

# **Dissertation**

## **Positional Cloning of New Disease Genes for Kallmann Syndrome and Potocki-Shaffer Syndrome**

Zur Erlangung des akademischen Grades Doktor  
der Naturwissenschaften (Dr. rer. nat.)

eingereicht im Fachbereich Biologie, Chemie, Pharmazie  
der Freien Universität Berlin

von  
Hyung-Goo Kim  
aus Busan, Korea

June 2012

**Positional Cloning of New Disease Genes for Kallmann Syndrome and  
Potocki-Shaffer Syndrome**



**Max-Planck-Institute  
for Molecular Genetics**

Hyung-Goo Kim

*Department of Human Molecular Genetics  
Max-Planck-Institute for Molecular Genetics  
Ihnestraße 63-73, 14195 Berlin, Germany*

First Reviewer: Prof. Dr. Hans-Hilger Ropers

Second Reviewer: Prof. Dr. Stephan Sigrist

# Contents

<b>Dedication</b> -----	vi
<b>Figures</b> -----	vii
<b>1. Aim of the thesis</b> -----	1
1.1 Kallmann syndrome-----	1
1.2 Potocki-Shaffer syndrome-----	1
<b>2. List of original publications</b> -----	2
2.1 Paper I-----	3
2.2 Paper II-----	4
2.3 Paper III-----	5
2.4 Paper IV-----	6
<b>3. ASHG Abstract</b> -----	7
<b>4. Introduction</b> -----	8
4.1 Balanced translocations: visible bridges between genotypes and phenotypes-----	8
4.2 Structural chromosome rearrangements-----	10
4.3 A historical perspective: balanced translocations and positional cloning-----	13
4.4 Balanced translocations associated with abnormal phenotypes--	16
4.5 Balanced translocations in recessive disorders-----	19
4.6 Positional effect-----	19
4.7 Balanced translocation as a “double hit”-----	21
4.8 Balanced translocations with discordant phenotype-----	21
4.9 Balanced translocations involving the X chromosome-----	23
<b>5. Kallmann syndrome</b> -----	26
5.1 Introduction-----	26
5.2 Genetics-----	28
5.3 Subjects-----	29
5.3.1 Positional cloning of t(7;8)(p12.3;p11.2)dn-----	29
5.3.2 Candidate gene approach with <i>CHD7</i> -----	29
5.3.3 Positional cloning of t(10;12)(q26.12;q13.11)dn-----	30

5.4 Hypothesis-----	30
5.4.1 Positional cloning of t(7;8)(p12.3;p11.2)dn-----	30
5.4.2 Candidate gene approach with <i>CHD7</i> -----	31
5.4.3 Positional cloning of t(10;12)(q26.12;q13.11)dn-----	32
5.5 Results and Discussion-----	32
5.5.1 Positional cloning of t(7;8)(p12.3;p11.2)dn-----	32
5.5.2 Candidate gene approach with <i>CHD7</i> -----	36
5.5.3 Positional cloning of t(10;12)(q26.12;q13.11)dn-----	40
<b>6. Potocki-Shaffer syndrome-----</b>	<b>45</b>
6.1 Introduction and Genetics-----	45
6.2 Subjects-----	46
6.2.1 DGAP012 patient with t(11;19)(p11.2;p13.2)dn-----	46
6.2.2 MCN1762 patient with t(1;11)(p13;p11.2)dn-----	47
6.2.3 GM03316-----	48
6.3 Hypothesis-----	49
6.4 Results-----	49
6.4.1 <i>PHF21A</i> is disrupted in unrelated subjects with chromosomal translocations, ID, and CFA-----	49
6.4.2 <i>PHF21A</i> maps to the critical PSS interval associated with ID and CFA-----	51
6.4.3 Murine <i>Phf21a</i> is expressed in the CNS and cranial bones-----	53
6.4.4 Suppression of zebrafish <i>phf21a</i> expression causes CFA and neuronal apoptosis-----	54
6.4.5 Disruption of <i>PHF21A</i> in the translocation subjects derepresses <i>SCN3A</i> -----	55
6.5 Discussion -----	56
<b>7. Outlook and Conclusion -----</b>	<b>61</b>
<b>8. Acknowledgements-----</b>	<b>62</b>
<b>9. Literature-----</b>	<b>63</b>
<b>10. Summary-----</b>	<b>76</b>
<b>11. Zusammenfassung-----</b>	<b>77</b>

<b>12. Appendix</b> -----	79
12.1 Paper I-----	80
12.2 Paper II-----	88
12.3 Paper III-----	102
12.4 Paper IV-----	137

부모님께

*This thesis is dedicated to my parents*

*Sang-Hwa Kim and Jeong-Ok Kim*

## Figures

Figure 1. Positional cloning using chromosome rearrangement----- 9

Figure 2. The genesis of a balanced chromosome translocation----- 12

Figure 3. The genesis of a pericentric balanced chromosome inversion----- 14

## 1. Aims of the thesis

The general objective of this thesis has been to provide new genetic insight into Kallmann syndrome (KS) and Potocki-Shaffer syndrome (PSS).

Specific aims have been:

### 1.1 Kallmann Syndrome

- ▶ To map the breakpoints of two balanced translocation patients with idiopathic hypogonadotropic hypogonadism (IHH) and KS
- ▶ To identify a new KS gene by a candidate gene approach
- ▶ To identify genetic loci and disease genes causing KS
- ▶ To identify KS causing gene mutations and to elucidate the function and structure of the protein encoded by the disease genes

### 1.2 Potocki-Shaffer Syndrome

- ▶ To perform *in silico* comparative deletion mapping at 11p11.2 to refine the candidate gene region for intellectual disability (ID) and craniofacial anomalies (CFA) in PSS.
- ▶ To map the breakpoints of two unrelated balanced translocation patients with ID and CFA
- ▶ To identify a genetic locus and a disease gene causing ID and CFA in PSS

## **2. List of original publications**

This thesis is based on the following four original peer-reviewed publications. A list of papers from the results presented in this thesis are given below. The copies of the published or accepted papers are provided in the appendix.

## 2.1 Paper I

**Hyung-Goo Kim**, Steven R. Herrick, Emma Lemyre, Shotaro Kishikawa, Joseph A. Salisz, Stephanie Seminara, Marcy E. MacDonald, Gail A. P. Bruns, Cynthia C. Morton, Brad J. Quade, James F. Gusella\*

\*Corresponding author

### ***Hypogonadotropic hypogonadism and cleft lip and palate caused by a balanced translocation producing haploinsufficiency for FGFR1***

Journal of Medical Genetics 2005 Aug;42(8):666-72.

#### **Hyung-Goo Kim's contribution:**

1. Conceiving the idea of 8p11.2 as a disease gene locus for IHH/KS
2. Mapping and delineation of the breakpoint region of 46,XY,t(7;8)(p12.3;p11.2)dn
3. Southern blot analysis to refine the breakpoint regions of 7p12.3 and 8p11.2
4. Cloning of two junction fragments from der(7) and der(8) chromosomes
5. Fusion genes amplification from t(7;8)(p12.3;p11.2)dn
6. Manuscript writing

## 2.2 Paper II

Hyung-Goo Kim, Ingo Kurth, Fei Lan, Irene Meliciani, Wolfgang Wenzel, Soo Hyun Eom, Gil Bu Kang, Georg Rosenberger, Mustafa Tekin, Metin Ozata, David P. Bick, Richard J. Sherins, Steven L. Walker, Yang Shi, James F. Gusella, and Lawrence C. Layman\*

\* Corresponding author

### ***Mutations in CHD7, Encoding a Chromatin-Remodeling Protein, Cause Idiopathic Hypogonadotropic Hypogonadism and Kallmann Syndrome***

American Journal of Human Genetics, 2008 Oct;83(4):511-9.

#### **Hyung-Goo Kim's contribution:**

1. Conceiving the idea of *CHD7* as a potential candidate gene for IHH/KS
2. Mutation screening of *CHD7* in 50 and 51 patients with KS and IHH respectively
3. Mutation screening of exons 6-10 in *CHD7* in additional 96 IHH/KS patients
4. RT-PCR in two patients (one with IHH and another with KS) with intronic mutations of *CHD7*, which cause exon skipping
5. Multiple protein-sequence alignment of *CHD7* with its orthologs
6. Semi-quantitative RT-PCR analysis in rat tissues
7. Manuscript writing and planning the experiments with other collaborators

## 2.3 Paper III

**Hyung-Goo Kim\***, Jang-Won Ahn, Ingo Kurth, Reinhard Ullmann, Hyun-Taek Kim, Anita Kulharya, Kyung-Soo Ha, Yasuhide Itokawa, Irene Meliciani, Wolfgang Wenzel, Deresa Lee, Georg Rosenberger, Metin Ozata, David P. Bick, Richard J. Sherins, Takahiro Nagase, Mustafa Tekin, Soo-Hyun Kim, Cheol-Hee Kim, Hans-Hilger Ropers, James F. Gusella, Vera Kalscheuer, Cheol Yong Choi, and Lawrence C. Layman\*

\*Corresponding author

### ***WDR11, a WD Protein that Interacts with Transcription Factor EMX1, Is Mutated in Idiopathic Hypogonadotropic Hypogonadism and Kallmann Syndrome***

American Journal of Human Genetics, 2010 Oct 8;87(4):465-479.

#### **Hyung-Goo Kim's contribution:**

1. Conceiving the idea of 10q26 as a disease gene locus for IHH/KS
2. Mapping and delineation of the breakpoint region of  
46,XY,t(10;12)(q26.12;q13.11)dn
3. Cloning of two junction fragments from der(10) and der(12) chromosomes
4. Mutation analysis of the following genes:  
*WDR11* in 201 IHH/KS patients, *FGFR2* in 123 IHH/KS patients, *PPAPDC1A* in 60 IHH/KS patients, *SEC23IP* in 120 IHH/KS patients.
5. Isolation of RNA from the lymphoblastoid cell lines of the translocation patients and rat tissues and ensuing quantitative RT-PCR
6. Manuscript writing and planning the experiments with other collaborators

## 2.4 Paper IV

**Hyung-Goo Kim\***, Hyun Taek Kim\*\*, Natalia Leach\*\*, Fei Lan\*\*, Reinhard Ullmann, Asli Silahdaroglu, Ingo Kurth, Anja Nowka, Ihn Sik Seong, Yiping Shen, Michael E. Talkowski, Douglas Ruderfer, Ji-Hyun Lee, Caron Glotzbach, Kyungsoo Ha, Susanne Kjærgaard, Alex V. Levin, Bernd F. Romeike, Tjitske Kleefstra, Oliver Bartsch, Sarah H. Elsea, Ethylin W. Jabs, Marcy E. MacDonald, David J. Harris, Bradley J. Quade, Hans-Hilger Ropers, Lisa G. Shaffer, Kerstin Kutsche, Lawrence C. Layman, Niels Tommerup, Vera M. Kalscheuer, Yang Shi, Cynthia C. Morton, Cheol-Hee Kim, James F. Gusella

\*Corresponding author

\*\* These authors contributed equally to this work

***Translocations disrupting PHF21A in the Potocki-Shaffer syndrome region are associated with intellectual disability and craniofacial anomalies***

In press in July 2012 issue of the American Journal of Human Genetics

### **Hyung-Goo Kim's contribution:**

1. Conceiving the idea of 11p11.2 as a disease gene locus for ID and CFA in PSS
2. *In silico* comparative deletion mapping to narrow down the candidate gene region of ID and CFA of PSS
3. Mapping and delineation of the breakpoint region of 46,XY, t(11;19)(p11.2;p13.2)dn
4. Southern blot analysis to refine the breakpoint regions of 11p11.2 and 19p13.2
5. Cloning of two junction fragments from der(11) and der(19) chromosomes from t(11;19)(p11.2;p13.2)dn
6. Fusion genes amplification from t(11;19)(p11.2;p13.2)dn
7. Identification and mapping as well as delineation of the breakpoint region of 46,XX, t(1;11)(p13;p11.2)dn
8. Cloning of two junction fragments from der(1) and der(11) chromosomes from t(1;11)(p13;p11.2)dn
9. Mutation screening of *PHF21A* in 200 patients
10. Manuscript writing and planning the experiments with other collaborators

### 3. ASHG Abstract

The American Society of Human Genetics Conference 2007 Poster

Kim HG, Norris K, Kulharya AS, Layman LC

#### ***The molecular analysis of apparently balanced chromosome translocations in two unrelated patients with hypogonadotropic hypogonadism***

Patients with idiopathic hypogonadotropic hypogonadism (IHH) present with absent puberty due to a hypothalamic-pituitary defect, and may be either normosmic (nIHH) or anosmic (Kallmann syndrome). Although mutations in genes such as FGFR1, KAL1, and GNRHR constitute the most commonly encountered etiology in IHH, the molecular basis for most patients remains unknown. Apparently balanced chromosomal rearrangements found in some patients may actually disrupt a gene at the breakpoint, thereby aiding in identification of the causative gene. We have characterized translocations in two unrelated IHH patients—one a 46,XY,t(10;12)(q26.12;q13.11) in a male with Kallmann syndrome, and the other a mos46,XY,t(3;12)(p13;p13)[18]/46,XY[3] in a male with normosmic IHH and cerebellar ataxia. The 10q26 breakpoint has been reported previously in a Kallmann syndrome patient with monosomy 10q26 and in several cases of translocations involving urogenital anomalies and hypogonadism. In our patient, the 10q26 translocation breakpoint was narrowed to 1.5 Mb region between RP11-717L13 and RP11-753P11 and we identified a breakpoint spanning BAC RP11-464D5 at 12q13.1.

In the second patient with the mosaic 3;12 translocation, homogeneous lymphoblastoid cell lines with the balanced translocation were successfully transformed from peripheral white blood cells. IHH and cerebellar ataxia often occur together and they are seen in Gordon-Holms syndrome and Boucher-Neuhauser syndrome. We hypothesize that one of the breakpoints of this translocation case is likely to harbor a gene responsible for this phenotype. A positional cloning technique was applied to clone each of the breakpoints. FISH mapping for a breakpoint at 12q13 in this patient is underway. These chromosome translocations afford the potential to define additional genes involved in IHH/Kallmann syndrome.

