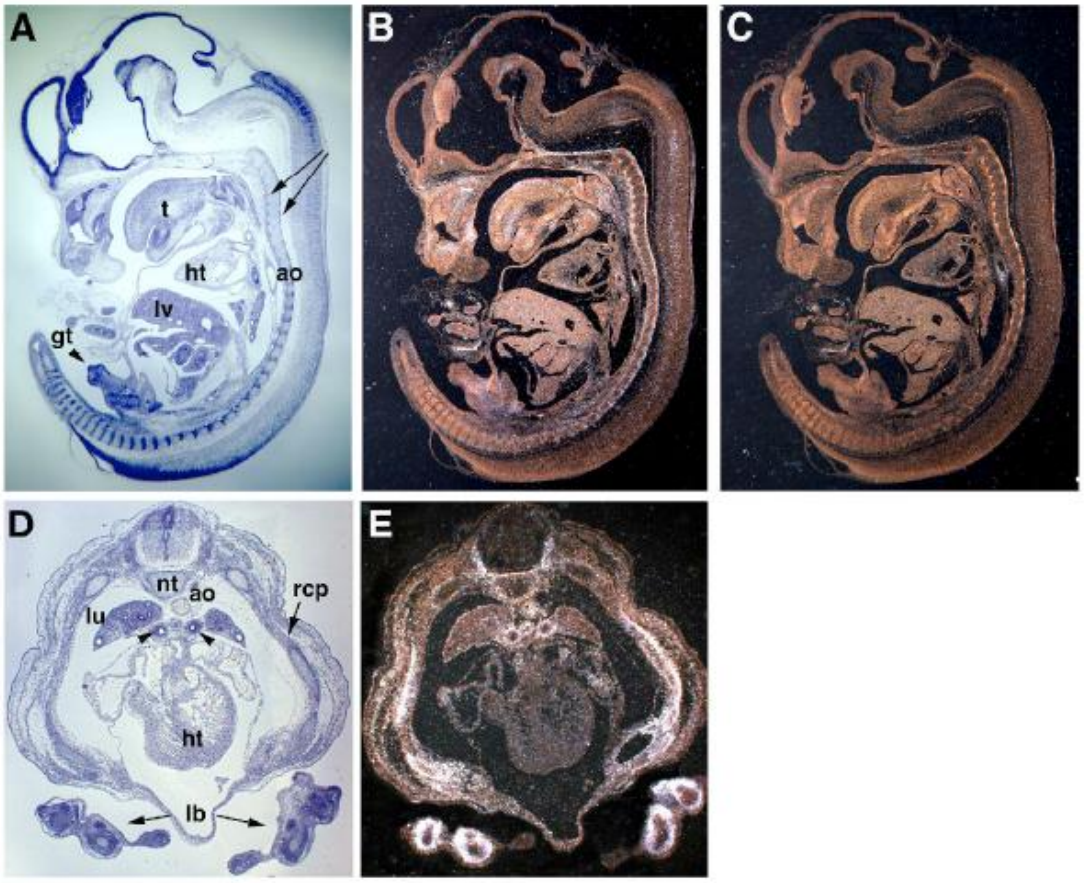


Figure 21 - Expression of *BNC2* mRNA in murine embryogenesis by radioactive section in situ hybridization. Sagittal (A, B, C) and transverse (D, E) sections of E13.5 mouse embryos. (B) Highest levels of expression of *BNC2* mRNA are seen in the tongue (t), genital tubercle (gt), structures within the developing nose, and tissue surrounding the cartilage primordium of the vertebrae (arrows). (A) represents bright-field image of (B). (E) Transverse section through the thorax show strongest expression of *BNC2* mRNA in the cartilage primordium of the ribs (rcp) and limb buds (lb). Expression is also seen in the bronchi of the lung (arrowheads) and epithelium surrounding the lung (lu). (C) represents bright field image of (D). (C) *BNC2* sense probe demonstrates no signal. ht, heart; ao, aorta; lv, liver; nt, notochord. (M. Maitra and V.Garg)

We sequenced the 7 major exons (1, 2, 2a, 3, 4, 5, 6) of *BNC2* (Vanhoutteghem and Djian 2007) in 10 unrelated patients with surgically-repaired Total- or Partial- APVR (Table 6). None of the six males in this sample had hypospadias.

Sample ID	Gender	Race	Diagnosis
RCHD-1	F	Hispanic	TAVPR
RCHD-21	F	Caucasian	TAPVR
ICHD-127	M	African-American	TAPVR
ICHD-151	F	Caucasian	PAPVR, ASD, Sinus venosus
2CHD-0017	M	Caucasian	PAPVR
2CHD-0033	M	Caucasian	PAPVR, PH, Hearing loss
2CHD-0040	F	Hispanic	PAPVR, ASD, Intracranial abscess
2CHD-0077	M	Hispanic	TAPVR
2CHD-0090	F	Caucasian/African-American	TAPVR
2CHD-0130	M	Hispanic	Abnormal pulmonary venous confluence, VSD, ASD

Table 6 – Patients screened for mutations in *BNC2*

Four different coding variations were observed in our patients (Table 7). Two of these variants were previously annotated: rs3739715, found in one patient, and rs3739714, found in two patients. A novel synonymous substitution at E2+81A→T was found in two patients, ICHD-151 and RCHD-1; this variant is not in a conserved region nor in a splice site. One novel non-synonymous SNP in patient 2CHD-0040, who was diagnosed with PAPVR and an ASD, was found at E6+94A→C. This SNP changes an aspartic acid that is evolutionarily conserved among *BNC2* orthologs to glutamic acid; however, this residue is not conserved in the paralog, *BNC1* (Figure 22). Using the Sorting Intolerance From Tolerance (SIFT) program this substitution has a score of 0.15, on a scale of 0-1 (Ng and Henikoff 2003). Scores with probabilities <0.05 are predicted to be deleterious. Additionally, this SNP is not within any known functional domain of *BNC2* (Vanhoutteghem and Djian 2004).

Position (Exon)	**Sequence Alteration**	**No of Patients**	**Syn/Non-Syn**	**Annotation**
chr9:16728403 (2)	A->A/T	2	Synonymous	
chr9:16425848 (6)	A->A/G	2	Non-synonymous	rs3739714
chr9:16425714 (6)	G ->A/G	1	Synonymous	rs3739715
chr9:16409553 (6)	C->C/A	1	Non-synonymous	

Table 7 - Sequencing results of *BNC2* from 10 patients with PAPVR

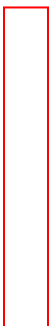


Figure 22 - Conservation of novel non-synonymous SNP in *BNC2*

Discussion

We characterized a de novo balanced 9;13 translocation in a man with unusually severe PAPVR, and mild hypospadias. The translocation falls within the critical region (9p22.2-p23) for monosomy 9p syndrome, which results in cardiac malformations, urogenital defects, mental retardation, trigonocephaly, and a distinctive facies (Kawara, Yamamoto et al. 2006). The 9p breakpoint was located in an intron of *BNC2*, a large (~470 kb) gene of unknown function with a highly complex pattern of alternative transcriptional processing. *BNC2* was discovered in 2004 through its similarity to *Basonuclin 1 (BNC1)*, a protein thought to be involved in RNA processing and cell proliferation (Romano, Li et al. 2004; Vanhoutteghem and Djian 2004). The two genes share 40% similarity.