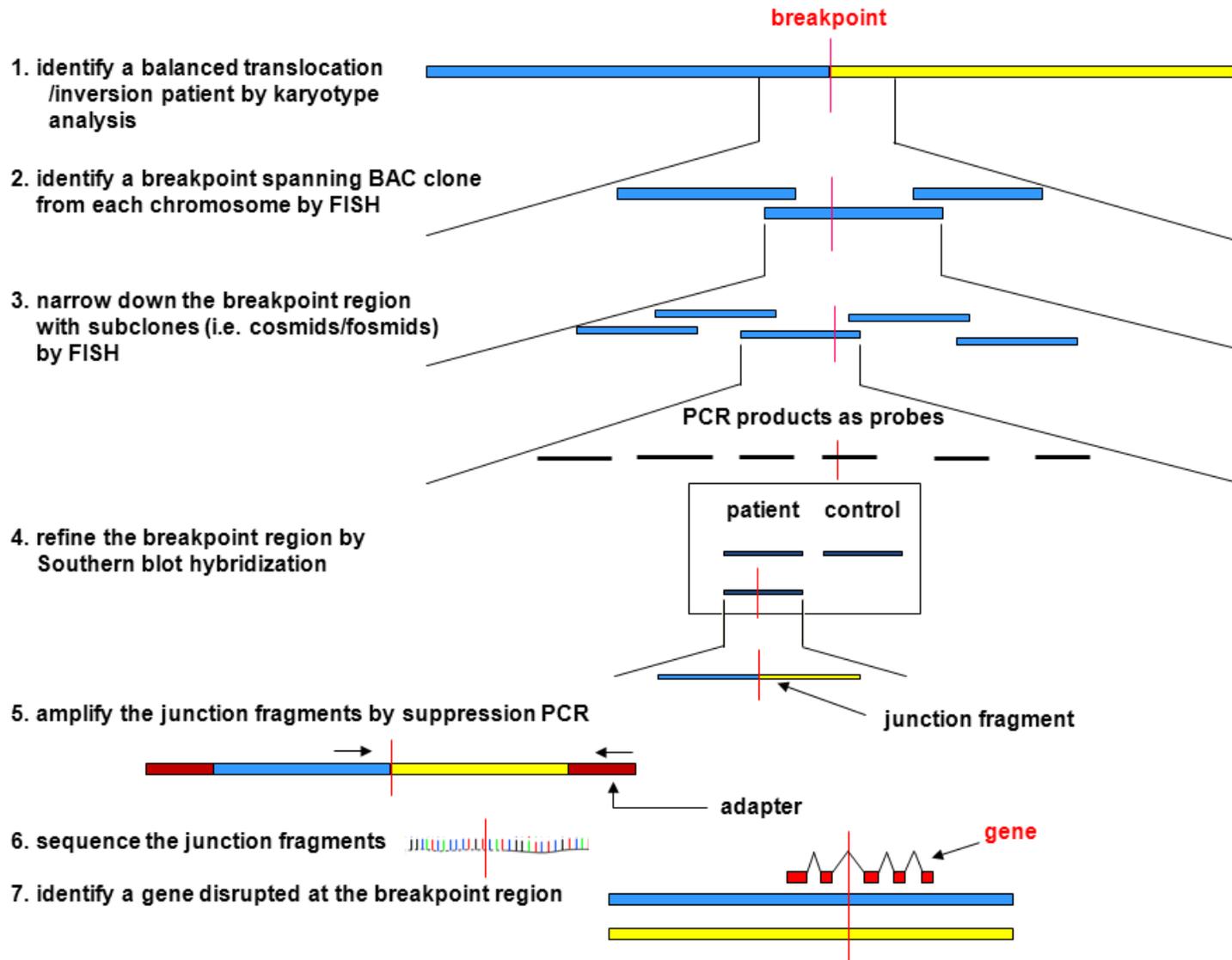
## 4. Introduction

### 4.1 Balanced translocations: visible bridges between genotypes and phenotypes

An ongoing goal of the Human Genome Project, following successful production of largely continuous genomic sequence for each of the 24 human chromosomes, is to identify all human genes and to elucidate their functions.<sup>1</sup> Of particular interest is the detection of disease genes and genetic sequence variations associated with specific human phenotypes. Although considerable progress has been made, the majority of genes causing monogenic and oligogenic disorders have yet to be identified. Discovery of genes underlying human developmental disorders has largely been accomplished through a candidate gene approach or positional cloning, which relies mainly on genetic linkage analysis. In studies of hereditary diseases, however, the resolution of linkage mapping is limited by the number of available meioses and informative marker density.<sup>2</sup> Furthermore, this approach relies on the availability of sufficiently large families. In rare disorders, this can be particularly challenging as affected individuals often have reduced genetic fitness due to their physical and mental handicaps. Moreover, affected persons may be deceased at a relatively young age. Thus, the paucity of familial cases has hampered traditional linkage mapping and the elucidation of genetic disorders.

Therefore, molecular cloning of chromosome translocation or inversion breakpoints in isolated patients offers an opportunity to identify novel disease genes that are disrupted or dysregulated by the chromosome rearrangement. Even a single balanced chromosome translocation or inversion has the potential to lead to the localization and cloning of the disease-causing gene through molecular analysis of the breakpoints (Figure 1).

Linkage analysis enables identification of a causative pathogenic mutation provided the linkage interval is small enough to sequence all relevant candidate genes. Additional patients with the same phenotype are therefore not required because the family used for linkage analysis should be sufficient to identify the mutation. In contrast, the identification of positional candidate genes from single patients with balanced chromosome rearrangements will mostly require subsequent mutation screening of unrelated patients to validate these findings and to establish that mutations in the gene cause the disease.



*Figure 1. Positional cloning using chromosome rearrangement.*

*After identification of a patient with balanced chromosome rearrangement the mapping of the breakpoint is initiated to identify the causative gene. FISH with BACs chosen from the cytogenetic bands of the karyotype will identify a breakpoint spanning clone showing three hybridization signals on metaphase chromosomes. Subsequent FISH with cosmids/fosmids contained in the breakpoint spanning BAC will narrow down the breakpoint region to less than 40 kb, which will be further refined to several kb by Southern blot hybridization. Using primers from the refined region and adaptor sequence in a suppression PCR, the junction fragments, consisting of the sequences of both chromosomes, will be cloned. A disease gene disrupted or dysregulated by a breakpoint will be identified.*

## **4.2 Structural chromosome rearrangements**

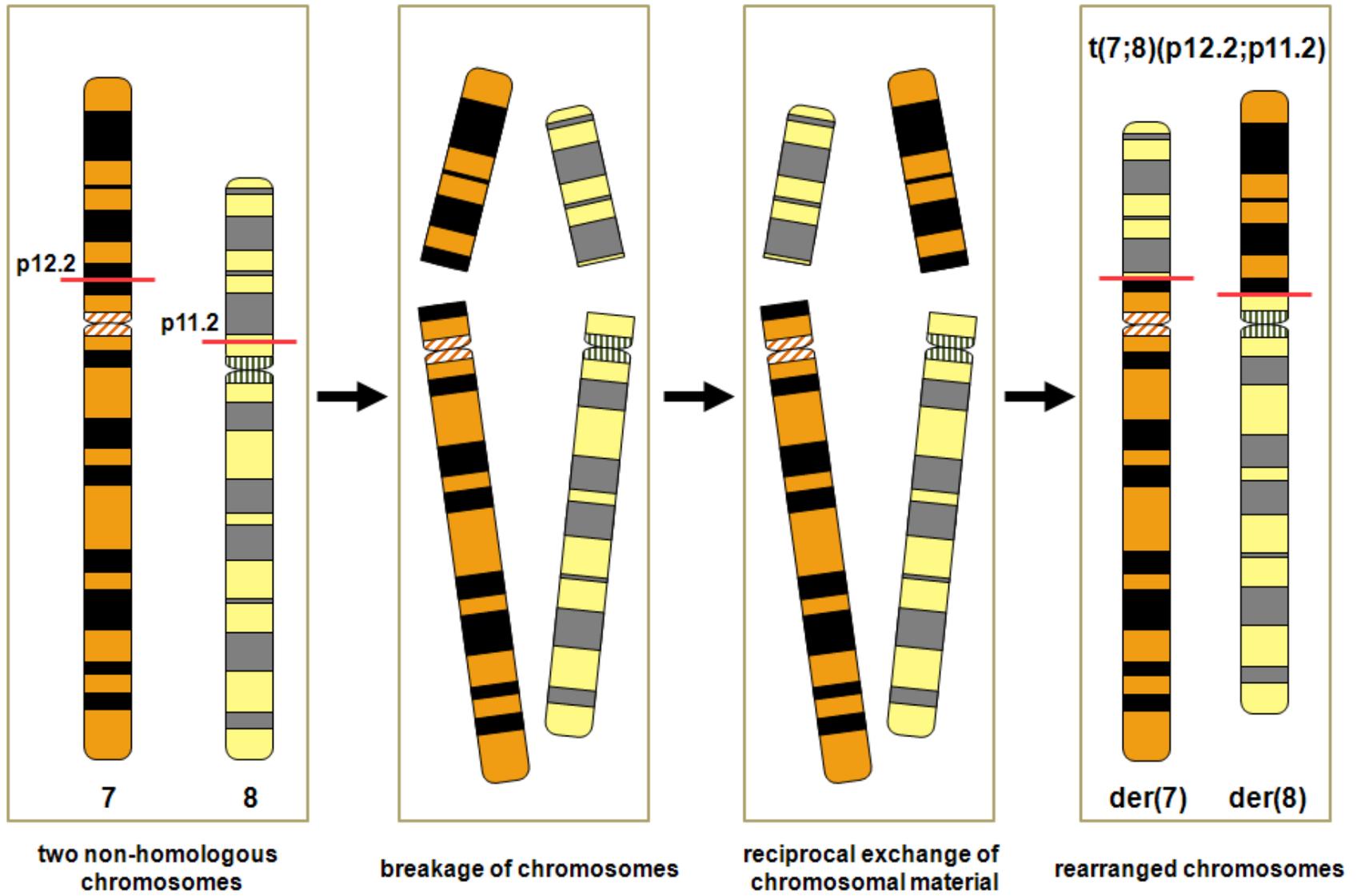
Human chromosomes can be distinguished because of their characteristic size, morphology and banding patterns, which serve as landmarks enabling the unambiguous identification of chromosome segments.<sup>3</sup> Compared with numerical chromosome aberrations, cytogenetically detectable balanced chromosomal structural changes are relatively rare, although theoretically there are countless ways in which chromosomes could restructure themselves from the normal chromosomes. Because chromosome breakage can theoretically occur anywhere within the human genome, the number of potential rearrangements that can result is immense.<sup>4</sup> Consequently most chromosome rearrangements are unique involving different cytogenetic bands and different molecular breakpoints even for a seemingly identical cytogenetic breakpoint.

Chromosome breaks sometimes occur spontaneously, or they may be induced by a mutagenic agent or clastogen such as ionizing radiation or DNA-damaging chemicals. Unlike normal chromosomal ends, broken ends tend to rejoin. Usually the broken ends repair at the same place where the break occurred, healing the break. However, a break may lead to structural rearrangements of chromosomes if the exchange occurs between different chromosomes.<sup>5</sup> In a reciprocal chromosome translocation, a two-way exchange of chromosomal material takes place between two non-homologous chromosomes. A break occurs in one arm of each chromosome and the portions of

chromosome material distal to the breakpoints switch positions.<sup>6</sup> A broken end of one chromosome attaches to a broken end of another chromosome. The rearranged chromosome is called a derivative chromosome and it is identified by the specific chromosomal centromere it possesses (Figure 2). The main forms of chromosomal rearrangements encountered in humans include translocation, inversion (Figure 3), insertion, deletion, duplication and ring formation. A chromosome translocation, the result of an exchange of chromatin between chromosomes following breakage, can give rise to two new chromosomes, which may function normally if each possesses a single centromere; chromosomes with more than one functional centromere rarely survive more than one mitosis. The two new products of such an exchange, which are called derivative chromosomes, may appear physically different from the normal homologues. However, if the amount of chromatin material exchanged is equivalent in size and structure, the translocation might be difficult to detect cytologically.<sup>7</sup>

All structural chromosome rearrangements necessitate at least two chromosomal breaks and subsequent aberrant reunion of the broken ends. These breaks and reunions may occur both within or in the vicinity of developmentally important genes. If genes critical in human development are interrupted or dysregulated by the chromosomal rearrangement and the amount of protein encoded from the corresponding intact gene is insufficient, individuals carrying the translocation may exhibit developmental disorders.

Mis-repair of DNA double-strand breaks (DSBs) can result in formation of chromosome rearrangement through a variety of different pathways. The two primary pathways are conservative homologous recombination (HR) and non-homologous end joining (NHEJ). In HR, broken ends of DNA can be aligned and joined using sequence homology, whereas in NHEJ broken ends are brought together and rejoined in the absence of tracks of sequence homology. In the third pathway through single-strand annealing (SSA), repair of DSBs occurs between two repeated sequences oriented in the same direction near a DSB. In the recombination between Alu elements in mammalian cells, Elliot *et al.*<sup>8</sup> demonstrated that NHEJ predominates as the chromosomal translocation mechanism when heterology is present at the DNA ends. When identical Alu elements are present at DNA ends, SSA predominates as the translocation mechanism.



*Figure 2. The genesis of a balanced chromosome translocation.*

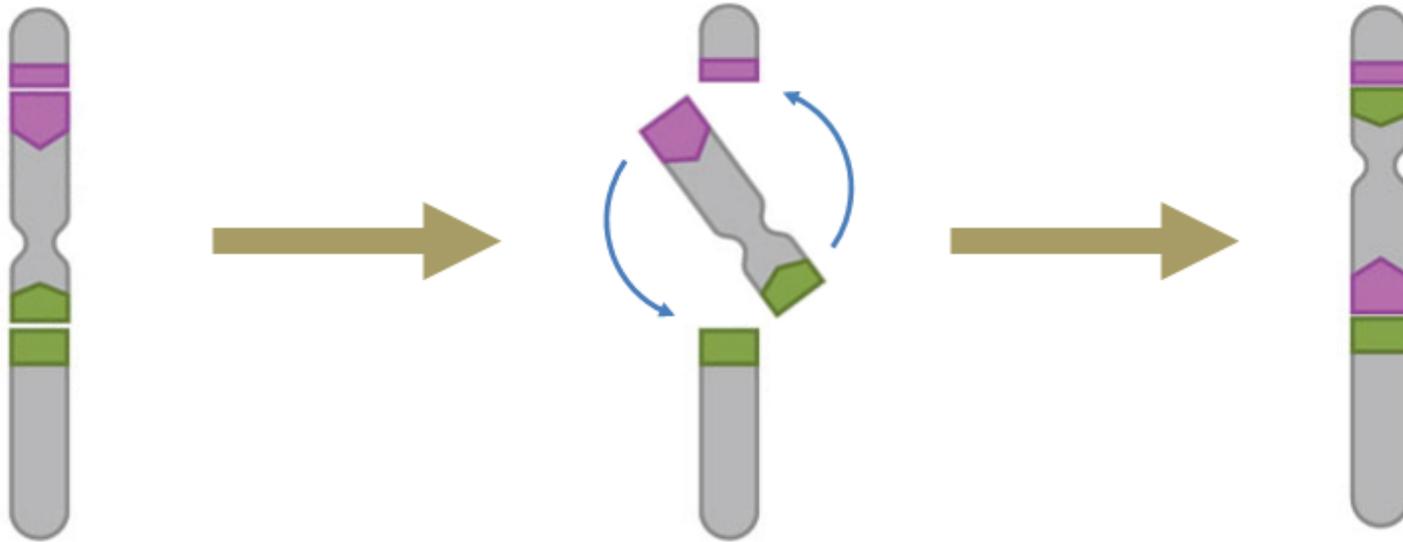
*Double-strand breaks at 7p12.2 and 8p11.2 in two non-homologous chromosomes are followed by a reciprocal exchange of chromosomal segments between partner chromosomes 7 and 8. This inappropriate repair leads to the generation of a balanced chromosome translocation  $t(7;8)(p12.2;p11.2)$ , resulting in two new derivative chromosomes 7 and 8.*

A homology-independent process may characterize the majority of chromosome translocations, and NHEJ is proposed as the possible mechanism involved.<sup>9</sup> Mapping studies of many constitutional  $t(11;22)$  translocations from unrelated families have demonstrated that the reciprocal exchange occurs between similar AT-rich palindromic repeat sequences at 11q23 and 22q11, which contain two inverted regions complementary to each other.<sup>8,9</sup>

#### **4.3 A historical perspective: balanced translocations and positional cloning**

The first description of a gene disruption at the chromosome breakpoint might be traced to a statement, “At the position where the break originally occurred, one or more genes may be damaged or altered and this may give rise to one or more mutations”. This sentence referred to chromosome breaks by ionizing radiation and was published in 1959 in a book entitled “Human Heredity” by Ashley Montague (The World Publishing Company page 321).

In 1972 W. Roy Breg and Dorothy A. Miller reported the detection of three independent mentally retarded females with facial dysmorphism and three phenotypically normal individuals in six balanced autosomal chromosome translocations using the quinacrine fluorescence chromosomal banding method. To explain the developmental abnormalities in the three cases they presumed either that 1) the translocations might be actually unbalanced, or that 2) the function of the translocated region might be altered.<sup>10</sup>



**two breakages  
in a chromosome**

**inversion of  
chromosome material**

**reinserted  
chromosome**

*Figure 3. The genesis of a pericentric balanced chromosome inversion.*

*A pericentric inversion occurs when a single chromosome undergoes double-strand breakage at two different locations of p and q arms followed by a rearrangement, in which the middle segment of a chromosome including a centromere is inverted 180 degrees in orientation before the breaks are aberrantly healed.*

Therefore, they appreciated the importance of balanced translocations disrupting developmental genes at the breakpoints. The postulation of Patricia A. Jacobs in 1974 that cytogenetically heterozygous apparently balanced chromosome translocations are associated with phenotypic abnormalities in particular in intellectual disability ushered in new era of positional cloning of balanced chromosome translocations with developmental disorders, since she explained the underlying mechanisms of deleterious effect as the following.<sup>11</sup>

- 1) The rearrangement is in reality unbalanced and carries a small deletion below the level of cytological detectability;
- 2) Chromosome breakage and exchange results in a gene mutation at one or both of the breakpoints; and
- 3) Certain rearrangements are associated with a position effect detrimental to their carriers.

Constitutional balanced translocations have facilitated positional cloning of a gene that is disrupted or mutated in Duchenne muscular dystrophy. A number of females carrying a balanced translocation who were affected with this genetic disorder provided an opportunity to clone the disease gene in 1987.<sup>12,13</sup> Since the breakpoints of these carrier females always occurred within chromosome band Xp21, the location of the breakpoints helped to identify the disease gene in combination with linkage and deletion data. This positional cloning approach using balanced translocations is historically important.<sup>14,15</sup> Numerous genes with critical functions in either developmental processes, organogenesis, oncogenesis, or normal metabolism have been identified by positional cloning following discovery of chromosomal rearrangements in affected patients.<sup>16</sup>

These structural chromosomal abnormalities can lead to disruption or dysregulation of proper gene function, which is critical in human developmental pathways.<sup>17</sup>

#### **4.4 Balanced translocations associated with abnormal phenotypes**

Chromosomal rearrangements such as translocations, inversions, deletions, insertions, and duplications play a critical role in disease gene discovery. Apparently balanced reciprocal translocations are common types of chromosomal rearrangements found in patients with an abnormal phenotype and even more frequently in clinically unaffected individuals.<sup>18</sup> Since most balanced rearrangements follow Mendelian inheritance patterns, naturally occurring human pathogenic mutations in the form of chromosome rearrangements are valuable biological tools for identification of monogenic disease genes.

Approximately 1 in 2000 embryos has a *de novo* balanced chromosomal translocation and 6.1 % of these translocation carriers have congenital anomalies, which include isolated defects ranging from intellectual disability, cleft palate/lip, abdominal wall defects, limb defects, cardiac abnormalities, genitourinary abnormalities, or they can occur as part of known clinical syndromes.<sup>19</sup> Given the birth rate in the U.S., it may be estimated that approximately 150 infants are born each year carrying *de novo*, balanced, reciprocal translocations or inversions in addition to congenital anomalies.

The overwhelming majority of individuals with balanced reciprocal translocations are phenotypically normal, although at higher risk of reproductive failure due to unbalanced segregations.<sup>20</sup> It is generally assumed that deleterious effects ensue rarely in most balanced chromosomal translocations, in which all the genetic information required by the cell is present even if it is in an abnormal position, because genes as potential targets of breakpoints comprise only a fraction of the human genome and a single copy of many genes is sufficient. It was reported that up to 1% of genes on the X chromosome in males can sustain loss-of-function variants with no apparent detrimental effect.<sup>21</sup> The number of genes with potential haploinsufficient effects in autosomes is expected to be larger, given that only one of two alleles is affected.

There are four possible reasons for a balanced chromosome translocation carrier to be apparently healthy: 1) both breakpoints do not disrupt or dysregulate any gene, 2)

heterozygous disruption or dysregulation of a gene at one breakpoint has no phenotypic consequence because for the respective proteins, 50% of the normal concentration is functionally sufficient, 3) the carrier of a balanced translocation actually has a subtle phenotype which is elusive and hard to detect, especially if it is a behavioral and/or psychiatric phenotype, and 4) the gene affected is maternally or paternally imprinted. It is possible that a gene disrupted or dysregulated at a breakpoint would be imprinted based on its parental origin, which would render the carrier either healthy or affected. This has important implications in familial balanced translocation cases with discordant phenotypes, , since parental imprinting might be more prevalent than expected.<sup>22</sup>

If an apparently balanced chromosome rearrangement is associated with a pathogenic phenotype five explanations are possible: 1) the finding is coincidental and there is no relationship between the chromosome translocation and phenotype (phenotype is translocation-unrelated and likely to be caused by a mutation of a gene somewhere else in a genome); 2) the rearrangement is in fact unbalanced with cryptic loss or gain of the genome; 3) a gene at one or more breakpoints is disrupted or dysregulated by a positional effect and causative; 4) coincidental linkage of a mutation of a gene near the breakpoint, and 5) two different genes with the same transcription direction at each breakpoint are truncated by a translocation or inversion to produce fusion genes exerting dominant-negative functions and sometimes endowing the fused transcripts with tumorigenic properties. Loss of function can be caused by gene truncation or positional effect, whereas gain-of-function can emerge from the generation of an aberrant fusion gene<sup>23</sup> or truncated protein.

Genomes with rearrangements may be balanced (whereby genetic material is not lost or gained) or unbalanced, when the chromosomal breakpoints or genomic region elsewhere exhibits additional complexity when studied in detail; copy number variations (CNV) like duplications or deletions or insertions may occur. In the case of unbalanced translocations, large chromosomal regions (usually detectable by traditional metaphase karyotyping), submicroscopic or even cryptic regions (detectable by high-resolution CGH arrays or molecular analysis) may be lost or gained. Although unbalanced translocations may be helpful in mapping genes to disease phenotypes, a large number

of genes may reside within the regions of CNVs, making identification of causative genes difficult. These unbalanced rearrangements, however, can assign gene loci with unique phenotypes in a specific chromosomal region, thereby helping to identify a disease gene when a balanced translocation or inversion involving this region and a similar phenotype are found.<sup>24,25</sup> These genomic imbalances, which are mostly in the form of deletions or duplications, are either at one of the translocation breakpoints<sup>26</sup> or unrelated to the breakpoint regions,<sup>27,28</sup> suggesting that the phenotypic effect is likely due to the gene deletions or duplications, rather than the disruption or dysregulation of a gene at or near the breakpoint. The causative gene(s) is mostly located in the deleted<sup>29</sup> or duplicated region,<sup>26</sup> but sometimes it is present outside of the deletion region, suggesting a positional effect.<sup>30</sup> Genomic deletions at the breakpoint appear to be more common than duplications, which is likely to be the result of under-reporting of duplications due to technical difficulties of detection.<sup>26</sup>

The majority of autosomal constitutional balanced chromosome rearrangements are heterozygous and the causative gene for supposedly dominant inherited disease lies at or near one breakpoint. However, in rare cases two genes from each breakpoint region explain the complex phenotype of the translocation carrier.<sup>31,32</sup> Albeit a rarity, two autosomal genes truncated by both breakpoints in a female chromosome translocation patient have been reported to be potential positional candidate genes for intellectual disability (ID) and dysmorphic features.<sup>33</sup> This was due to the identification of several rare heterozygous non-synonymous missense nucleotide variants in both genes in a cohort of ID patients with unknown etiology suggesting that alterations of either gene are likely to be causative.

Dissecting the contribution of seemingly benign balanced chromosomal rearrangements to disease susceptibility in monogenic and complex genetic diseases as well as human diversity is a gargantuan goal of human genetics but it requires more knowledge of SNP involvement in genetic and phenotypic variation. This could be achieved by studying the frequency of the rearrangements in a large population with registered phenotypes and the establishment of reference databases. Some chromosomal structural variants might

predispose to the genetic disease only when combined with other genetic variants or SNPs.

#### 4.5 Balanced translocations in recessive disorders

Translocations, like inversions, may exist in either a homozygous or a heterozygous state.<sup>3</sup> Rare homozygous balanced chromosome translocations can be used to identify genes involved in recessive disease. In all known cases the homozygous translocation carriers were the offspring of consanguineous unaffected parents with heterozygous translocations and the breakpoint truncated the disease gene or caused a positional effect in both alleles.<sup>34-39</sup>

To become clinically manifest, a balanced heterozygous rearrangement disrupting a recessive locus would require the structurally normal chromosome to carry a mutant allele for the recessive phenotype.<sup>40</sup> A unique heterozygous chromosome translocation or inversion unmasking recessive compound heterozygous mutations in an underlying gene has been reported in a patient of Alström syndrome.<sup>41</sup> This patient had one copy of the gene disrupted by the translocation and the other copy disrupted by an intragenic mutation, suggesting that an unknown clastogen might have induced both chromosome breaks and a gene mutation. In another apparently balanced translocation - 46,XX,t(2;8)(p16.3or21;p23.1)mat – a 163 kb deletion involving exons 9 and 10 of *FSHR* was detected at the breakpoint of 2p16.3 and ensuing *FSHR* sequence analysis in this translocation patient revealed a novel missense point mutation p.Pro587His. These findings suggested that compound heterozygous mutations of *FSHR* are the cause of primary amenorrhea and incomplete pubertal development combined with hypergonadotropic hypogonadism in this patient.<sup>42</sup>

#### 4.6 Positional effect

Some studies show that it may not always be the gene truncated by a breakpoint, but rather a neighboring gene that is causative for the phenotype seen in a chromosome translocation carrier.<sup>43-46</sup> This suggests a positional effect, possibly mediated by genomic separation of genes from their *cis*-acting regulatory elements resulting in alteration of expression level, seen roughly in 10% of positional cloning studies.<sup>47-50</sup>

This position effect is unique in that it does not disrupt gene function, but instead

disturbs its regulation. The longest-range position effect found in the field of human genetics was 17q breakpoint mapping ~1.3 Mb downstream of *SOX9* in a fetus with a *de novo* complex balanced translocation and campomelic dysplasia.<sup>51</sup> This position effect represents a significant hindrance to positional cloning research since the directly disrupted gene is not the culprit gene for the phenotype. Generally the disease phenotypes produced by position effects are similar to those produced by loss-of-function mutations within the gene, suggesting that these position effect mutations result from decreased gene expression.<sup>52</sup> Therefore, the gene knockdown or knockout animal model would likely recapitulate the human phenotype from this position effect.

Rarely, position effects of two different genes from each breakpoint of a balanced translocation are the cause of complex phenotype of the carrier.<sup>31</sup> Furthermore, the positional effect could exert an effect upon more than one gene in the vicinity of the breakpoint in the affected region with high gene density. One of the typical regions containing multiple disease genes in tandem is chromosome 2q31.1 where nine *HOXD* genes, in the order *HOXD13*, *HOXD12*, *HOXD11*, *HOXD10*, *HOXD9*, *HOXD8*, *HOXD4*, *HOXD3*, *HOXD1* from centromere to telomere direction, are clustered on the 98 kb of genomic sequence.

The three different breakpoints located either centromeric or telomeric to this gene cluster have been reported, causing diverse limb malformations in three unrelated patients with chromosomal rearrangements. A positional effect resulting in the misregulation of *HOXD* expression is likely the underlying mechanism for each case due to separation of different regulatory elements from this gene cluster.<sup>53</sup>

Another region is in Xp11.22 where *KDM5C* (formerly *JARD11C*) for XLID,<sup>54</sup> *IQSEC2* for nonsyndromic XLID,<sup>55</sup> *SMC1A* for X-linked Cornelia de Lange syndrome,<sup>56</sup> *HUWE1* for nonsyndromic XLID,<sup>57</sup> and *PHF8* for XLID<sup>58</sup> cluster together. If there is a breakpoint within or outside of this gene cluster, it is not easy to determine which single gene or genes is/are ascribed to the phenotype of a patient with a chromosome rearrangement. In this case it is desirable to measure expression pattern of each gene by qRT-PCR and western blot analysis to find which candidate gene is dysregulated by a positional effect. If more than one gene is affected by a position effect, a gradient has been observed with genes closest to the translocation breakpoint displaying the most pronounced outcome.<sup>52</sup>

#### **4.7 Balanced translocation as a “double hit”**

It is difficult to assess whether a gene disruption at a breakpoint is causal for the phenotype or just coincidental, especially if only a single translocation patient is reported to have the breakpoint. Therefore, *de novo* translocations, familial translocations segregating with the disease in a family,<sup>59-61</sup> or “double hit”<sup>60</sup> cases, in which independent breakpoints in two individuals with related phenotypes occur in the same chromosomal region,<sup>60,62,63</sup> take priority for positional cloning to maximize the likelihood that the chromosomal disruption is causative.