Both genes have three separated pairs of zinc fingers, a nuclear localization signal (NLS), and a serine-rich region. *BNC1* is less conserved, and thought to have evolved from a distant duplication of *BNC2*. The most conserved isoform of human *BNC2* shares 97.2% identity at the amino acid level with mouse *BNC2*, which suggests that the gene has an essential function. Some forms of *BNC2* have been implicated in mRNA processing based on its co-localization with the processing machinery, but its true function remain unknown (Vanhoutteghem and Djian 2004). BNC2 is expressed ubiquitously, with most forms that contain the NLS found in the nucleus and those without spread between the nucleus and cytoplasm (Vanhoutteghem and Djian 2006).

It was recently published that the formation of over 90,000 mRNAs and over 2,000 different proteins of *BNC2* is possible. Much of this complexity results from alternate splicing of the 23 known exons, all of which are alternate (Figure 23). The exons are named historically, with the major exons (1 to 6) being the first discovered, as more exons were found they have been named in relation to the original six. Therefore it is possible to create a transcript that has both exons 2, 2a, and 2b, or one with no exon 2 version at all.

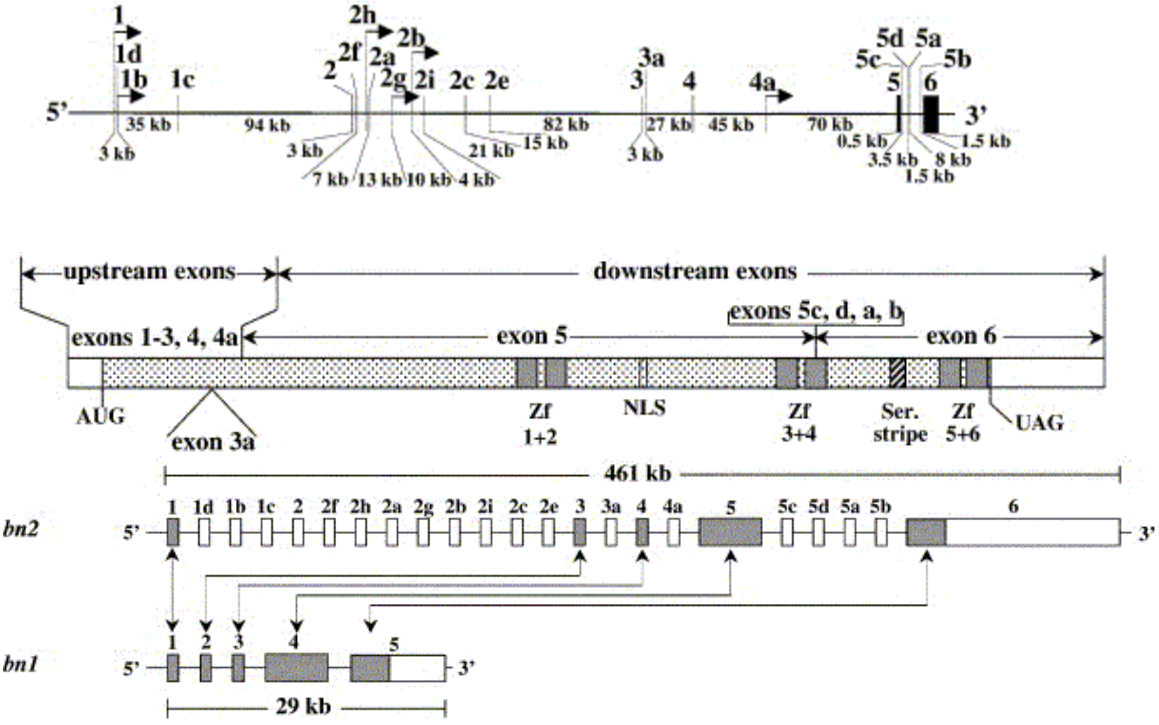


Figure 23 – Exon/Intron structure of *BNC2* shows great complexity and allows for the formation of over 90,000 mRNA and over 2,000 protein products (Vanhoutteghem and Djian 2007)

There are forms which lack any of the known domains, including any of the three zinc finger domains. There are also one major and five minor promoters, and three poly-A addition sites; this complex pattern has yet to be explored in terms of tissue specificity or other forms of regulation (Vanhoutteghem and Djian 2007). *BNC2* has never previously been implicated in human disease, although complete lack of *BNC1* prevents development of embryos beyond the two-cell stage (Ma, Zeng et al. 2006).

In previous work with *BNC2* it has been described through Northern blot analysis that in the mouse expression was highest in the kidney, ovary, skin, testis, and embryo (Romano, Li et al. 2004). Additionally, a Northern blot on human tissues has shown highest expression of *BNC2* in testis, uterus, and intestine (Vanhoutteghem and Djian 2004).

The 13q breakpoint fell in an intergenic region, ~400 kb 3' to the transcriptional unit of *LMO7*, a transcription factor. *LMO7* contains PDZ and LIM protein-protein interaction domains, and is a putative transcriptional regulator. Interestingly, targeted mutations of two other LIM/PDZ proteins, *actin-associated LIM protein (ALP)* and *ZASP/Cypher/Oracle*, have been found to cause congenital heart and muscle pathology (Guy, Kenny et al. 1999; Nakagawa, Hoshijima et al. 2000).

LMO7 is found in the nucleus, cytoplasm and cell surface, with especially high concentrations at adherans junctions.(Ooshio, Irie et al. 2004) It has also been shown to bind afadin and the E-cadherin-catenin system through alpha-actinin, further suggesting a role in cell-cell adhesion (Ooshio, Irie et al. 2004). A mouse with an engineered 800 kilobase deletion that encompassed *LMO7* and *ubiquitin C-terminal hydrolase L3 (UCHL3)* demonstrated defects in viability, post-natal growth, and degeneration of muscle and retina. (Semenova, Wang et al. 2003) The *LMO7/UCHL3*-null mice were born at expected Mendelian ratios, but only about two-thirds reach adulthood and these are significantly runted. The

cardiac and skeletal muscles of the null mice had reduced numbers of nuclei, and contained inclusion in older mice (Semenova, Wang et al. 2003).

These signs of muscular degeneration may be explained by the discovery that LMO7 is a binding partner for emerin, one of the causative genes in X-linked Emery-Dreifuss muscular dystrophy (X-EDMD) (Holaska, Rais-Bahrami et al. 2006). X-EDMD patients have progressive skeletal muscle weakening, contractures of major tendons, and cardiac conduction system defects (Emery 2000). It has been proposed that LMO7 is a nuclear shuttle that relays information about the status of emerin. Specifically, LMO7 forms a positive feedback loop with emerin to respond to changes in cell adhesions or mechanical force by regulating tissue-specific gene expression (Holaska, Rais-Bahrami et al. 2006).

The expression patterns of *LMO7* in adult and fetal tissues were initially interesting because of the involvement of the heart, fetal heart, lungs, skeletal muscle and fetal penis in the malformations in our patient. In the initial paper describing *LMO7* it was shown through Northern blots to be expressed in the heart, lung, and skeletal muscle, which agrees with our findings (Putilina, Jaworski et al. 1998).

Although BNC2 and LMO7 are both relatively highly expressed in human fetal heart and penis, we considered dysregulation of LMO7 expression unlikely to cause the observed phenotypes. The breakpoint was 3' of the gene,

whereas long-range position effects are more often associated with breakpoints in the 5' regulatory region (Kleinjan and van Heyningen 2005). Also, roughly equal expression of both LMO7 alleles was observed in the proband's lymphocytes. Both BNC2 alleles also showed roughly equal expression in the proband's lymphocytes. However, due to limited availability of informative expressed SNPs, we were unable to assay specifically those BNC2 isoforms containing the last two exons, the ones whose expression is most likely to be affected by the translocation. Additionally, this translocation is predicted to result in a tail-to tail orientation of BNC2 exons 1-5 and LMO7, making it unlikely to create a fusion transcript.