If two independent cases of balanced translocations or inversions with different karyotypes share cytogenetic breakpoint regions and have a similar phenotype, it is likely that the breakpoints affect the same disease gene. These so-called “double-hit” cases provide unique opportunities for identifying the causative genes. The association of two or more cases with a common phenotype and common cytogenetic breakpoint is a strong indication that a causative gene is located at the common breakpoint.<sup>64</sup> However, it is impossible to locate molecular breakpoints by conventional karyotype analysis alone, and seemingly identical breakpoints may turn out to be different.<sup>24</sup> Therefore, in order to find “double hit” cases, patients with breakpoints in neighboring cytogenetic bands still warrant consideration. Mutant mice or zebrafish knockdowns with similar human phenotypes mapping to the syntenic region of cytogenetic breakpoint of a human chromosome translocation also provide important supportive evidence.

If the cytogenetic breakpoint of a balanced chromosome translocation or inversion is compatible with linkage results from other affected families,<sup>65-67</sup> this could be also considered as a priority. Balanced chromosome rearrangements sometimes facilitate assignment of phenotype/genotype relationships in contiguous gene deletion syndromes.<sup>68-76</sup> At times, the causative gene has been found in another deleted region in the same translocation patient independent of the breakpoint region.<sup>77</sup>

#### **4.8 Balanced translocations with discordant phenotype**

The risk of the abnormal phenotype is always higher for an individual with an apparently balanced *de novo* rearrangement than for an individual who has inherited a similar rearrangement from a normal parent.<sup>4</sup> Sometimes the balanced heterozygous translocation carrier in the family has a discordant phenotype—one is asymptomatic and

the other has a phenotype. Even if a *de novo* balanced chromosome translocation or an inversion strongly suggests that a chromosomal rearrangement is the cause of symptoms seen in the carrier, this rule may not apply for non-Mendelian disorders. In some autosomal human genes, only the paternal or maternal copy of the gene is expressed. Consequently, for a gene that is normally imprinted with paternal silencing, a disruption in the maternal copy of the gene by a breakpoint will result in disease, whereas the same disruption in the paternal copy will have no effect. This can occur in a familial balanced translocation, in which one translocation carrier is affected and another is unaffected. This discordant phenotype might be due to additional cryptic CNVs, uniparental disomy, or unmasking of a recessive allele on the normal chromosome, which does not exist in an asymptomatic carrier in the same family. Even if the same balanced rearrangement of a healthy carrier parent is detected in a fetus at the time of prenatal diagnosis, it is not possible to exclude completely an increased risk for a phenotypic abnormality in the offspring.

For example, regarding an imprinted gene with paternal silencing, a balanced chromosome rearrangement disrupting a maternal gene will result in disease, whereas the same inversion from the paternal copy will have no effect. The Angelman syndrome gene *UBE3A* was identified by cloning the breakpoint at 15q11.2 from a child with Angelman syndrome and a balanced inversion, who inherited the same inversion from a healthy mother. The authors hypothesized that the inversion may inactivate a maternally active gene at the breakpoint, and that the lack of a functional maternal allele may lead to Angelman syndrome.<sup>78</sup> In this case one inversion carrier is healthy, whereas another inversion carrier in the same family is phenotypically affected. Thus, familial balanced chromosome rearrangements with discordant phenotypes may sometimes offer an excellent opportunity to clone paternally or maternally imprinted disease genes. In Prader-Willi syndrome, the translocation breakpoint interrupting the disease gene on the paternal copy causes the disease.<sup>79</sup>

#### 4.9 Balanced translocations involving the X chromosome

Balanced translocations involving the X chromosome have a unique place in positional cloning history. With the very first positional cloning technique started in 1976 using a female balanced translocation patient, t(X;21)(p21;p12) with Duchenne muscular dystrophy, the causative gene, DMD, was discovered at Xp21.1 in 1987.<sup>15,80,81</sup> The use of balanced translocations to identify disease genes, such as DMD, has followed a timeline similar to the use of genetic linkage. For example, although the Huntington disease gene was linked to chromosome 4 using DNA markers for the first time in 1983<sup>82</sup>, the gene was finally identified in 1993.<sup>83</sup> The use of balanced translocations to understand basic cellular mechanisms and heredity has also been a focus of study. In half a dozen independent cases of X;autosome balanced translocation cases with Duchenne muscular dystrophy, the translocated X chromosome was of paternal origin.<sup>84,85</sup> The underlying mechanism is unknown. The parental origin of the translocated X chromosomes, as well as the autosomes in other X;autosome translocations remain to be elucidated.

Constitutional balanced X;autosome translocations in males are very rare. However, since *de novo* X;autosome translocations are usually paternal in origin,<sup>86,87</sup> male parents will transmit this X;autosome translocation to their daughters.<sup>6</sup> Thus most of the carriers of these anomalies are females. X inactivation in somatic cells of females is usually random, but this is not true in females with balanced X;autosomal translocations. In the majority of cells of such individuals, the translocated X chromosome remains active, whereas the normal X chromosome is inactive.<sup>88</sup> It generally is assumed that the nonrandom X inactivation observed in such cases is the result of a selection process operating against cells in which X inactivation has spread into the autosome and has inactivated autosomal genes.<sup>89</sup> If the derivative chromosome with the translocated X chromosome was erroneously inactivated in some cells, then a portion of the autosome would also likely be inactivated, causing functional autosomal monosomy, as well as functional X chromosome disomy resulting from the portion of X chromosome translocated onto the autosome.<sup>90</sup>

Before mapping the breakpoint on chromosome X of a balanced X;autosome translocation from an affected female, it is important to determine the X-inactivation pattern. Skewed X inactivation would be expected with the normal X chromosome preferentially inactivated so that the disruption of a gene at the X;autosome translocation breakpoint is responsible for X-linked disorders. In several cases of X;autosome balanced translocations in females, a small percentage of cells have been observed in which the normal X is early replicating and therefore, presumably active.<sup>13</sup>

Most females who carry a balanced X;autosome translocation show a skewed X inactivation pattern, with 100% of cells having the same normal X chromosome selectively inactivated.<sup>91,92</sup> This erroneous exclusive inactivation of the normal X chromosome leaves the carrier of a balanced X;autosome translocation suffering from X-linked disorders. Since one copy of the disease gene is inactivated on the normal chromosome X and another copy of the same disease gene is interrupted at the breakpoint in the derivative chromosome X, this leads to a functional nullisomy in this female translocation carrier. This implies that the causative gene for the disorder is more likely affected by the chromosome X breakpoint than by the autosomal breakpoint. In X-linked recessive diseases, males typically are affected and females are usually carriers. However, a female with a balanced X;autosome translocation with skewed X inactivation is also affected if the causative gene is disrupted at the X breakpoint<sup>80,93</sup> (examples include Kallmann syndrome and Duchenne muscular dystrophy).

It is apparent from evidence derived from both human and mouse translocations that autosomal segments tend to become inactivated when in close physical contact with inactivated X material. This is one of the reasons why in female balanced translocation carriers, it is usually the normal X which is preferentially inactivated.<sup>94</sup> Another reason is that due to the interruption of the physical continuity of the X chromosome by the reciprocal translocation, only one of the two fragments of the rearranged X chromosome, which is physically linked to an X-inactivation center (XIC) at Xq13.2, will be guaranteed for inactivation. If one translocated chromosome bears the active XIC, which is necessary and sufficient for chromosome inactivation, inactivation would spread into the adjacent autosomal region, causing functional monosomy of partial autosomal genes. On the other derivative chromosome without inactivation center, functional disomy of X chromosomal genes ensues for the loci involved. This means that all cells with an

inactive XIC on the normal X chromosome, there will be a functional imbalance, which the cell is likely to select against to improve survival and function.

## 5. Kallmann syndrome

### 5.1 Introduction

Human puberty is a dynamic process that initiates complex interactions of the hypothalamic-pituitary-gonadal axis, the purpose of which is to produce sex steroids for reproductive maturity and gametes for fertility. Any disruption of the development or regulation of this system, for which the hypothalamus serves as the master regulator through its pulsatile release of gonadotropin releasing hormone (GnRH), can produce deleterious consequences for successful reproduction.<sup>95</sup>

Idiopathic hypogonadotropic hypogonadism (IHH, MIM 146110), one of the most commonly inherited forms of hypogonadism, results from deficient hypothalamic GnRH release or action.<sup>96</sup> Patients with IHH show clinical signs and symptoms of GnRH deficiency: delayed puberty due to low sex steroid production along with low levels of serum gonadotropins. IHH patients are usually discovered during adolescence or adulthood present with absent or impaired sexual development due to sex-steroid-hormone deficiency, low serum levels of the pituitary gonadotropins follicle-stimulating hormone (FSH) and luteinizing hormone (LH), and infertility.<sup>96</sup> Individuals affected by hypogonadism have external genitalia, genital ducts, and gonads of the appropriate sex but do not function normally, particularly with respect to secondary sexual characteristics such as absent to sparse sexual hair. In females breast development may be absent or greatly retarded and they usually have primary amenorrhea. In males the voice may not deepen and the physique fails to develop along muscular lines because of deficient testosterone. The adolescent male who is obese and slow to mature is a familiar example.<sup>97</sup> Males may have small testes, micropenis (hypogenitalism, microphallus) and gynecomastia.

Kallmann syndrome is a human developmental disorder characterized by anosmia resulting from agenesis of the olfactory bulbs and hypogonadism secondary to deficiency of hypothalamic GnRH. In addition, cleft lip and palate, unilateral renal agenesis, dental agenesis, and neurologic abnormalities such as synkinesia and

cerebellar dysfunction have been described.<sup>95</sup> Its prevalence has been estimated at 1/10,000 in males and 1/50,000 in females. In a minority of cases there are inactivating mutations of *KAL1*, an X linked gene encoding a putative adhesion molecule thought to mediate embryonic neuronal migration.<sup>98,99</sup> The identification of additional disease genes involved in IHH/KS by linkage analysis has been hampered by the lack of a sufficient number of large families since most affected individuals do not have affected relatives and/or they do not seek treatment for their inherent reproductive failure.<sup>96</sup> Thus, the traditional positional cloning approach using linkage analysis has been substituted by positional cloning of breakpoints in rare IHH/KS patients harboring chromosomal translocations<sup>98,100</sup> or deletions.<sup>101,102</sup> Many constitutional chromosomal rearrangements involving IHH and KS have been reported, but the disrupted genes have not been identified.<sup>25,95,103,104</sup> In this project we are using both positional cloning and candidate gene approach as described below to identify new genes involved in IHH and KS.

- 1) A balanced translocation t(7;8)(p12.3;p11.2)dn from a male subject with IHH as well as cleft lip and palate to define the chromosomal breakpoint on chromosome 8 and identify a positional candidate gene at 8p11.2.
- 2) A candidate gene approach with *CHD7* as a potential causative gene for IHH and KS based on overlapping clinical features between KS and CHARGE (coloboma, heart anomaly, choanal atresia, retardation, genital and ear anomalies) syndrome. *CHD7* will be screened in more than 100 patients of IHH and KS to identify causative mutations to validate this gene for IHH and KS.
- 3) A balanced translocation t(10;12)(q26.12;q13.11)dn from a male subject with KS to define the chromosomal breakpoint on chromosome 10 and scan positional candidate genes in its vicinity at 10q26.12 to identify a disease gene. Since currently there are no known causative genes for IHH and KS at the 10q26 chromosomal region this study has a potential to identify a new gene involved in IHH and KS.

## 5.2 Genetics

IHH and KS are genetically heterogeneous. Identification of genetic abnormalities related to IHH has provided major insights into the pathways critical for the development, maturation, and function of the reproductive axis.<sup>105</sup> X-linked (*KAL1* at Xp22.3) recessive, autosomal dominant, and recessive loci have been described. The discovery of a rare family with males displaying both X-linked KS and ichthyosis led to the identification of *KAL1* by positional cloning,<sup>98,102</sup> while the characterization of deletions<sup>101</sup> and a balanced translocation<sup>100</sup> involving chromosome 8 facilitated the cloning of *FGFR1*, associated with both IHH and KS.<sup>106</sup> Positional cloning in consanguineous autosomal-recessive IHH families revealed *KISS1R* encoding GPR54<sup>107,108</sup> *TAC3*<sup>109</sup>, and *TACR3*<sup>109</sup> and candidate-gene approaches identified mutations in *GNRHR*,<sup>110,111</sup> *NELF*,<sup>112</sup> *CHD7*,<sup>113</sup> *FGF8*,<sup>114</sup> and *GNRH1*,<sup>115,116</sup>; analogous mouse phenotypes pointed to *PROKR2*,<sup>117</sup> and *PROK2*,<sup>117</sup> as well as *KISS1R*.<sup>108</sup> The IHH phenotype associated with mutations in *GNRHR1*, *KISS1R*, *GNRH1*, *TAC3*, and *TACR3* is limited to the consequences of the resulting gonadotropin and sex steroid deficiencies. These patients rarely have associated clinical manifestations such as those seen in KS or in more complex syndromic forms of IHH such as CHARGE syndrome.<sup>105</sup>

IHH and KS have been considered as monogenic disorders with a Mendelian inheritance pattern, but recently a digenic pattern has been reported for *PROKR2/KAL1*,<sup>117</sup> *FGFR1/NELF* and *FGFR1/GNRHR1*,<sup>118</sup> *PROK2/PROKR2*,<sup>119</sup> *GNRH1/FGFR1/PROKR2* and *GNRH1/NROB1*,<sup>116</sup> *KAL1/NELF/PROK2*,<sup>120</sup> *KAL1/TACR3*,<sup>120</sup> *NELF/TACR3*,<sup>120</sup> *GNRHR1/KAL1*,<sup>120</sup> *WDR11/KAL1*,<sup>120</sup> and *WDR11/GNRHR1*.<sup>120</sup> Digenic disease is seen in 3-10% of cases.<sup>118,121</sup> To date, *KAL1* mutations cause only KS,<sup>95</sup> whereas mutations of *GNRHR*, *KISS1R* and *GNRH1* are restricted to normosmic IHH.<sup>122,123</sup> Mutations in *FGFR1*,<sup>106</sup> *CHD7*,<sup>113</sup> *FGF8*<sup>124</sup> and *WDR11*<sup>24</sup> are known to cause both normosmic IHH and KS.

Despite these significant advances in the past two decades, however, mutations in each of these genes account for a small fraction of IHH patients and the genetic etiology remains unknown for about two-thirds of all IHH and KS patients.

## 5.3 Subjects

### 5.3.1 Positional cloning of t(7;8)(p12.3;p11.2)dn

The subject is a white man who was aged 24 years at the time of initial diagnosis. He had a history of cleft lip and palate, corrected by surgery. He had no outstanding medical problems other than delayed sexual development and a feminine sounding voice. He had his growth spurt at age 18–19 years, developed sparse armpit hair at age 20, and pubic hair at 16–17, but no penile or testicular enlargement. He displayed infantile testes, child-like facial hair, sparse axillary adult appearing hair, and prepubertal chest hair. Based on the presence of cleft palate and hypogonadism, a tentative diagnosis of Kallmann syndrome was reached, though the subject did not complain of anosmia. He was prescribed a regimen of testosterone injections, which successfully induced secondary sexual characteristics. At the age of 31, he was seen by a different physician for azoospermia and infertility, and cytogenetic analysis was ordered for the possibility of Klinefelter syndrome. The analysis revealed an apparently balanced chromosomal translocation with the karyotype, 46,XY,t(7;8)(p12.3;p11.2). Informed consent for the generation of a lymphoblastoid cell line was obtained in accordance with institutional policies.

### 5.3.2 Candidate gene approach with *CHD7*

For all patients screened with no chromosomal anomalies, IHH was diagnosed in males who were 18 years old or older and had delayed puberty, testosterone levels <100 ng/dl (normal is 300–1100 ng/dl), and low or normal serum gonadotropins. In females, IHH was defined as primary amenorrhea, nearly always with absent breast development at 17 years of age and low estradiol (<30 pg/ml).<sup>96,125</sup> All patients had normal pituitary function, including normal thyroid-stimulating hormone, thyroxin, cortisol, and prolactin. No pituitary tumor was present by radiographic imaging. Complete IHH or KS is a more severe phenotype defined as the complete lack of puberty with absent breast development (Tanner 1) in females and testis size  $\leq 3$  ml bilaterally in males. Incomplete IHH or KS was defined as partial breast development in females and testis size  $\geq 4$  ml bilaterally in males.<sup>96</sup> Olfaction was either tested with the University of Pennsylvania Smell Identification Test when available or documented by history.

Lymphoblastoid cell lines were generated from patients, and DNA, RNA, and/or protein was extracted by standard methods as described previously.<sup>96</sup> All patients signed an informed consent approved by the Human Assurance Committee of the Georgia Health Sciences University.

### **5.3.3 Positional cloning of t(10;12)(q26.12;q13.11)dn**

We have screened 76 patients with IHH and KS for karyotype abnormalities and identified one male KS patient with an apparently balanced chromosome translocation revised as 46, XY, t(10;12)(q26.12;q13.11)dn.<sup>96</sup> The finding of this unique *de novo* translocation with a rare phenotype strongly suggests that the translocation is the underlying cause of this disorder. From the blood of this proband we have established a lymphoblastoid cell line. Since this balanced translocation occurred *de novo*, it is likely this structural chromosome rearrangement is the cause of KS seen in the carrier. The size of left and right testis was 2X1 cc and he had complete IHH and anosmia (KS).

## **5.4 Hypothesis**

### **5.4.1 Positional cloning of t(7;8)(p12.3;p11.2)dn**

There are a couple of reported autosomal chromosome translocation cases associated with Kallmann syndrome from the year 1990.<sup>25,103,104</sup> But no positional cloning of a autosomal gene involved with KS from the chromosome anomaly is reported to date. Chromosome rearrangements associated with human disease can tip off geneticists to the location of important genes (at or near the translocation breakpoints) whose function has been disrupted. Therefore such chromosome rearrangements provide a unique resource for applying the positional cloning strategy to identify the crucial disease genes whose functions have been disrupted by rearrangement. Since a KS patient with a microdeletion involving the same 8p11.2 region had been reported,<sup>126</sup> we sought to identify the chromosome 8 gene disrupted in this reciprocal translocation as a likely candidate for the cause of autosomal KS as well as of isolated hypogonadotropic hypogonadism. Among the candidate genes in this microdeletion region *FGFR1* appears to be particularly promising in our review based on its likely functional role in the context

of the phenotypes-CL/P and azoospermia-of our patient. *FGFR1* is a potential candidate to be involved in craniofacial development with a particular emphasis in syndromic or nonsyndromic CL/P due to its homolog *FGFR2*, an activating mutation of which causes cleft palate in Apert syndrome.<sup>127</sup> There is also evidence that the FGF/FGFR signaling pathway plays an important role in the initiation and maintenance of spermatogenesis<sup>128</sup> and *FGFR1* and *FGFR2* are in the same pathway with *TWIST* and *CBFA1*.<sup>129</sup> Therefore we set out to analyze the 8p11.2 breakpoint region in more detail by a positional cloning strategy to examine whether the translocation might have physically affected the *FGFR1* locus of this patient. While this breakpoint in *FGFR1* was being characterized, Dode *et al.* identified *FGFR1* mutations in several patients, establishing that protein truncating mutations of *FGFR1* can cause autosomal dominant KS.<sup>101</sup>

#### **5.4.2 Candidate gene approach with *CHD7***

CHARGE syndrome and KS are two distinct developmental disorders sharing overlapping features of impaired olfaction and hypogonadism. KS is a genetically heterogeneous disorder consisting of IHH and anosmia, and is most commonly due to *KAL1* or *FGFR1* mutations. CHARGE syndrome, a multisystem autosomal-dominant disorder, is caused by *CHD7* mutations. *CHD7* is expressed in the disease-associated organs of CHARGE syndrome, but also in KS-relevant tissues including the olfactory epithelium<sup>130</sup> and pituitary<sup>131</sup> in mice, as well as the olfactory nerve and bulb, hypothalamus, and pituitary in humans.<sup>132</sup> Prior to the identification of *CHD7*, hypogonadism was reported to be associated with CHARGE syndrome, including some patients old enough to be diagnosed with IHH, as defined below.<sup>133</sup> Several other studies implicate a possible connection between the KS phenotype and CHARGE syndrome, although it is important to emphasize that these patients with CHARGE syndrome are not yet of pubertal age. Anosmia or hyposmia has been identified in children with CHARGE syndrome,<sup>134</sup> as well as abnormal olfactory bulbs by magnetic resonance imaging (MRI).<sup>135</sup> Most boys studied had micropenis and/or cryptorchidism suggestive of IHH, although all subjects were prepubertal.<sup>135</sup> None of seven females under 12 years of age began puberty spontaneously, and hormonal data suggested that IHH could ensue as they got older<sup>135</sup>. In fact, recently one female with CHARGE syndrome due to a *CHD7* mutation and with some features of KS has been reported.<sup>136</sup>

However, there has been no systematic evaluation of IHH/KS patients without a diagnosis of CHARGE syndrome for the presence of CHD7 mutations. We hypothesized that *CHD7* would be involved in the pathogenesis of IHH and KS (IHH/KS) without the CHARGE phenotype and that IHH/KS represents a milder allelic variant of CHARGE syndrome.

#### **5.4.3 Positional cloning of t(10;12)(q26.12;q13.11)dn**

The balanced chromosome translocation in this KS patient is *de novo*, which involves the 10q26 region. Firstly, we hypothesized that the translocation likely has disrupted a developmentally important novel gene for KS in this patient, since most individuals with balanced translocations are asymptomatic and his parents were both healthy without this balanced rearrangement. The chromosome 10q26 region has been associated previously with abnormal male genital development resulting from interstitial or terminal deletions as well as a balanced translocation. Abnormal development has included cryptorchidism,<sup>137,138</sup> small testes,<sup>139</sup> sperm defects and infertility,<sup>140</sup> micropenis and hypogonadism,<sup>141</sup> hypogonitalism<sup>142</sup> and sparse sexual hair.<sup>143</sup> All of these phenotypic features overlap with characteristics of IHH and KS. Importantly, a trisomy involving 10q26 has been reported in a KS patient with an unbalanced chromosome translocation,<sup>25</sup> suggesting the presence of a new KS gene, which can involve dysregulation as a result of either reduced or increased dosage, in this critical region. Therefore, we postulated that a single disease gene disrupted or dysregulated by a position effect<sup>144</sup> is located at or near the 10q26 breakpoint in this KS patient with a balanced translocation.

### **5.5 Results and Discussion**

#### **5.5.1 Positional cloning of t(7;8)(p12.3;p11.2)dn**

To identify the genes potentially disrupted in the patient, we first mapped the translocation breakpoints using FISH, which was carried out using BACs chosen from within the region of 8p12 and 8p11.2 to narrow the candidate region until a breakpoint crossing BAC clone was identified. The BAC, RP11-100B16 (chr8:38,239,682-

38,403,260: UCSC, hg19), hybridized to the normal chromosome 8 and both der(7) and der(8) chromosomes, indicating that it spans the translocation breakpoint. Additionally two BAC clones that partially overlap with RP11-100B16—RP11-265K5 (chr8:38,108,125-38,285,474: UCSC, hg19) and RP11-359P11 (chr8:38,308,658-38,483,854: UCSC, hg19)—mapped distal and proximal to the breakpoint, respectively, refining the breakpoint to 24 kb at 8p11.2. Subsequent Southern blot analysis narrowed the breakpoint to 1557 bp on the BAC restriction map and suggested that the breakpoint is between exons 2 and 3 of *FGFR1* isoform 1. Cloning and sequencing the junction fragments from both derivative chromosomes revealed that the translocation is perfectly balanced, without the gain or loss of any sequence. Based on partial sequence of chromosome 7 in junction fragments the breakpoint on chromosome 7 was found to be located within BAC RP11-549I23 (chr7:47,303,986-47,489,559: UCSC, hg19) at 7p12.3. The location of the translocation breakpoint predicted the chromosome 7 breakpoint lies in intron 15 of *TENS1* while the breakpoint on chromosome 8 is in intron 2 of *FGFR1*.

The chromosome 8 breakpoint maps within a SINE/Alu repetitive sequence while the breakpoint in *TENS1* occurs in unique intronic sequence with no apparent homology to the chromosome 8 breakpoint region. The location of the two breakpoints predicted two putative reciprocal in-frame fusion transcripts *TENS1/FGFR1* and *FGFR1/TENS1* and RT-PCR detected only the *TENS1-FGFR1* fusion transcript with the skipping of *FGFR1* exon 3. The 5624 bp *TENS1* $\Delta$ ex16-26/*FGFR1* $\Delta$ ex1-3 transcript encodes a fusion protein of 1586 amino acids comprising the first 883 amino acids of *TENS1* joined to the final 703 amino acids of *FGFR1*. Mutation analysis of the second non-translocated *FGFR1* allele from the patient, done by SSCP and direct sequencing, identified only a heterozygous nucleotide difference, 345 C>T in exon 3, a known SNP (NCBI reference SNP ID: rs2915665) which does not alter the Ser amino acid encoded at this site. Thus the presence of the translocated allele of the *FGFR1* results in a disease phenotype without a corresponding coding sequence mutation in the alternate *FGFR1* allele.

The *TENS1* locus encodes tensin-like SH2 domain-containing protein 1 (also known as tumour endothelial marker 6, tensin 3), a 1445 amino acid protein named for its similarity with tensin, an actin filament crosslinking protein found in focal adhesions.<sup>145,146</sup> *TENS1*

protein contains a protein tyrosine phosphatase domain in the amino-terminal region and Src homology 2 (SH2) and phosphotyrosine binding domains near its carboxyl-terminus. *FGFR1* encodes several different isoforms of a transmembrane protein, the extracellular moiety of which interacts with fibroblast growth factors (FGFs), setting in motion a cascade of downstream signals, ultimately influencing mitogenesis and differentiation; it is characterised by two or three extracellular immunoglobulin-like loops (depending on inclusion of exon 3), a transmembrane domain, and an intracellular tyrosine kinase domain. The predicted TENS1/FGFR1 fusion protein lacks the FGFR1 signal peptide and the first Ig-like domain, but contains Ig-like domains 2 and 3, which are sufficient for specific FGF binding, and an intact tyrosine kinase domain region, suggesting the potential for functionality.<sup>147</sup>

However, the related KS phenotype in a patient hemizygous for 8p due to the deletion region suggests that the translocation produces hypogonadotropic hypogonadism as a result of haploinsufficiency for FGFR1.<sup>126</sup> Consistent with this view, there was no evidence for an FGFR1 mutation on the second allele in SSCP/sequence analysis in any of 18 exons and splice junctions. The apparent functional hemizygosity for FGFR1 in the translocation patient probably reflects a failure to direct the FGFR1 functional domains to the proper location in the plasma membrane owing to the absence of the appropriate signal peptide and the presence of the large TENS1 moiety. While this work was being completed, Dode *et al*,<sup>101</sup> and subsequently Sato *et al*,<sup>148</sup> reported several truncating and missense *FGFR1* mutations in KS patients, some with cleft lip and palate, consistent with the haploinsufficiency in the patient reported here.<sup>101,148</sup> However, the absence of frank anosmia in the current patient indicates that sufficient FGFR1 function may have been maintained to prevent the degree of agenesis of the olfactory bulbs typical in KS. As the translocation patient displays no obvious phenotypes distinct from those seen in patients with KS associated *FGFR1* point mutations or microdeletions, the disruption of *TENS1* does not seem to contribute to the patient's abnormalities. This suggests that heterozygous inactivation of TENS1 is without dramatic consequence, but the possibility that the predicted fusion proteins effectively provide normal TENS1 function cannot be excluded.

The X linked form of KS is associated with inactivating mutation of the *KAL1* gene, encoding anosmin 1, a secreted proteoglycan binding protein with similarities to

neuronal cell adhesion molecules.<sup>98,149</sup> Anosmin 1 interacts with heparan and chondroitin sulphates to promote cell adhesion and neuronal outgrowth, and has been implicated in the migration of gonadotropin releasing hormone (GnRH) producing neurones and olfactory axonal fibres, though the receptor system through which it acts remains uncertain.<sup>150</sup> Notably, FGFR1 activation by binding to FGF ligands involves receptor dimerisation that also requires heparan sulphate proteoglycan binding.<sup>151</sup> Indeed, FGF2 ligand and FGFR1 have been co-crystallised with heparin, and the structure of the complex defined.<sup>152</sup> The common association with heparin sulphates and the similar effects of *KAL1* and *FGFR1* inactivating mutations support the suggestion that the FGFR1 signalling pathway participates directly in mediating anosmin 1 function.<sup>101,153</sup>

The translocation patient reported here and the KS patients reported by others also support the view that haploinsufficiency for FGFR1 is a cause of cleft lip and palate.<sup>101,126,148</sup> Interestingly, *FGFR1* gain of function mutations have previously been associated with the craniosynostosis of Pfeiffer's syndrome and in the Jackson–Weiss syndrome.<sup>154,155</sup> These syndromes can also be caused by mutations in *FGFR2*, which has also been associated with cleft palate in Apert's syndrome, indicating that the two receptors may function in the same signalling pathway.<sup>129</sup> This suggests that *FGFR2*, located at 10q26, may be an excellent candidate for an additional KS or idiopathic hypogonadotropic hypogonadism locus. Hence it would be of interest to determine whether *FGFR2* is disrupted by translocation in a KS patient with a *de novo* unbalanced der(1)t(1;10)(q44;q26).<sup>25</sup> It is noteworthy that a variety of fusion proteins involving FGFR1 underlie the 8p11 myeloproliferative syndrome (EMS)/stem cell leukaemia–lymphoma syndrome, presumably because of constitutive activity of the tyrosine kinase domain.<sup>156-159</sup> However, neither Pfeiffer's syndrome nor the Jackson–Weiss syndrome shows a myeloproliferation defect, suggesting that the gain of function in these cases is insufficient or inappropriate to transform target cells. The constitutional translocation reported here creates a predicted fusion protein that has not produced a myeloid disorder, despite being likely to mislocalise a portion of FGFR1 containing the tyrosine kinase domain. This reinforces the view that both the fusion partner and the site of the breakpoint are likely to be critical in producing constitutive tyrosine kinase activity in a manner that leads to malignancy. This is the first demonstration that constitutional

translocation of *FGFR1* can lead to abnormal development rather than to myeloid disorder, and provides a basis for more detailed structure–function comparison of the respective fusion proteins.