We did not investigate BNC2 as a candidate for hypospadias because mutations in this would likely be just one of many etiologies for such a common birth defect. PAPVR necessitating surgical correction is much rarer than hypospadias. We attempted to corroborate a role for BNC2 in pulmonary venous development by screening rare patients with surgically corrected TAPVR or PAPVR for BNC2 coding mutations. We were able to collect 10 unrelated patients. One of patients with PAPVR had a nonsynonymous substitution of unclear significance. She had sporadic disease, and family members were unavailable for study.

Chapter 5

Conclusions and Future Directions

Patients with balanced chromosomal translocations and disease have previously been used to identify novel disease-causing genes by examining the genes affected by the chromosomal breakpoints. We have identified a man with unexplained MODY-like diabetes mellitus and a balanced 7q22;10p12 translocation, and a man carrying a balanced translocation associated with partial anomalous pulmonary venous return (PAPVR) and hypospadias.

Through the investigation of this unique patient with MODY-X and a balanced translocation, we were able to propose one candidate gene worthy of further study, *MPP7*. Although the breakpoint did not map to a specific disease-causing gene, these findings may contribute to future investigations aimed at the identification of novel MODY genes. There are a number of well-studied mechanisms by which a translocation may cause an effect without an intragenic break, which could be occurring in this case. In a recent study of 47 chromosomal breakpoints in patients with reportedly-balanced translocations and congenital abnormalities, only 34 (72%) directly disrupted genes (Higgins, Alkuraya et al. 2008). It is also possible that the MODY and translocation are unrelated. In the general population one in 2000 patients have *de novo* balanced translocations, and

the prevalence of MODY is around one in 3000, therefore the chance of a translocation and coincidental MODY phenotype is about one in 6,000,000.

Screening of MODY-X patients for *MPP7* mutations may yield other mutations that would more firmly cement the causality of gene disruptions in *MPP7* and MODY. It would even be useful to screen the binding partners of *MPP7*, *DGL1* and *LIN7*, for mutations. (Bohl, Brimer et al. 2007) The greatest yield of new mutations may be found in non-white populations, as the proportion of MODY-X in Japanese patients is as high as 80%, but is only about 10% in Caucasians.

It would be informative to learn more about the specific expression of *MPP7* within the pancreas, and especially if it is affected by changes in blood glucose concentration. If *MPP7* is highly expressed in the connections between the lateral surfaces of the islet cells, that would lend support to our theory about its role in maintaining a glucose gradient between the basal and apical surfaces of the cells.

Additionally, generation of *MPP7* knockout mice would test the hypothesis that abnormal epithelial cell polarity plays a role in the pathogenesis of MODY and perhaps also impaired glucose-stimulated insulin secretion in common type 2 diabetes mellitus. These *MPP7*-deficient mice could be metabolically tested in both heterozygote and homozygote state and provide valuable information about the pathology of aberrant tight junctions. The

histology of the islet-cells under high and low glucose states would be especially interesting to observe.

Additionally, we used a novel CGH-based approach to facilitate mapping of a balanced translocation associated with PAPVR and hypospadias. The translocation disrupted *BNC2*, a large, complex gene of unknown function whose expression is enriched in fetal heart and penis. One missense mutation in *BNC2* of unclear significance was identified in 11 patients with sporadic TAPVR or PAPVR. The possible role of *BNC2* in the etiology of hypospadias and disorders of pulmonary venous return warrants further study, including more extensive screening of mutations in patients with these conditions and creation of animal models.

It has been well-described that CHD is a multifactorial disease, and many genes have been found that play a role to its incidence; therefore large numbers of patients with CHD may need to be screened to fully elucidate the contribution of BNC2. Animal models, including a mouse deficient in *BNC2*, could provide valuable tools for examining the role of *BNC2* over the span of development in multiple organ systems. Mouse embryos lacking *BNC1* do not develop past the two-cell stage, so a conditional knock-out strategy may be the most informative (Ma, Zeng et al. 2006).

It would be especially interesting to examine the expression of BNC2 in our tissues of interest (pulmonary vein, heart, lungs, skeletal muscle, and penis) over the course of embryonic and fetal development. Since the defects in our patient were all present at birth, it would be likely that any disruption of BNC2 would have its effects in the prenatal period. Although we did explore the tissue-specific expression of BNC2 in the human fetus, these only gave us a snapshot into one time point of development. It would also be useful to perform a more thorough examination of the tissue specificity of the many different isoforms of BNC2, especially those that specifically include or exclude the known motifs, such as the zinc fingers or nuclear localization sequence.

It is also possible that the translocations in either of our probands may have lead to epigenetic changes in the surrounding regions that caused dysregulation of the transcription of the genes in the breakpoint regions. There are many well-described cases that involve breakpoints that have one break in a heterochromatic region and another in a euchromatic region that lead to inactivation of the genes in the euchromatic region through aberrant methlylation patterns. (Rouyer, Simmler et al. 1987) This may not be evident in our allelic discrimination assays if it affects only tissue- or developmentally-specific control of those genes. It would be ideal to be able to perform these same experiments with RNA extracted from the tissues affected, i.e. the pancreas or pulmonary veins, but access to these samples would be prohibitively difficult to obtain until

autopsies are performed. This is another field of experimentation that would benefit from mouse models of these gene disruptions, as any tissue or developmental stage could be assayed.

Additionally, we examined the breakpoint areas for other functional elements that may have contributed to their phenotypes, such as enhancers, inactivators, insulators, and large non-coding RNAs, but found no evidence or them, mainly due to the lack of significant conservation in these areas. Large non-coding RNAs are of special interest in these cases because they have been shown to both enhance and inhibit genes in mammals. (Shamovsky and Nudler 2006) They have been shown to be important in the repression of multiple hormone pathways, including those important for glucose regulation. (Hatchell, Colley et al. 2006) It is thought that they provide a scaffold where transcriptional complexes can assemble and stabilize. Although it is thought that there are as-of-yet undiscovered large non-coding RNAs, all that have been previously described are in ultra-conserved regions of the genome, making it very unlikely that any of our breakpoint regions are affecting these RNA elements. (Shamovsky and Nudler 2006)

It is difficult to know with certainty if the pathologies of either of our probands are directly caused by their balances translocations. In our MODY-X patient, the breakpoint does not fall within a gene, which makes it less likely, but certainly not impossible, for it to have an effect. Also, since only one organ

system is involved, it is more likely that other etiologies may have been the cause. Additionally, the importance of MPP7 has yet to be determined, suggesting that perhaps, even if one copy is perturbed, it may have no effect on the organism. If further patients are found that have defects in *MPP7* and MODY, or a *MPP7* heterozygous gene-deficient mouse recapitulates his phenotype, these may allow us to establish direct causality of his breakpoint and disease process. However, in our patient with PAPVR, a very well-conserved gene, *BNC2*, is directly disrupted; it is thought to be a vital protein, since mouse embryos lacking *BNC1* do not develop past the two-cell stage (Ma, Zeng et al. 2006). This patient also has multiple congenital anomalies involves at least three organ systems (heart, penis, and skeletal muscle/bones), which is more indicative of a genetic cause.