### 5.5.2 Candidate gene approach with *CHD7*

We first looked at *Chd7* mRNA expression in three different mouse GnRH neuronal cell lines, two migratory (GN11 and NLT<sup>160</sup>) and one postmigratory (GT1-7<sup>161</sup>). *Chd7*-mRNA expression was confirmed by cloning and sequencing of the PCR products. Next, RT-PCR on RNA extracted from KS-relevant rat tissues demonstrated expression in the olfactory bulb, medial basal hypothalamus, and pituitary. Therefore, we undertook an extensive mutation analysis of *CHD7* in IHH/KS patients without clinical features of CHARGE syndrome. DNA sequencing of the entire 37 protein-coding exons and splice junctions of the *CHD7* gene from 101 affected individuals (50 with KS and 51 with normosmic IHH) revealed six heterozygous mutations—one splice-donor site alteration and five missense mutations for a prevalence of 6%. Another heterozygous splice-donor site mutation was identified in one of 96 additional IHH/KS patients screened for mutations only in exons 6–10. This gene region was selected for a focused mutation screen because these exons encode the highly conserved and important functional chromodomains. All cases with mutations were sporadic, and none of the probands had any affected family members or relatives. No identical nucleotide changes were identified in 180 controls or listed in the single-nucleotide polymorphism (SNP) database. In addition, the two *CHD7* mutations identified in probands of Turkish descent were also absent in 96 Turkish controls. Six new SNPs (submitted into the SNP database in NCBI) in the *CHD7* were also detected. None of the patients with a *CHD7* mutation were known to possess a mutation in another gene involved in IHH/KS.

Both intronic mutations impaired mRNA splicing in lymphoblasts of affected individuals and are predicted to interfere with protein function. An IVS8+5G>A mutation identified in a normosmic IHH male without any other anomalies resulted in exon 8 skipping, as determined by RT-PCR of *CHD7* exons 4–10 from patient and control lymphoblastoid RNA, subsequent subcloning, and DNA sequencing. This mutant introduces 16 aberrant out-of frame AA residues, causing a frameshift and subsequent premature termination codon at residue 849 located 49 bp downstream of the junction of

36

exons 7 and 9. This results in a CHD7 protein predicted to be truncated more than 70% of the C terminus. This transcript is not a normal splice variant because it was absent in four control lymphoblastoid cell lines, and nor was the same G>A transition identified in 180 normal controls. Importantly, this mutation removes about half of the first and all of the second chromodomains, as well as other important domains of the protein, thereby predicting a nonfunctional protein.

In a KS female with mild sensorineural deafness and cleft lip and palate, a *de novo* heterozygous intronic transversion, IVS6+5G>C, was identified. The mutation was absent in both parents. For direct examination of whether this variant affects splicing, CHD7 exons 4–9 were similarly analyzed by RT-PCR. In addition to the normally spliced transcript, a transcript with reduced size demonstrated exon 6 skipping not observed in four controls. This results in an in-frame deletion of 22 of 66 AAs of chromodomain 1. Chromodomains are evolutionarily conserved<sup>162</sup> and known to interact with histone tails.<sup>163,164</sup> Interestingly, this mutation has also been reported in a patient with CHARGE syndrome.<sup>165</sup> The patient with CHARGE syndrome has additional severe phenotypic findings, including the absence of earlobes, triangular concha, vestibular disturbance, autism-spectrum disorder, and mental retardation<sup>165</sup>, which are absent in our KS patient.

The five missense mutations affect highly conserved AA residues when compared with known CHD7 orthologs. Four of five point mutations were predicted to be deleterious by SIFT (Ser834Phe, Lys2948Glu, Pro2880Leu, and His55Arg), whereas Ala2789Thr was tolerated. Ser834Phe, located in DNA-binding chromodomain 1, a highly conserved sequence motif observed in a variety of animal and plant species, had the highest confidence measure. Further supportive evidence of the deleterious effect of Ser834Phe comes from the report of the same mutation in three patients from a family with a severe CHARGE syndrome phenotype.<sup>166</sup>

Structural models for the relevant domains suggest that Ala2789Thr and Pro2880Leu, located in the spacer sequence between the BRK2 and leucine-zipper regions, as well as Lys2948Glu, are also detrimental. All three AA residues are located in loop regions so that mutations of these residues will most likely affect structural and binding properties of the domains to their interaction partners. The local secondary structure of the region around Pro2880 is a random coil, and the Pro2880Leu mutation

induces helix formation in this region, predicting its deleterious effect. The Ser834 residue is also located in a loop region, coordinating strongly with adjacent residues in the neighboring helix (Tyr881, context PDYV), so Phe834 will therefore strongly affect protein stability. The Ala2789 residue is conserved in seven of eight orthologs, and Lys2948 is relatively conserved.

SIFT was also used to characterize the functional significance of a 22 amino acid (ESVDAEGPVVEKIMSSRSVKKQ) in-frame deletion of exon 6 (IVS6+5G>C mutant) by predicting deleterious effects for almost all possible AA substitutions in this region. The splice mutant IVS8+5G>A truncates the functional part of the protein starting at AA 810, including the region from AAs 920–1490. The latter is highly homologous to the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54 (PDB code: 1Z3I)<sup>167</sup> and therefore an essential DNA-binding domain of the wild-type. Collectively, involvement of the conserved chromodomain by three mutations, two of which have been identified in patients with CHARGE syndrome,<sup>165,166</sup> as well as SIFT AA conservation and protein structural analysis, indicates that these nucleotide alterations are pathogenic mutations.

To investigate developmental expression of *Chd7* further in the mouse embryo, in situ hybridization analysis was performed from embryonic day 10.5 (E10.5) to postnatal day 0 (P0). Already by E10.5 and E11, high *Chd7* expression was preferentially observed in the developing nervous system and its derivatives. The entire neuroepithelium was strongly labeled at that stage. At E10.5 and E11, high levels of expression were particularly noteworthy in the olfactory placode, which gives rise to olfactory and GnRH neurons. With progressing differentiation, the label intensity declined and the signal became restricted to specific locations in the developing central and peripheral nervous systems. At E14, high signal intensity was seen in the olfactory epithelium, cochlea, anterior pituitary, and the spinal cord. This pattern of expression is consistent with involvement of *Chd7* in the development of the olfactory pathway and the GnRH-positive neurons. In the adult brain, intense *Chd7* expression was restricted to the granule cell layer of the cerebellum, hippocampal formation, and hypothalamus. Signals were absent when the sense control was used.

The molecular basis for 70%–75% of IHH/KS patients remains unknown.<sup>95</sup> To date, only *FGFR1* mutations have been reported to cause either normosmic IHH families or KS families.<sup>106</sup> Although a homozygous *PROKR2* deletion was seen in a single family

comprising both normosmic and anosmic patients, this represents variable expressivity within the same family.<sup>168</sup> In our case, we have three unrelated probands with KS and four unrelated probands with IHH with *CHD7* mutations, demonstrating that *CHD7* is involved in either IHH or KS. We present new evidence for a role of *CHD7* in the pathophysiology of both normosmic IHH and KS patients without a CHARGE phenotype.

We first demonstrate *Chd7* mRNA expression in both migratory and postmigratory GnRH neuronal cell lines. We also document mRNA expression in the hypothalamus, pituitary, and olfactory bulb in the rat—all of which are IHH/KS-relevant organs. By in situ hybridization, we confirm mRNA expression at the appropriate time when GnRH and olfactory neurons migrate from the olfactory placode region to the hypothalamus. This begins on E10–E11 in the mouse and is virtually completed by E18.5.<sup>169</sup> In humans, GnRH migration begins at E5.5–E6.5 weeks and finishes 6–8 weeks postnatally.<sup>169</sup> Whether or not *CHD7* affects this important developmental neuron migration requires future study.

Furthermore, we demonstrate sporadic heterozygous *CHD7* mutations, which were not present in  $\geq 180$  controls or the SNP database, in humans with both normosmic IHH and anosmic IHH (KS). Additionally, two mutations in Turkish patients were also absent in 96 Turkish controls (for a total of 276 controls). The prevalence of *CHD7* mutations of ~6% is similar to that of apparently sporadic *KAL1* mutations<sup>95</sup> and somewhat less than the 10% reported rate of *FGFR1* mutations in IHH/KS.<sup>106</sup> Two of our *CHD7* mutations alter mRNA splicing in lymphoblast RNA and predict deletions of chromodomains, whereas another missense mutation within chromodomain affects a highly conserved Ser residue. The disruption of these important evolutionarily conserved chromodomains is highly likely to result in deleterious consequences,<sup>162</sup> because chromodomains are known to interact with histone tails.<sup>163,164</sup> Chromodomain deletion has also been reported to impair nucleosome binding and remodeling by CHD proteins,<sup>170</sup> indicating that their disruption will be detrimental to *CHD7* function.

Further supportive evidence that our *CHD7* mutations are deleterious comes from SIFT analysis, which indicates that four of the five missense mutations (Ser834Phe, Lys2948Glu, Pro2880Leu, and His55Arg) involve highly conserved AA residues among known species and, therefore, are not likely to be tolerated by their observed

substitutions.<sup>171</sup> These findings were also corroborated by protein structural analysis of the AA variants Ala2789Thr, Pro2880Leu, and Lys2948Glu, which were predicted to alter structural and binding properties of the domains. Taken together, both AA conservation and protein structural modeling provide additional support that these missense substitutions are highly likely to be deleterious mutations.

Importantly, our one IHH and one KS patient, who both lack the CHARGE phenotype, possess the same mutations (Ser834Phe and IVS6+5G>C) reported previously in patients with CHARGE syndrome,<sup>165,166</sup> further demonstrating the allelic relationship of both syndromes. The KS patient with the IVS6+5G>C mutation does not fulfill Blake's criteria for CHARGE syndrome,<sup>172</sup> although she does have hearing impairment and cleft lip and palate. This also indicates that the effects of modifying genes may determine whether the patient has the more severe CHARGE phenotype rather than the milder IHH/KS phenotype.

Interestingly, our mutations have been localized to regions around four exons-2, 6, 8, and 38-suggesting the possibility of hotspots for IHH/KS mutation. Because of these findings and the absence of nonsense mutations, which often occur in CHARGE syndrome,<sup>173</sup> we provide the first convincing evidence that IHH/KS represents a milder allelic variant of CHARGE syndrome. Although its precise function is uncertain, CHD7 appears to be important in GnRH and olfactory neuron migration to their embryologic destination in the hypothalamus. CHD7 is the first chromatin-remodeling protein involved in normal puberty in humans and is the second gene (after *FGFR1*) identified that results in both normosmic IHH and KS.

### **5.5.3 Positional cloning of t(10;12)(q26.12;q13.11)dn**

We first performed a microarray analysis of DNA of translocation patient. Affymetrix Human Mapping 500K Array SNP oligonucleotide microarray analysis (SOMA) of patient DNA was negative and excluded copy-number variation (CNV) as the cause of the IHH or KS phenotype. This result suggested the disruption of a gene at one translocation breakpoint is likely the cause of the KS phenotype in this patient. To identify potentially disrupted genes, we mapped the translocation breakpoints using FISH.

We first performed FISH analysis using the BAC clones as probes, which identified breakpoint spanning BAC clones RP11-254K03 (chr10:122,064,742-40

122,240,093: UCSC, hg19) and RP11-464D5 (chr12:47,594,539-47,787,685: UCSC, hg19), narrowing the breakpoint to 175 kb and 193 kb in 10q26.12 and 12q13.11, respectively. Ensuing array painting refined both breakpoints to 2.7 kb and 4.2 kb in 10q26.12 and 12q13.11, respectively. Breakpoint cloning results demonstrate that no gene is directly disrupted by the translocation breakpoints at 10q26.12 and 12q13.11. To select candidate genes for mutation analysis, RT-PCR was performed for four positional candidate genes-*FGFR2*, *PPAPDC1A*, *SEC23IP*, and *WDR11*-in the IHH/KS relevant rat tissues ovary, testis, olfactory bulb, hypothalamus, pituitary. All four genes showed different levels of expression in these tissues. No mutations were present in *FGFR2*, *PPAPDC1A*, and *SEC23IP*. However, we have shown the expression of *WDR11* is markedly reduced in RT-qPCR and western blot analysis, suggesting a position effect of the translocation breakpoint.

We screened 201 individuals with IHH/KS for mutations in *WDR11* and identified a total of five different heterozygous missense mutations (R395W, A435T, R448Q, H690Q, F1150L) in six patients. Our mutation detection rate was about 3 %. Three of the missense changes altering amino acid residues R395, H690, and F1150 are completely conserved in all higher vertebrates. Two unrelated patients had an identical heterozygous missense alteration F1150L and haplotype analysis of these two patients (C17 and C100) with seven informative SNP markers indicated that they share a haplotype around *WDR11* and are therefore likely to be descended from a recent common ancestor.

Our mapping, cloning, and sequencing of the breakpoints of the *de novo* balanced chromosome translocation from a KS male with karyotype 46,XY,t(10;12)(q26.12;q13.11)dn and the consequent detection of multiple independent missense variants in *WDR11* in KS and IHH patients argues strongly for a causative role for *WDR11* in this disorder. Regrettably, DNA samples of the patients' parents are not available, and all efforts to track the parents have been unsuccessful. However, because all *WDR11* mutations were heterozygous, autosomal-dominant inheritance of the IHH phenotype is likely. Consistent with this interpretation, no *WDR11* mutation was found in any of 29 exons or splice junctions on the untranslocated second allele from our

translocation patient, suggesting that a reduction in functional WDR11 as a result of dysregulation of the gene via a position effect of the translocation is the cause of KS in this subject. The absence of truncating nonsense and frameshift mutations indicates that these might produce a more severe phenotype or embryonic lethality.

*WDR11* was originally identified as a potential tumor suppressor by positional cloning of a somatically acquired t(10;19) chromosome translocation that generated an intragenic deletion in human glioblastoma cells.<sup>174</sup> This deletion disrupted *WDR11* and fused it to the *ZNF320* gene (MIM 606427). Interestingly, other genes associated with mammalian puberty are also known for their involvement in tumorigenesis.<sup>175</sup> For example, hypothalamic expression of certain tumor suppressor genes is increased at puberty in monkeys or decreased in mice with delayed puberty.<sup>176</sup> Similarly, *KISS1* (MIM 603286), a tumor metastasis suppressor gene<sup>177</sup> in melanomas<sup>178</sup> and breast carcinomas<sup>179</sup> encodes the peptide ligand of G-protein-coupled receptor 54 (GPR54),<sup>180</sup> which plays a critical role in the initiation of puberty. *KISS1R* (encoding GPR54) mutations cause autosomal-recessive IHH in consanguineous families and mice.<sup>107,108</sup>

*WDR11*'s function is unknown, but it contains twelve WD domains and is highly conserved throughout vertebrate evolution. The gene was previously annotated as *BRWD2* (Bromodomain and WD repeat domain containing 2) but has recently been more appropriately designated *WDR11* because it has WD domains but no Bromo domains.<sup>181</sup> Mutations of other WD proteins have been associated previously with mammalian reproduction. Both genders of *Repro5* mice, which have ENU-induced *Brwd1* (*Wdr9*) mutations, demonstrate infertility. In addition, patients with mutations of *BRWD3* (MIM 300553) have undescended testes and minimal facial or axillary hair.<sup>182</sup> However, this is the first report of a specific role for a WD protein in IHH or KS, and it implicates protein-protein interaction mediated by WD repeats of *WDR11* as a critical requirement for normal puberty.

Proteins with repeated WD domains, each of which consists of four antiparallel  $\beta$  strands, form  $\beta$  propeller structures to support interactions with protein-binding partners and to organize and stabilize multiprotein complexes<sup>183</sup>. A  $\beta$  propeller is characterized by 4–8 bladeshaped  $\beta$  sheets arranged around a central axis; each sheet of four antiparallel  $\beta$  strands is twisted so that the first and fourth sheets are close to

perpendicular. The last  $\beta$  strand of one WD repeat, and the first three  $\beta$  strands of the WD repeat form a blade of the  $\beta$  propeller.

Three of the *WDR11* missense mutations leading to R395W, H690Q, and F1150L alter amino acid residues that are completely conserved in all 13 available mammalian and avian orthologs (human, chimpanzee, cow, horse, panda, pig, dog, rat, mouse, rabbit, opossum, chicken, and finch), and a fourth change, A435T, is shared in 11 out of 13 species, suggesting that these substitutions in five independent sporadic patients are very likely to be detrimental. Three of the missense alterations are located directly in the WD domains important for  $\beta$  propeller formation and protein-protein interaction: A435T and R448Q are in the sixth WD domain, and H690Q is in the ninth WD domain. These three mutations are predicted by SPPIDER to alter protein-protein binding domains defined by protein modeling and therefore are likely to disrupt normal protein function.

The similarity in structural modeling between *WDR11* and the known structure of *AIP1* suggests that *WDR11*, like *AIP1*, may form a dimer stabilized by interaction with two zinc ions. Because the two protein structures are not identical, deviations arising from their alignment make the position of the Zn ion uncertain to a few angstroms. However, *WDR11* has the required residues (Asp377, Glu384, His501, His508, and Glu510) for zinc binding in the vicinity of the putative zinc position. The R448Q mutation is less than 5 angstroms from the predicted zinc binding site. Arg residues near Zn coordination sites do not directly interact with the Zn, but they stabilize their environment because Arg is highly positively charged. Replacing Arg with the much smaller Gln residue could influence the zinc-binding propensity of *WDR11* and affect its dimer formation and interactions, including a potential actin interaction predicted by analogy with *AIP1*.<sup>184</sup>

In view of the importance of protein-protein interactions for the function of WD proteins, we sought binding partners for *WDR11* and identified the transcription factor *EMX1* as a novel interactor by a yeast two-hybrid screen. *EMX1* is a homeobox transcription factor involved in specifying cell fates in the developing central nervous system,<sup>185</sup> and it participates in the development of olfactory neurons.<sup>186</sup> This putative transcription factor has been shown to be one of the downstream target genes for Gli-Kruppel family member 3 (*Gli3*, MIM 165240) transcription factor,<sup>187,188</sup> which is a part

of the Sonic hedgehog-Patched-Gli (Shh-Ptch-Gli) signaling pathway<sup>189</sup> important in endocrine signaling.

We examined the subcellular localization of both EMX1 and WDR11 in a physiologically relevant human cell system of U2OS and FCNB4-hTERT. When these cells were treated with nuclear export inhibitor leptomycin B, both EMX1 and WDR11 colocalized in the nucleus, suggesting that WDR11 might be shuttling between the nucleus and cytoplasm. Analysis of the expression patterns of WDR11 in mouse and zebrafish development revealed overlapping patterns of expression with EMX1 in regions critical for formation of the hypothalamus, supporting the opportunity for the two proteins to interact *in vivo* and to act together during development. Deletion analysis revealed that WDR11 interacts with EMX1 via both its N terminus and its central region, where four (R395W, A435T, R448Q, and H690Q) of the five WDR11 missense alterations were found. R448Q reduced and both A435T and H690Q abolished binding to EMX1, physically decreasing the opportunity for productive interaction. Interestingly, the mutant R395W did not appear to affect EMX1 binding, but this represents alteration of an amino acid that is invariant in higher vertebrates, suggesting that this alteration permits binding but impairs the as-yet undefined functional consequences of the WDR11-EMX1 interaction.

Taken together, our genetic and functional data provide strong evidence for missense sequence variants of *WDR11* as a cause of IHH and KS in a proportion of cases of this genetically heterogeneous condition. This adds to the growing list of genes known to be mutated in IHH and KS and will open new investigative routes for understanding the development of normal human puberty and reproduction. Importantly, the identification of EMX1 as a binding partner of WDR11 whose interaction can be disrupted by IHH- and KS-associated mutations is significant for two principal reasons. First, it places WDR11, whose biological function is not well understood, as a potential player in the Sonic hedgehog-Patched-Gli-Emx signaling pathway via its interaction with EMX1. Second, the demonstration of a developmental role for WDR11 in IHH and KS suggests a possible connection between Shh signaling and pubertal development. The potential for defects in this pathway in IHH and KS is worthy of further exploration.

## 6. Potocki-Shaffer syndrome

### 6.1 Introduction and Genetics

Potocki-Shaffer syndrome (PSS; MIM 601224) is a rare contiguous gene deletion syndrome caused by a heterozygous interstitial deletion of band p11.2 of chromosome 11 and is characterized by developmental defects that include intellectual disability (ID), multiple exostoses (ME), parietal foramina (PF), and craniofacial anomalies (CFA) as the main features.<sup>190,191</sup> The full spectrum of PSS manifests when deletions are at least 2.1 Mb in size, spanning the segment from *D11S1393* to *D11S1385/D11S1319* in 11p11.2.<sup>192</sup> This suggests that the clinical expression of PSS in particular cases is determined by the precise location and extent of the deletion because specific genes, when rendered hemizygous, are responsible for the individual phenotypic features of PSS.

The challenging aspect of positional cloning in such contiguous gene syndromes is to assign individual phenotypes to individual genes in the deleted region. As the size and location of contiguous gene syndrome deletions may vary from patient to patient, comparison of overlapping regions to define a minimal candidate region associated with a particular phenotype has often preceded candidate gene analysis to identify the causative gene.<sup>101,173</sup>

So far this strategy has led to two genes implicated in the pathogenesis of PSS. *ALX4* has been identified as the gene from the PSS deletion associated with parietal foramina, which consists of symmetrical and oval defects in the parietal bone situated on either side of the sagittal suture and separated by a narrow bridge of bone.<sup>193-195</sup> This gene encodes a mesenchymal homeodomain transcription factor that is a key regulator of calvarial development.<sup>196</sup> *EXT2* was identified as a gene from the PSS interval responsible for multiple exostoses, benign cartilaginous tumors (enchondromata) on the long bones.<sup>197,198</sup> Linkage analysis in large families with multiple exostoses and deletion analysis in PSS families defined the candidate region,<sup>190</sup> which helped to identify this presumed tumor suppressor gene.<sup>199,200</sup> *EXT2* encodes an endoplasmic reticulum-localized type II transmembrane glycoprotein that possesses or is tightly associated with glycosyltransferase activities involved in the polymerization of heparan sulfate.<sup>201,202</sup>

Craniofacial abnormalities (CFA) are present in about one-third of all human congenital defects<sup>203</sup> and intellectual disability (ID) affects about 2%-3% of the human population.<sup>204</sup> In many cases, these phenotypes are present in combination rather than as distinct separable entities. Here, we have complemented *in silico* deletion mapping using three published PSS patients without ID and CFA,<sup>205-207</sup> which narrowed the minimal candidate interval of ID/CFA in PSS to ~1.1 Mb.

Further, we identified two individuals with *de novo* balanced translocations, with both breakpoints in the 1.1 Mb of 11p11.2 critical PSS region, who display ID/CFA as well as other typical PSS phenotypes, in the absence of multiple exostoses and parietal foramina. Cytogenetic analysis disclosed an apparently balanced *de novo* translocation between chromosomes 11 and 19 in one boy, resulting in a 46,XY,t(11;19)(p11.2;p13.2)dn karyotype. In another female an apparently balanced *de novo* translocation involving chromosomes 1 and 11, 46,XX,t(1;11)(p13;p11.2)dn was detected. These two chromosomal rearrangements are instrumental to identify the autosomal genetic etiology of syndromic intellectual disability involving craniofacial anomalies.

## 6.2 Subjects

### 6.2.1 DGAP012 patient with t(11;19)(p11.2;p13.2)dn

At birth, the DGAP012 subject, a white male, was small for gestational age, but not misproportioned. He suffered from supraventricular tachycardia, hyperbilirubinemia and hypoglycemia all of which resolved. During pregnancy, the mother declined amniocentesis, which was offered because ultrasound showed nuchal thickening, raising the possibility of Down syndrome. After birth, chromosome analysis revealed an apparently balanced chromosomal translocation with the karyotype, 46,XY,t(11;19)(p11.2;p13.3)dn, modified by molecular analysis here to 46,XY,t(11;19)(p11.2;p13.2)dn. The family history was unremarkable and parental chromosomes of both parents were normal. Serum amino acids, mucopolysaccharides, venous blood gas, lactate, ammonium, urine organic acids, oligo TLC, CBC, and TSH were all normal.

A CT scan of the head at age one year showed an element of bifrontal biparietal atrophy with some extension of atrophy along the anterior aspect of the interhemispheric fissure. At 15 months of age, the subject displayed significant global developmental delay. Systemic examination of DGAP012 revealed digitalized thumbs, brachycephaly, microcephaly (head circumference 46 cm; 10<sup>th</sup> percentile), small down turned mouth, mild midfacial hypoplasia, flat mid-face, narrow nasal bridge, very small nose, large ears, and bilateral epiblepharon (extra skin fold medially under each lower lid) without trichiasis, small hands and feet and absence of emotional expression. He also displayed hand flapping and had feeding problems prior to the age of three. An ocular examination revealed an anterior insertion of the conjunctiva on the cornea, a common non-specific minor malformation. He showed good fixation and following movements with either eye, although at times he was somewhat inattentive especially to distant targets. He had strabismus, but no nystagmus, wandering movements or photophobia which might otherwise characterize a child with a severe ocular disorder. The remainder of the eye examination was normal. MRI at the age of two revealed prominent CSF. EEG was within normal limits. At the age of 5 years, DGAP012 displayed hypotonia and dystonic movement and could not walk. Eye movements were full but cycloplegic refraction showed significant nearsightedness and astigmatism (-3.00 +3.00 x 90 OU) for which glasses were prescribed. Dilated retinal examination was normal with the exception of some diffuse pigmentary mottling suggesting a possible retinal dystrophy. Electroretinography showed diminishment of rod/cone and cone responses but normal rod responses. No further follow-up is available.

### **6.2.2 MCN1762 patient with t(1;11)(p13;p11.2)dn**

This female subject, presently aged 42 y, is of Caucasian origin. She was delivered at term after an uneventful pregnancy with a birth weight of 3400 g. During her first months she was lethargic and had feeding difficulties. At age 3½ y she was referred for evaluation of developmental delay. Her gross motor development was normal, but her speech development was delayed, and she also had behavioural problems such as hyperactivity and poor concentration. Verbal IQ test (Terman-Merrill) indicated a delay of one year, and non-verbal IQ test (Leiter) was normal. The following dysmorphic features were noticed: brachycephaly, microcephaly, down slanting palpebral fissures, epicanthal

folds, narrow nose, mild midfacial hypoplasia, down turned mouth, prominent ear lobes, short broad hands and knock-knees. Eye examination was normal apart from mild myopia (-2/-2.5). G-banded karyotyping displayed an apparently balanced reciprocal translocation, 46,XX,t(1;11)(p13;p11)dn. Her parents and two sibs are healthy, and parental karyotypes are normal. She has attended a school for special needs and is able to read. Her linear growth and pubertal development have been normal. Her final height is 173 cm. She has developed truncal obesity. She has had viral meningitis and a traumatic bone fracture without sequelae. Presently, she appears mildly mentally retarded; she manages to live independently and to have a sheltered part time job. The following investigations have been performed and were all normal: audiometry, urine metabolic screening, s-calcium, s-phosphate, alkaline phosphatases and s-phenylalanine. Brain imaging has not been done.

### **6.2.3 GM03316**

GM03316 was sent to NIGMS in January of 1978 (long before Potocki-Shaffer syndrome was identified). It was a blood sample from a 3 year old South American girl. The specimen arrived from Venezuela. The submitter called it an apparently balanced *de novo* (X;11) translocation. The clinical symptoms were listed as: global retardation (GQ = 60) with bizarre dysmorphism syndrome including epicanthus, hypertelorism, oblique palpebral fissures, trigonocephaly and micrognathia. At 5 years of age her GQ corresponds to that of a girl of almost 3 years old. She has a good memory and when she wishes she can feed herself. She got control of her sphincters both day and night by age 3 and a half. Her vocabulary is progressing well. She is shy and easily frightened. She is clumsy with hands and legs. She has three dental carries. Her principal problem is the lack of concentration. Her karyotype was 46,XX,t(X;11)(q11.1;p11.2). The normal X was late replicating. Microarray analysis identified a del(11)(p11.2p12) with 3.6 Mb including the *PHF21A* gene.

### 6.3 Hypothesis

PSS is an autosomal dominant contiguous gene deletion disorder due to an interstitial deletion of band p11.2 of chromosome 11 due to haploinsufficiency of multiple, functionally unrelated yet physically contiguous genes in the deleted region. PSS is characterized by multiple exostoses, parietal foramina, intellectual disability (ID), and craniofacial anomalies (CFA). It is likely that the clinical manifestations of this disorder arise from the combined impact of haploinsufficiency for multiple genes.<sup>208</sup> Despite the identification of individual genes responsible for multiple exostoses and parietal foramina in PSS, the identity of the causative gene(s) for the ID/CFA phenotypes has remained elusive.

Next, we identified two independent balanced chromosome translocation patients with ID and CFA, DGAP012 and MCN1762. A child with a *de novo* balanced translocation between chromosomes 11 and 19, and a female with a *de novo* balanced translocation between chromosomes 1 and 11, shared with a breakpoint in the 11p11.2 critical PSS region. Both probands display ID and CFA reminiscent of PSS in the absence of multiple exostoses and parietal foramina.