Three different techniques were utilized to characterize the breakpoints of our probands: aCGH, FISH, and somatic cell hybridization. We were able to use aCGH to find the breakpoints in a carrier of a balanced translocation because a carrier of the unbalanced segregant was available through the products of conception of his wife's miscarriage. Using this method we were able to obtain, within a few hundred bps, the location of the breakpoints in a few weeks and minimal cost. Also, when the unbalanced segregant is available, no other materials need to be generated, which decreases the cost and time required. The data provided from aCGH also requires only minimal interpretation, and the location of the breakpoints is immediately evident. It is a post-genomic, unbiased

view of the entire chromosome or genome, therefore if the original cytogenetics location of the breakpoints is incorrect, it will not have any impact on the results. However, there is a slight chance that additional rearrangements or other cytogenetics perturbations may exist in the unbalanced segregant, and aCGH results must be confirmed through FISH on the cell of the balanced translocation carrier.

Sequential FISH is the second method used to identify the breakpoints, this does not require that an unbalanced segregant be available, but is more costly and time-consuming. Depending on the size of the breakpoint area that was defined through cytogenetics, many rounds of iterative FISH may be necessary to localize the breakpoint at a useful level. Each round may take between one and two weeks, with significant down time while each new order of BACs is shipped. This method is also dependent on the cytogenetic interpretation of the location of the FISH signals on the derivative chromosomes, which may be subject to error during the assignment of a signal to an individual chromosome. This is a viable and useful method when no unbalanced segregant is available, and also as a confirmation method for other techniques, but has significant drawbacks due to the cost and time required.

The third method utilized in this work, somatic cell hybridization with iterative PCRs on the isolated derivative chromosome, is unique amongst these three because it allows the molecular characterization of the breakpoints. It is also

inexpensive and fast to find the breakpoint once the derivative chromosome has been isolated from the normal copy, however, the somatic cell hybridization and subsequent passages necessary to obtain the chromosome are more difficult. Once the cells have been fused, it may take as long as two to three weeks before colonies become large enough to utilize, and then ten passages must be completed to allow time for the chromosomal content of those lines to stabilize. There is also a large element of chance in this process, as it is possible to obtain many hybrid colonies, and not have any that contain the required chromosomes. This technique is ideal for the molecular characterization of the breakpoints, but requires a significant amount of time to perform, and may result only in hybrids with useless chromosome contents.

An additional method for mapping translocation breakpoints has recently been described, in which flow-sorted derivative chromosomes are assayed using multiplex sequencing-by-synthesis technology. This technology allows a breakpoint to be molecularly characterized within weeks, at a resolution beyond what traditional FISH mapping provides. However, due to the current high cost of this type of screening it may not be practical for large-scale studies. (Chen, Kalscheuer et al. 2008)

The characterization of balanced translocations has led to the discovery of many novel disease-causing genes, even in the modern era with the ability of researchers to pursue genome-wide linkage and association studies. As with the

previously-described MODY genes, mutations in genes that affect only a small percentage of the population with a disease can lead to great insights into the pathophysiology of the condition, including new targets for drug discovery. The small number of patients affected by these mutations may not be large enough to make genome-wide studies useful. In addition, the time and resources required to perform genome-wide studies are much greater than those needed to characterize translocation breakpoints, making it a feasible alternative for individual labs or rare diseases. As with many studies focused on a single case, the problem of definitively proving a causal relationship between the translocation and disease can be a major hurdle to overcome. We have attempted to use mutation screening of the gene of interest in other patients with a similar condition to illustrate an additional link between the gene and disease, however, if the mutations in this gene only account for a small percentage of total cases the number of patients needed to be screened may be prohibitive. Despite this issue, history has shown that the characterization of balanced translocations in patients with a variety of diseases can be a successful and useful technique in the discovery of novel disease-causing genes.

Allera, A., M. A. Herbst, et al. (1995). "Mutations of the androgen receptor coding sequence are infrequent in patients with isolated hypospadias." J Clin Endocrinol Metab 80(9): 2697-9.

Altshuler, D., J. N. Hirschhorn, et al. (2000). "The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes." Nat Genet 26(1): 76-80.

Ashcroft, F. M., D. E. Harrison, et al. (1984). "Glucose induces closure of single potassium channels in isolated rat pancreatic beta-cells." Nature 312(5993): 446-8.

Bagwell, A. M., J. L. Bento, et al. (2005). "Genetic analysis of HNF4A polymorphisms in Caucasian-American type 2 diabetes." Diabetes 54(4): 1185-90.

Bartsocas, C. S. and A. Gerasimidi-Vazeou (2006). "Genetics of type 1 diabetes mellitus." Pediatr Endocrinol Rev 3 Suppl 3: 508-13.

Baskin, L. S. (2000). "Hypospadias and urethral development." J Urol 163(3): 951-6.

Bell, G. I. and K. S. Polonsky (2001). "Diabetes mellitus and genetically programmed defects in beta-cell function." Nature 414(6865): 788-91.

Bento, J. L., N. D. Palmer, et al. (2004). "Association of protein tyrosine phosphatase 1B gene polymorphisms with type 2 diabetes." Diabetes 53(11): 3007-12.

Bleyl, S., L. Nelson, et al. (1995). "A gene for familial total anomalous pulmonary venous return maps to chromosome 4p13-q12." Am J Hum Genet 56(2): 408-15.

Bleyl, S., H. D. Ruttenberg, et al. (1994). "Familial total anomalous pulmonary venous return: a large Utah-Idaho family." Am J Med Genet 52(4): 462-6.

Bleyl, S. B., L. D. Botto, et al. (2006). "Analysis of a Scottish founder effect narrows the TAPVR-1 gene interval to chromosome 4q12." Am J Med Genet A 140(21): 2368-73.

Boesgaard, T. W., S. I. Castella, et al. (2007). "A -243A-->G polymorphism upstream of the gene encoding GAD65 associates with lower levels of

body mass index and glycaemia in a population-based sample of 5857 middle-aged White subjects." Diabet Med 24(7): 702-6.

Bohl, J., N. Brimer, et al. (2007). "The stardust family protein MPP7 forms a tripartite complex with LIN7 and DLG1 that regulates the stability and localization of DLG1 to cell junctions." J Biol Chem 282(13): 9392-400.

Bottini, N., L. Musumeci, et al. (2004). "A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes." Nat Genet 36(4): 337-8.

Brackenridge, A., E. R. Pearson, et al. (2006). "Contrasting insulin sensitivity of endogenous glucose production rate in subjects with hepatocyte nuclear factor-1beta and -1alpha mutations." Diabetes 55(2): 405-11.

Budarf, M. L., J. Collins, et al. (1995). "Cloning a balanced translocation associated with DiGeorge syndrome and identification of a disrupted candidate gene." Nat Genet 10(3): 269-78.

Calabrese, A., D. Caton, et al. (2004). "Differentiating the effects of Cx36 and E-cadherin for proper insulin secretion of MIN6 cells." Exp Cell Res 294(2): 379-91.

Candib, L. M. (2007). "Obesity and diabetes in vulnerable populations: reflection on proximal and distal causes." Ann Fam Med 5(6): 547-56.

Chandley, A. C. (1991). "On the parental origin of de novo mutation in man." J Med Genet 28(4): 217-23.

Charollais, A., A. Gjinovci, et al. (2000). "Junctional communication of pancreatic beta cells contributes to the control of insulin secretion and glucose tolerance." J Clin Invest 106(2): 235-43.

Chen, W., V. Kalscheuer, et al. (2008). "Mapping translocation breakpoints by next-generation sequencing." Genome Res.

Chiba, H., T. Gotoh, et al. (2003). "Hepatocyte nuclear factor (HNF)-4alpha triggers formation of functional tight junctions and establishment of polarized epithelial morphology in F9 embryonal carcinoma cells." Exp Cell Res 286(2): 288-97.

Chomyn, A., A. Martinuzzi, et al. (1992). "MELAS mutation in mtDNA binding site for transcription termination factor causes defects in protein synthesis and in respiration but no change in levels of upstream and downstream mature transcripts." Proc Natl Acad Sci U S A 89(10): 4221-5.

Cinquetti, R., I. Badi, et al. (2008). "Transcriptional deregulation and a missense mutation define ANKRD1 as a candidate gene for total anomalous pulmonary venous return." Hum Mutat.

Cohen, O., C. Cans, et al. (1996). "Cartographic study: breakpoints in 1574 families carrying human reciprocal translocations." Hum Genet 97(5): 659-67.

Collares-Buzato, C. B., C. P. Carvalho, et al. (2004). "Upregulation of the expression of tight and adherens junction-associated proteins during maturation of neonatal pancreatic islets in vitro." J Mol Histol 35(8-9): 811-22.

Cremer, T., M. Cremer, et al. (2006). "Chromosome territories--a functional nuclear landscape." Curr Opin Cell Biol 18(3): 307-16.

Damcott, C. M., N. Hoppman, et al. (2004). "Polymorphisms in both promoters of hepatocyte nuclear factor 4-alpha are associated with type 2 diabetes in the Amish." Diabetes 53(12): 3337-41.

De Baere, E., Y. Fukushima, et al. (2000). "Identification of BPESC1, a novel gene disrupted by a balanced chromosomal translocation, t(3;4)(q23;p15.2), in a patient with BPES." Genomics 68(3): 296-304.

Deeb, S. S., L. Fajas, et al. (1998). "A Pro12Ala substitution in PPARgamma2 associated with decreased receptor activity, lower body mass index and improved insulin sensitivity." Nat Genet 20(3): 284-7.

DeLuca, S. A. (1984). "Partial anomalous pulmonary venous return." Am Fam Physician 30(4): 127-9.

Dubos, A., S. Pannetier, et al. (2008). "Inactivation of the CDKL3 gene at 5q31.1 by a balanced t(X;5) translocation associated with nonspecific mild mental retardation." Am J Med Genet A 146(10): 1267-79.

Duckett, J. W. (1995). "The current hype in hypospadiology." Br J Urol 76 Suppl 3: 1-7.

Easom, R. A. (1999). "CaM kinase II: a protein kinase with extraordinary talents germane to insulin exocytosis." Diabetes 48(4): 675-84.

Edghill, E. L., C. Bingham, et al. (2006). "Mutations in hepatocyte nuclear factor-1beta and their related phenotypes." J Med Genet 43(1): 84-90.

Edwards, J. E. (1953). "Pathologic and developmental considerations in anomalous pulmonary venous connection." Proc Staff Meet Mayo Clin 28(17): 441-52.

Egger, G., G. Liang, et al. (2004). "Epigenetics in human disease and prospects for epigenetic therapy." Nature 429(6990): 457-63.

Elliott, D. A., E. P. Kirk, et al. (2003). "Cardiac homeobox gene NKX2-5 mutations and congenital heart disease: associations with atrial septal defect and hypoplastic left heart syndrome." J Am Coll Cardiol 41(11): 2072-6.

Emery, A. E. (2000). "Emery-Dreifuss muscular dystrophy - a 40 year retrospective." Neuromuscul Disord 10(4-5): 228-32.

Escudero, T., M. Lee, et al. (2000). "Female gamete segregation in two carriers of translocations involving 2q and 14q." Prenat Diagn 20(3): 235-7.

Fajans, S. S., G. I. Bell, et al. (2001). "Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young." N Engl J Med 345(13): 971-80.

Fajans, S. S. and J. W. Conn (1960). "Tolbutamide-induced improvement in carbohydrate tolerance of young people with mild diabetes mellitus." Diabetes 9: 83-8.

Ferner, R. E. (1992). "Drug-induced diabetes." Baillieres Clin Endocrinol Metab 6(4): 849-66.

Florez, J. C., J. Hirschhorn, et al. (2003). "The inherited basis of diabetes mellitus: implications for the genetic analysis of complex traits." Annu Rev Genomics Hum Genet 4: 257-91.

Florez, J. C., K. A. Jablonski, et al. (2006). "TCF7L2 polymorphisms and progression to diabetes in the Diabetes Prevention Program." N Engl J Med 355(3): 241-50.

Frayling, T. M. (2007). "A new era in finding Type 2 diabetes genes-the unusual suspects." Diabet Med 24(7): 696-701.

Frayling, T. M., J. C. Evans, et al. (2001). "beta-cell genes and diabetes: molecular and clinical characterization of mutations in transcription factors." Diabetes 50 Suppl 1: S94-100.

Frayling, T. M., M. I. McCarthy, et al. (2000). "No evidence for linkage at candidate type 2 diabetes susceptibility loci on chromosomes 12 and 20 in United Kingdom Caucasians." J Clin Endocrinol Metab 85(2): 853-7.

Fredell, L., L. Iselius, et al. (2002). "Complex segregation analysis of hypospadias." Hum Genet 111(3): 231-4.

Freedom, R. M. (2004). "The natural and modified history of congenital heart disease."

Froguel, P., H. Zouali, et al. (1993). "Familial hyperglycemia due to mutations in glucokinase. Definition of a subtype of diabetes mellitus." N Engl J Med 328(10): 697-702.

Fukami, M., Y. Wada, et al. (2006). "CXorf6 is a causative gene for hypospadias." Nat Genet 38(12): 1369-71.

Fukami, M., Y. Wada, et al. (2007). "Mastermind-like domain containing 1 (MAMLD1 or CXorf6) transactivates the Hes3 promoter, augments testosterone production, and contains the SF1 target sequence." J Biol Chem.

Gabellini, D., M. R. Green, et al. (2002). "Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle." Cell 110(3): 339-48.

Gallentine, M. L., A. F. Morey, et al. (2001). "Hypospadias: a contemporary epidemiologic assessment." Urology 57(4): 788-90.

Gimelbrant, A., J. N. Hutchinson, et al. (2007). "Widespread monoallelic expression on human autosomes." Science 318(5853): 1136-40.

Giuffrida, F. M. and A. F. Reis (2005). "Genetic and clinical characteristics of maturity-onset diabetes of the young." Diabetes Obes Metab 7(4): 318-26.

Gloyn, A. L., S. Ellard, et al. (2002). "Maturity-onset diabetes of the young caused by a balanced translocation where the 20q12 break point results in disruption upstream of the coding region of hepatocyte nuclear factor-4alpha (HNF4A) gene." Diabetes 51(7): 2329-33.

Gloyn, A. L., M. N. Weedon, et al. (2003). "Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes." Diabetes 52(2): 568-72.

Gomis, A., J. V. Sanchez-Andres, et al. (1996). "Oscillatory patterns of electrical activity in mouse pancreatic islets of Langerhans recorded in vivo." Pflugers Arch 432(3): 510-5.

Gragnoli, C., T. Lindner, et al. (1997). "Maturity-onset diabetes of the young due to a mutation in the hepatocyte nuclear factor-4 alpha binding site in the promoter of the hepatocyte nuclear factor-1 alpha gene." Diabetes 46(10): 1648-51.

Grant, S. F., G. Thorleifsson, et al. (2006). "Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes." Nat Genet 38(3): 320-3.

Guy, P. M., D. A. Kenny, et al. (1999). "The PDZ domain of the LIM protein enigma binds to beta-tropomyosin." Mol Biol Cell 10(6): 1973-84.

Hanis, C. L., E. Boerwinkle, et al. (1996). "A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2." Nat Genet 13(2): 161-6.

Hansen, S. K., M. Parrizas, et al. (2002). "Genetic evidence that HNF-1alpha-dependent transcriptional control of HNF-4alpha is essential for human pancreatic beta cell function." J Clin Invest 110(6): 827-33.

Hansen, S. K., C. S. Rose, et al. (2005). "Variation near the hepatocyte nuclear factor (HNF)-4alpha gene associates with type 2 diabetes in the Danish population." Diabetologia 48(3): 452-8.

Harris, D. L., B. L. Siu, et al. (2004). "Mosaic ring 12p and total anomalous pulmonary venous return." Am J Med Genet A 131(1): 91-3.

Hatchell, E. C., S. M. Colley, et al. (2006). "SLIRP, a small SRA binding protein, is a nuclear receptor corepressor." Mol Cell 22(5): 657-68.

Hattersley, A. T. and E. R. Pearson (2006). "Minireview: pharmacogenetics and beyond: the interaction of therapeutic response, beta-cell physiology, and genetics in diabetes." Endocrinology 147(6): 2657-63.

Healey, J. E., Jr. (1952). "An anatomic survey of anomalous pulmonary veins: their clinical significance." J Thorac Surg 23(5): 433-44.

Henquin, J. C. (2000). "Triggering and amplifying pathways of regulation of insulin secretion by glucose." Diabetes 49(11): 1751-60.

Higgins, A. W., F. S. Alkuraya, et al. (2008). "Characterization of apparently balanced chromosomal rearrangements from the developmental genome anatomy project." Am J Hum Genet 82(3): 712-22.

Holaska, J. M., S. Rais-Bahrami, et al. (2006). "Lmo7 is an emerin-binding protein that regulates the transcription of emerin and many other muscle-relevant genes." Hum Mol Genet 15(23): 3459-72.

Holder, J. L., Jr., N. F. Butte, et al. (2000). "Profound obesity associated with a balanced translocation that disrupts the SIM1 gene." Hum Mol Genet 9(1): 101-8.

Horikawa, Y., N. Iwasaki, et al. (1997). "Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY." Nat Genet 17(4): 384-5.

Hoveyda, N., J. P. Shield, et al. (1999). "Neonatal diabetes mellitus and cerebellar hypoplasia/agenesis: report of a new recessive syndrome." J Med Genet 36(9): 700-4.

Hyvarinen, A. K., J. L. Pohjoismaki, et al. (2007). "The mitochondrial transcription termination factor mTERF modulates replication pausing in human mitochondrial DNA." Nucleic Acids Res 35(19): 6458-74.

Jaeckel, E., M. Manns, et al. (2002). "Viruses and diabetes." Ann N Y Acad Sci 958: 7-25.

Jamieson, R. V., N. Farrar, et al. (2007). "Characterization of a familial t(16;22) balanced translocation associated with congenital cataract leads to identification of a novel gene, TMEM114, expressed in the lens and disrupted by the translocation." Hum Mutat 28(10): 968-77.

Jamieson, R. V., R. Perveen, et al. (2002). "Domain disruption and mutation of the bZIP transcription factor, MAF, associated with cataract, ocular anterior segment dysgenesis and coloboma." Hum Mol Genet 11(1): 33-42.

Jenkins, K. J., A. Correa, et al. (2007). "Noninherited risk factors and congenital cardiovascular defects: current knowledge: a scientific statement from the American Heart Association Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics." Circulation 115(23): 2995-3014.

Jirasek, J. E., J. Raboch, et al. (1968). "The relationship between the development of gonads and external genitals in human fetuses." Am J Obstet Gynecol 101(6): 830-3.

Joslin, E. P., C. R. Kahn, et al. (2005). Joslin's diabetes mellitus. Philadelphia, Pa., Lippincott Williams & Willkins, c2005.

Katoh, M. (2004). "Identification and characterization of human MPP7 gene and mouse Mpp7 gene in silico." Int J Mol Med 13(2): 333-8.

Katoh, M. (2007). "Networking of WNT, FGF, Notch, BMP, and Hedgehog signaling pathways during carcinogenesis." Stem Cell Rev 3(1): 30-8.

Kawara, H., T. Yamamoto, et al. (2006). "Narrowing candidate region for monosomy 9p syndrome to a 4.7-Mb segment at 9p22.2-p23." Am J Med Genet A 140(4): 373-7.

Kehrer-Sawatzki, H., J. Haussler, et al. (1997). "The second case of a t(17;22) in a family with neurofibromatosis type 1: sequence analysis of the breakpoint regions." Hum Genet 99(2): 237-47.

Kim, H. G., S. R. Herrick, et al. (2005). "Hypogonadotropic hypogonadism and cleft lip and palate caused by a balanced translocation producing haploinsufficiency for FGFR1." J Med Genet 42(8): 666-72.

Kleinjan, D. A. and V. van Heyningen (2005). "Long-range control of gene expression: emerging mechanisms and disruption in disease." Am J Hum Genet 76(1): 8-32.

Kulkarni, R. N. and C. R. Kahn (2004). "Molecular biology. HNFs--linking the liver and pancreatic islets in diabetes." Science 303(5662): 1311-2.

Kurahashi, H. and B. S. Emanuel (2001). "Long AT-rich palindromes and the constitutional t(11;22) breakpoint." Hum Mol Genet 10(23): 2605-17.

Lang, J. (1999). "Molecular mechanisms and regulation of insulin exocytosis as a paradigm of endocrine secretion." Eur J Biochem 259(1-2): 3-17.

Leiter, E. H. and C. H. Lee (2005). "Mouse models and the genetics of diabetes: is there evidence for genetic overlap between type 1 and type 2 diabetes?" Diabetes 54 Suppl 2: S151-8.

Ligon, A. H., S. D. Moore, et al. (2005). "Constitutional rearrangement of the architectural factor HMGA2: a novel human phenotype including overgrowth and lipomas." Am J Hum Genet 76(2): 340-8.

Love-Gregory, L. D., J. Wasson, et al. (2004). "A common polymorphism in the upstream promoter region of the hepatocyte nuclear factor-4 alpha gene on chromosome 20q is associated with type 2 diabetes and appears to contribute to the evidence for linkage in an ashkenazi jewish population." Diabetes 53(4): 1134-40.

Lucassen, A. M., C. Julier, et al. (1993). "Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment of DNA spanning the insulin gene and associated VNTR." Nat Genet 4(3): 305-10.

Ma, J., F. Zeng, et al. (2006). "Basonuclin: a novel mammalian maternal-effect gene." Development 133(10): 2053-62.

Maassen, J. A., L. M. t Hart, et al. (2007). "Lessons that can be learned from patients with diabetogenic mutations in mitochondrial DNA: implications for common type 2 diabetes." Curr Opin Clin Nutr Metab Care 10(6): 693-7.

Maestro, M. A., S. F. Boj, et al. (2003). "Hnf6 and Tcf2 (MODY5) are linked in a gene network operating in a precursor cell domain of the embryonic pancreas." Hum Mol Genet 12(24): 3307-14.

Malecki, M. T., U. S. Jhala, et al. (1999). "Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus." Nat Genet 23(3): 323-8.

Mateos-Langerak, J., S. Goetze, et al. (2007). "Nuclear architecture: Is it important for genome function and can we prove it?" J Cell Biochem 102(5): 1067-75.

McCarthy, M. I. (2003). "Growing evidence for diabetes susceptibility genes from genome scan data." Curr Diab Rep 3(2): 159-67.

Menten, B., N. Maas, et al. (2006). "Emerging patterns of cryptic chromosomal imbalance in patients with idiopathic mental retardation and multiple congenital anomalies: a new series of 140 patients and review of published reports." J Med Genet 43(8): 625-33.

Meyre, D., N. Bouatia-Naji, et al. (2005). "Variants of ENPP1 are associated with childhood and adult obesity and increase the risk of glucose intolerance and type 2 diabetes." Nat Genet 37(8): 863-7.

Munne, S., M. Bahce, et al. (1998). "Case report: chromatid exchange and predivision of chromatids as other sources of abnormal oocytes detected by preimplantation genetic diagnosis of translocations." Prenat Diagn 18(13): 1450-8.

Nakagawa, N., M. Hoshijima, et al. (2000). "ENH, containing PDZ and LIM domains, heart/skeletal muscle-specific protein, associates with cytoskeletal proteins through the PDZ domain." Biochem Biophys Res Commun 272(2): 505-12.

Neill, C. A. (1956). "Development of the pulmonary veins; with reference to the embryology of anomalies of pulmonary venous return." Pediatrics 18(6): 880-7.

Nerup, J., P. Platz, et al. (1974). "HL-A antigens and diabetes mellitus." Lancet 2(7885): 864-6.

Ng, P. C. and S. Henikoff (2003). "SIFT: Predicting amino acid changes that affect protein function." Nucleic Acids Res 31(13): 3812-4.

Nichols, D. G. (2006). Critical heart disease in infants and children. Philadelphia, Mosby.

Nothwang, H. G., H. G. Kim, et al. (2001). "Functional hemizygosity of PAFAH1B3 due to a PAFAH1B3-CLK2 fusion gene in a female with mental retardation, ataxia and atrophy of the brain." Hum Mol Genet 10(8): 797-806.

O'Shaughnessy, P. J., P. J. Baker, et al. (2007). "Developmental changes in human fetal testicular cell numbers and messenger ribonucleic acid levels during the second trimester." J Clin Endocrinol Metab 92(12): 4792-801.

Odom, D. T., N. Zizlsperger, et al. (2004). "Control of pancreas and liver gene expression by HNF transcription factors." Science 303(5662): 1378-81.

Onengut-Gumuscu, S. and P. Concannon (2005). "The genetics of type 1 diabetes: lessons learned and future challenges." J Autoimmun 25 Suppl: 34-9.

Ooshio, T., K. Irie, et al. (2004). "Involvement of LMO7 in the association of two cell-cell adhesion molecules, nectin and E-cadherin, through afadin and alpha-actinin in epithelial cells." J Biol Chem 279(30): 31365-73.

Orci, L., F. Malaisse-Lagae, et al. (1975). "Cell contacts in human islets of Langerhans." J Clin Endocrinol Metab 41(5): 841-4.

Orci, L., R. H. Unger, et al. (1973). "Structural coupling between pancreatic islet cells." Experientia 29(8): 1015-8.

Owen, K. R. and M. I. McCarthy (2007). "Genetics of type 2 diabetes." Curr Opin Genet Dev 17(3): 239-44.

Page, L. A. (1979). "Inheritance of uncomplicated hypospadias." Pediatrics 63(5): 788-90.

Page, S. L. and L. G. Shaffer (1997). "Nonhomologous Robertsonian translocations form predominantly during female meiosis." Nat Genet 15(3): 231-2.

Paulozzi, L. J., J. D. Erickson, et al. (1997). "Hypospadias trends in two US surveillance systems." Pediatrics 100(5): 831-4.

Pearson, E. R., S. F. Boj, et al. (2007). "Macrosomia and hyperinsulinaemic hypoglycaemia in patients with heterozygous mutations in the HNF4A gene." PLoS Med 4(4): e118.

Pearson, E. R., I. Flechtner, et al. (2006). "Switching from insulin to oral sulfonylureas in patients with diabetes due to Kir6.2 mutations." N Engl J Med 355(5): 467-77.

Pierpont, M. E., C. T. Basson, et al. (2007). "Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics." Circulation 115(23): 3015-38.

Prentki, M. and C. J. Nolan (2006). "Islet beta cell failure in type 2 diabetes." J Clin Invest 116(7): 1802-12.

Pugliese, A. (2004). "Genetics of type 1 diabetes." Endocrinol Metab Clin North Am 33(1): 1-16, vii.

Putilina, T., C. Jaworski, et al. (1998). "Analysis of a human cDNA containing a tissue-specific alternatively spliced LIM domain." Biochem Biophys Res Commun 252(2): 433-9.

Raeder, H., S. Johansson, et al. (2006). "Mutations in the CEL VNTR cause a syndrome of diabetes and pancreatic exocrine dysfunction." Nat Genet 38(1): 54-62.

Raisher, B. D., S. B. Dowton, et al. (1991). "Father and two children with total anomalous pulmonary venous connection." Am J Med Genet 40(1): 105-6.

Ramer, J. C., P. N. Mowrey, et al. (1990). "Five children with del (2)(q31q33) and one individual with dup (2)(q31q33) from a single family: review of brain, cardiac, and limb malformations." Am J Med Genet 37(3): 392-400.

Ray, P. N., B. Belfall, et al. (1985). "Cloning of the breakpoint of an X;21 translocation associated with Duchenne muscular dystrophy." Nature 318(6047): 672-5.

Reya, T., A. W. Duncan, et al. (2003). "A role for Wnt signalling in self-renewal of haematopoietic stem cells." Nature 423(6938): 409-14.

Romano, R. A., H. Li, et al. (2004). "Identification of Basonuclin2, a DNA-binding zinc-finger protein expressed in germ tissues and skin keratinocytes." Genomics 83(5): 821-33.

Rouyer, F., M. C. Simmler, et al. (1987). "A sex chromosome rearrangement in a human XX male caused by Alu-Alu recombination." Cell 51(3): 417-25.

Royer-Pokora, B., L. M. Kunkel, et al. (1986). "Cloning the gene for an inherited human disorder--chronic granulomatous disease--on the basis of its chromosomal location." Nature 322(6074): 32-8.

Rudolph, A. M. (2001). Congenital Diseases of the Heart: Clinical Physiological Considerations. Armonk, NY, Futura Publishing Company Inc.

Sander, T. L., D. B. Klinkner, et al. (2006). "Molecular and cellular basis of congenital heart disease." Pediatr Clin North Am 53(5): 989-1009, x.

Satoda, M., F. Zhao, et al. (2000). "Mutations in TFAP2B cause Char syndrome, a familial form of patent ductus arteriosus." Nat Genet 25(1): 42-6.

Saxena, R., B. F. Voight, et al. (2007). "Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels." Science 316(5829): 1331-6.

Sellick, G. S., K. T. Barker, et al. (2004). "Mutations in PTF1A cause pancreatic and cerebellar agenesis." Nat Genet 36(12): 1301-5.

Semenova, E., X. Wang, et al. (2003). "An engineered 800 kilobase deletion of Uchl3 and Lmo7 on mouse chromosome 14 causes defects in viability, postnatal growth and degeneration of muscle and retina." Hum Mol Genet 12(11): 1301-12.

Semino, M. C., E. E. de Gagliardino, et al. (1987). "Islet cells-tight junctions: changes in its number induced by glucose." Acta Physiol Pharmacol Latinoam 37(4): 533-9.

Servitja, J. M. and J. Ferrer (2004). "Transcriptional networks controlling pancreatic development and beta cell function." Diabetologia 47(4): 597-613.

Shaffer, L. G. and J. R. Lupski (2000). "Molecular mechanisms for constitutional chromosomal rearrangements in humans." Annu Rev Genet 34: 297-329.

Shalaby, F., J. Ho, et al. (1997). "A requirement for Flk1 in primitive and definitive hematopoiesis and vasculogenesis." Cell 89(6): 981-90.

Shamovsky, I. and E. Nudler (2006). "Gene control by large noncoding RNAs." Sci STKE 2006(355): pe40.

Silver, R. I. and D. W. Russell (1999). "5alpha-reductase type 2 mutations are present in some boys with isolated hypospadias." J Urol 162(3 Pt 2): 1142-5.

Sladek, R., G. Rocheleau, et al. (2007). "A genome-wide association study identifies novel risk loci for type 2 diabetes." Nature 445(7130): 881-5.

Solymar, L., K. G. Sabel, et al. (1987). "Total anomalous pulmonary venous connection in siblings. Report on three families." Acta Paediatr Scand 76(1): 124-7.

Spath, G. F. and M. C. Weiss (1998). "Hepatocyte nuclear factor 4 provokes expression of epithelial marker genes, acting as a morphogen in dedifferentiated hepatoma cells." J Cell Biol 140(4): 935-46.

Srivastava, D. (2006). "Genetic regulation of cardiogenesis and congenital heart disease." Annu Rev Pathol 1: 199-213.

Stankiewicz, P. and J. R. Lupski (2002). "Genome architecture, rearrangements and genomic disorders." Trends Genet 18(2): 74-82.

Steiner, D. F., H. S. Tager, et al. (1990). "Lessons learned from molecular biology of insulin-gene mutations." Diabetes Care 13(6): 600-9.

Stoffers, D. A., J. Ferrer, et al. (1997). "Early-onset type-II diabetes mellitus (MODY4) linked to IPF1." Nat Genet 17(2): 138-9.

Stucke, V. M., E. Timmerman, et al. (2007). "The MAGUK protein MPP7 binds to the polarity protein hDlg1 and facilitates epithelial tight junction formation." Mol Biol Cell 18(5): 1744-55.

Su, A. I., T. Wiltshire, et al. (2004). "A gene atlas of the mouse and human protein-encoding transcriptomes." Proc Natl Acad Sci U S A 101(16): 6062-7.

Sutherland, R. W., J. S. Wiener, et al. (1996). "Androgen receptor gene mutations are rarely associated with isolated penile hypospadias." J Urol 156(2 Pt 2): 828-31.

Triggs-Raine, B. L., R. D. Kirkpatrick, et al. (2002). "HNF-1alpha G319S, a transactivation-deficient mutant, is associated with altered dynamics of diabetes onset in an Oji-Cree community." Proc Natl Acad Sci U S A 99(7): 4614-9.

Trucco, M. and J. S. Dorman (1989). "Immunogenetics of insulin-dependent diabetes mellitus in humans." Crit Rev Immunol 9(3): 201-45.

Ueda, H., J. M. Howson, et al. (2003). "Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease." Nature 423(6939): 506-11.

Unger, R. H., R. E. Dobbs, et al. (1978). "Insulin, glucagon, and somatostatin secretion in the regulation of metabolism." Annu Rev Physiol 40: 307-43.

Van Dyke, D. L., L. Weiss, et al. (1983). "The frequency and mutation rate of balanced autosomal rearrangements in man estimated from prenatal genetic studies for advanced maternal age." Am J Hum Genet 35(2): 301-8.

Vanhoutteghem, A. and P. Djian (2004). "Basonuclin 2: an extremely conserved homolog of the zinc finger protein basonuclin." Proc Natl Acad Sci U S A 101(10): 3468-73.

Vanhoutteghem, A. and P. Djian (2006). "Basonuclins 1 and 2, whose genes share a common origin, are proteins with widely different properties and functions." Proc Natl Acad Sci U S A 103(33): 12423-8.

Vanhoutteghem, A. and P. Djian (2007). "The human basonuclin 2 gene has the potential to generate nearly 90,000 mRNA isoforms encoding over 2000 different proteins." Genomics 89(1): 44-58.

Velho, G. and P. Froguel (1998). "Genetic, metabolic and clinical characteristics of maturity onset diabetes of the young." Eur J Endocrinol 138(3): 233-9.

Wang, C. C., E. T. Wu, et al. (2008). "Scimitar syndrome: incidence, treatment, and prognosis." Eur J Pediatr 167(2): 155-60.

Warburton, D. (1984). "Outcome of cases of de novo structural rearrangements diagnosed at amniocentesis." Prenat Diagn 4 Spec No: 69-80.

Warburton, D. (1991). "De novo balanced chromosome rearrangements and extra marker chromosomes identified at prenatal diagnosis: clinical significance and distribution of breakpoints." Am J Hum Genet 49(5): 995-1013.

Watanabe, R. M., M. H. Black, et al. (2007). "Genetics of gestational diabetes mellitus and type 2 diabetes." Diabetes Care 30 Suppl 2: S134-40.

Weedon, M. N., V. J. Clark, et al. (2006). "A common haplotype of the glucokinase gene alters fasting glucose and birth weight: association in six studies and population-genetics analyses." Am J Hum Genet 79(6): 991-1001.

Weedon, M. N., K. R. Owen, et al. (2004). "Common variants of the hepatocyte nuclear factor-4alpha P2 promoter are associated with type 2 diabetes in the U.K. population." Diabetes 53(11): 3002-6.

Weiman, D. S., K. Lee, et al. (1985). "Partial anomalous pulmonary venous return: a ten-year experience." Tex Heart Inst J 12(3): 239-43.

Wieczorek, D., M. Krause, et al. (2000). "Unexpected high frequency of de novo unbalanced translocations in patients with Wolf-Hirschhorn syndrome (WHS)." J Med Genet 37(10): 798-804.

Winckler, W., M. N. Weedon, et al. (2007). "Evaluation of common variants in the six known maturity-onset diabetes of the young (MODY) genes for association with type 2 diabetes." Diabetes 56(3): 685-93.

Wu, C. H., W. L. Hwu, et al. (2001). "Deletion of 11q24.2-qter with agenesis of unilateral internal carotid artery and total anomalous pulmonary venous return." Am J Med Genet 103(3): 245-8.

Xu, G. M. and M. A. Arnaout (2002). "WAC, a novel WW domain-containing adapter with a coiled-coil region, is colocalized with splicing factor SC35." Genomics 79(1): 87-94.

Yamagata, K., H. Furuta, et al. (1996). "Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1)." Nature 384(6608): 458-60.

Yamagata, K., N. Oda, et al. (1996). "Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3)." Nature 384(6608): 455-8.

Yoshiura, K., J. Machida, et al. (1998). "Characterization of a novel gene disrupted by a balanced chromosomal translocation t(2;19)(q11.2;q13.3) in a family with cleft lip and palate." Genomics 54(2): 231-40.