The independent breakpoints in two individuals with related phenotypes occurring in the same chromosomal region *de novo* strongly indicates that the chromosomal disruption is causative. We hypothesized that ID and CFA in both patients are caused by haploinsufficiency of an unknown gene disrupted at a breakpoint 11p11.2 within a narrowed interval of 1.1 Mb region we refined by *in silico* deletion mapping. To this end we focused on mapping the location of the 11p11.2 breakpoint in these two translocations.

### 6.4 Results

#### 6.4.1 *PHF21A* is disrupted in unrelated subjects with chromosomal translocations, ID, and CFA

The Developmental Genome Anatomy Project (DGAP) is a collaborative effort to identify genes of developmental importance by studying individuals with apparently balanced chromosomal abnormalities and developmental defects.<sup>144</sup> Identification of multiple

cases where the same gene is disrupted in independent subjects with *de novo* translocations and similar phenotypes provides particularly strong evidence of the causative nature of the lesion. Through DGAP, we identified a subject (DGAP012) with an apparently balanced *de novo* translocation between chromosomes 11 and 19, resulting in a 46,XY,t(11;19)(p11.2;p13.2)dn karyotype. A second subject, MCN1762 (MCN19730002-227), identified through the Mendelian Cytogenetic Network database, had an apparently balanced *de novo* translocation between chromosomes 1 and 11, initially reported as 46,XX,t(1;11)(p13;p11)dn. They both display evidence of ID with CFA, as well as other features typical of PSS but without multiple exostoses and parietal foramina, suggesting that the disruption in each case might affect the same gene in or near the PSS region in 11p11.2.

To map precisely the translocation in DGAP012, we first used FISH to bracket a candidate region and then to define a breakpoint crossing BAC.<sup>209</sup> After DNA blotting to refine the breakpoints, we used suppression cloning<sup>210</sup> and targeted PCR for subsequent isolation and sequencing of junction fragments. The chromosome 19 breakpoint lies within a *SINE/Alu* sequence in intron 5 of *ELAVL1*, while the chromosome 11 breakpoint interrupts a *LINE/L2* repetitive element in intron 14 of *PHF21A*. At the der(19) breakpoint, there was a five nucleotide CTCCT deletion of chromosome 11 sequence and a five nucleotide TTCAG deletion of chromosome 19 sequence, while on the der(11) there was no loss or gain of nucleotides due to the translocation. The breakpoints were also confirmed independently using a multiplexed targeted capture and sequencing approach as previously described.<sup>211</sup>

For the second subject, MCN1762, we used FISH and array painting followed by PCR amplification to isolate the breakpoints, resulting in revision of the previously designated karyotype to 46,XY,t(1;11)(p21.1;p11.2)dn. The chromosome 1 breakpoint lies within a small 48 bp non-repetitive sequence surrounded by *LINE/L1* sequences in a region devoid of annotated genes, 635 kb distal from the 5' end of *PRMT6* (MIM 608274). The chromosome 11 breakpoint is in intron 5 of *PHF21A*. At the der(1) breakpoint, there was a two bp (TT) deletion of chromosome 1 sequence and an eight bp CTCCAAAT insertion, while on the der(11) there was a three bp TTA deletion of chromosome 1 sequence.

Immunofluorescence in cultured cells using an antibody specific to the amino-terminal segment of PHF21A revealed that the majority of the protein resides in nuclei, consistent with a role for PHF21A in nuclear processes, such as chromatin association and transcriptional regulation. As expected from the sites of the translocations in both DGAP012 and MCN1762, disruption of *PHF21A* resulted in reduced protein levels of full-length PHF21A, as detected by western blot analysis. In both subjects, the *PHF21A* promoter could theoretically drive expression of a truncated PHF21A protein, either alone or, in the case of DGAP012, as part of a fusion protein. However, in neither case would such a protein product contain the plant homeodomain (PHD) required for binding H3K4me0.<sup>212</sup>

Localization of the chromosome 11 breakpoint from each subject within the same gene with consequent reduction in the protein product strongly supports a role for disruption of *PHF21A* in generating the ID and CFA phenotypes. This possibility was further supported by a third translocation subject, a young girl who displayed both ID and CFA in conjunction with phenotypes of reminiscent of Gillespie syndrome (MIM 206700). In this female subject, the translocation directly disrupted both *PHF21A* and the X chromosome gene *ARHGAP6* (MIM 300118), encoding Rho-GTPase-activating protein 6.<sup>213,214</sup> The precise breakpoints in this case were not determined.

#### **6.4.2 *PHF21A* maps to the critical PSS interval associated with ID and CFA**

Mapping of the translocation sites in *PHF21A* places them in proximity to the genomic deletion region in 11p11.2 associated with PSS, in which phenotypic manifestations depend upon the precise location and extent of the deletion.<sup>192</sup> The full phenotypic spectrum of PSS is manifest when deletions are at least 2.1 Mb in size,<sup>192</sup> containing 16 annotated genes and spanning the segment from *D11S1393* to *D11S1385/D11S1319*. The three translocation subjects share phenotypes with PSS subjects, with the notable exceptions of multiple exostoses and parietal foramina, which are known to be caused by two genes (*EXT2* and *ALX4* respectively) distal to *PHF21A*. Notably, neither of the two translocation subjects investigated here showed a second alteration on the non-translocated chromosome by exon sequencing of *PHF21A*.

A critical reassessment of the *in silico* comparative deletion mapping of published deletion subjects, including those who display multiple exostoses and parietal foramina without ID and CFA<sup>205-207</sup> excludes the distal portion of the PSS interval. This leaves an ~1.1 Mb ID/CFA candidate interval between *D11S554* and *D11S1319* encompassing 12 annotated genes (*TSPAN18*, *TP53I11*, *PRDM11*, *SYT13*, *CHST1*, *SLC35C1*, *CRY2*, *MAPK8IP1*, *C11orf94*, *PEX16*, *GYLTL1B*, *PHF21A*). *PHF21A* maps at the centromeric end of this candidate interval, with markers *D11S1385* and *D11S1319* being located within intron 5.

To test whether *PHF21A* disruption occurred in PSS with ID and CFA, we investigated 5 subjects clinically diagnosed with PSS and 1 with PSS-like phenotype in association with deletion in 11p11.2 where the extent of deletion had not been previously resolved. For 5 subjects (PSS02, PSS08, PSS10, GC14361 and GM03316), we were able to perform FISH with BAC clone RP11-618K13, which contains the telomeric 65 kb of *PHF21A* and spans the breakpoint in DGAP012. This BAC revealed heterozygous deletion in all five subjects. In addition, we carried out array CGH analysis on PSS02, PSS08, GC14361, GM03316, and one additional PSS subject (PSS-Romeike) reported recently,<sup>215</sup> which confirmed the loss of *PHF21A* in the first four and also established it for the final subject. Thus, *PHF21A* is hemizygous in all six subjects, five of whom were clinically diagnosed with PSS and all of whom displayed both ID and CFA phenotypes comparable to the translocation subjects.

To test the frequency of *PHF21A* hemizygosity in apparently normal individuals, we obtained a CNV data set for 13,991 independent control individuals from the International Schizophrenia Consortium (ISC), collected across multiple studies,<sup>216-221</sup> and from Cooper et al.<sup>222</sup> In this collection, the only structural variations (SVs) that crossed an exon of *PHF21A* were two identical 590 bp duplications that overlapped portions of exon 1 and intron 1 (the coding sequence of *PHF21A* begins in exon 3). We also examined the Database of Genomic Variants (DGV), which showed six SVs reported in this region, including four small deletions within introns of *PHF21A*, sized from 444 bp to 5230 bp (intron 1, two intron 5, intron 6), and one example of the aforementioned 590 bp 5'UTR duplication. Data from the 1,000 Genomes project also showed multiple CNVs at each of the exon 1 and intron 5 locations, along with a single

545 bp CNV in the 3'UTR.<sup>223</sup> In addition to these presumably benign SVs, there is a single report in DGV of a large 75 kb deletion encompassing 6 genes, including a portion of *PHF21A*. We are not in a position to validate that deletion or to confirm the absence of phenotype; however, there is precedent for even well-established disease-associated CNVs being non-penetrant in some individuals.

#### **6.4.3 Murine *Phf21a* is expressed in the CNS and cranial bones**

To determine whether the pattern of expression of *PHF21A* supports a role in craniofacial and neuronal development, we performed *in situ* hybridization experiments for the orthologous mouse gene. Predominant expression of *Phf21a* is detected in the developing central nervous system at early stages. At mouse embryonic days 13.5 and 14.5, the roof of the neopallidal cortex (rnc) and thus the developing cerebral cortex, as well as the roof of the midbrain (rom) and the spinal cord (sc), showed highest expression levels of *Phf21a*. The intraventricular portion of the cerebellar primordium (cp) also expressed *Phf21a* at E14.5. At early embryonic stages facial bone and viscerocranial ossification initiates and, with ongoing ossification, high levels of *Phf21a* transcripts were found at E17.5 in the palatine bone and the orbito-sphenoidal bone as well as in the calvaria (cal). Signals observed in bone with *Phf21a* antisense probes were essentially restricted to cranial bones, suggesting a particular function for *Phf21a* in craniofacial development. In adult mouse brain, the most abundant expression of *Phf21a* was observed in the neuronal layers of the hippocampus, the granule cell layer of the cerebellum and the main olfactory bulb. Taken together, these findings indicate that expression of *Phf21a* is consistent with an important role in central nervous system and craniofacial skeletal development and in adult neuronal function. Interestingly, a single report of a mouse knock-out for *Phf21a* described no gross morphological abnormality, although the potential for CFA was not specifically evaluated. Neonatal mice died due to an inability to suckle properly, which was interpreted as a likely defect in neuronal control of milk-sucking behavior.<sup>224</sup>

#### 6.4.4 Suppression of zebrafish *phf21a* expression causes CFA and neuronal apoptosis

To test directly the developmental importance of PHF21A protein, we isolated the zebrafish *phf21a* orthologue, examined its expression pattern, and performed gain- and loss-of-function experiments in this model organism. The zebrafish *phf21a* is highly related to human and mouse PHF21A proteins, exhibiting an AT-hook domain, a PHD, and two coiled-coil domains. Using RT-PCR, we confirmed that zebrafish *phf21a* showed maternal and zygotic transcripts during embryonic development. Whole-mount *in situ* hybridization analyses revealed *phf21a* transcripts ubiquitously distributed throughout the embryo during cleavage, blastula, gastrula, and early segmentation stages. Expression in the head region was increased from later somitogenesis and continued to 24 and 48 hpf.

To investigate the function of *phf21a* in zebrafish development, we tested the effect of *phf21a* knock-down by anti-sense morpholino oligonucleotide (MO). Injection of *phf21a* MO, but not a standard control MO or no MO, caused a small head phenotype and facial dysmorphism with a pronounced defect in growth of the lower jaw at three days post-fertilization (dpf), reminiscent of the microcephaly and dysmorphism in the translocation subjects.

We examined the head structure of *phf21a* morphants in more detail by using Alcian blue to visualize the extent of cartilage development in larval fishes. At five dpf, Meckel's and palatoquadrate cartilages were severely distorted in their size and shape in *phf21a* morphants. Such defects were already manifest during early stages, since we also observed defects in *dlx2a*-positive pharyngeal arch cartilage progenitor cells in *phf21a* morphants at two dpf. Defects in cranial cartilage formation were also observed in the zebrafish *headless* mutation, which is known to be involved in the signaling pathway of vertebrate head formation and patterning.<sup>225</sup> To investigate further whether these defects also involved other arch-associated structures, we injected *phf21a* MO into *flk1:GFP* transgenic zebrafish in which the vascular endothelial cells were visualized by GFP fluorescence.<sup>226</sup> At four dpf, the aortic arches of the *phf21a* MO-injected *flk1:GFP* transgenic embryos were found to be hypoplastic, showing poor development of capillary networks associated with pharyngeal arches. In vertebrates, Meckel's and palatoquadrate cartilages form the embryonic jaw apparatus.<sup>129,227</sup> Thus, it would be

interesting to examine whether the *Phf21a/Bhc80*-deficient mice, which display a failure to suckle,<sup>224</sup> may also have a defect in jaw structure.

We also examined the effects of gain or loss of *phf21a* function on neuronal development but did not see any prominent change. However, injection of *phf21a* MO, but not a standard control MO or *PHF21a* mRNA, caused apoptosis in the developing brain region at 36 hpf. Importantly, this apoptosis and the small head phenotype can be rescued by introduction of wild-type human *PHF21A* mRNA, suggesting that the ID phenotype in humans may be due to a requirement for PHF21A in the function of neuronal cell survival in the developing brain.

Overall the *phf21a*-antisense morpholino caused craniofacial, morphological, and growth defects in the developing zebrafish embryo, as depicted by the notable ventral curvature of the body and small head phenotype relative to the control. The body axis of the zebrafish normally straightens from its early curvature in the pharyngula stage (24-48 hr) of development, at which *phf21a* is strongly expressed in the spinal cord. Therefore, *phf21a* depletion in the spinal cord might impair the straightening process, resulting in persistent spinal and tail curvature. Two processes that have previously been implicated in such ventral curvature are dorsal midline/axis development and cilia development but the relationship of this phenotype to human PSS is unclear. Notably, injection of either wildtype *phf21A* or *PHF21A* mRNA rescued both the spinal curvature and small head phenotypes, demonstrating that they result from a lack of a conserved function of the protein.

#### **6.4.5 Disruption of *PHF21A* in the translocation subjects derepresses**

##### ***SCN3A***

PHF21A is known to be a component, along with LSD1, of the CoREST-related protein complex, BRAF-HDAC complex (BHC), which participates in neuron-specific gene repression, presumably by regulating histone demethylation activity.<sup>228,229</sup> PHF21A specifically recognizes unmethylated histone H3 lysine 4 residues and is required for LSD1 mediated transcriptional repression and LSD1 occupancy at target promoters.<sup>212</sup> To investigate whether LSD1-mediated transcriptional repression is functionally altered by the *PHF21A*-disrupting translocations, we first examined transcription levels of several reported LSD1 targets, including *SCN1A*, *SCN3A* and *SYN1*, in lymphoblastoid

cell lines derived from normal male, normal female, and the translocation subjects DGAP012 and MCN1762. We found that *SYN1* and *SCN1A* respectively have high and moderate levels of expression even in normal lymphoblastoid cells, indicating that they are not epigenetically repressed, as they are in HeLa cells.<sup>228,229</sup> However, *SCN3A* mRNA was expressed at a lower level as measured by RT-PCR in lymphoblastoid cells from normal controls, suggesting that this LSD1 target gene is transcriptionally repressed. An increase of *SCN3A* transcript was readily detected in DGAP012 and MCN1762 cells, indicating that a single functional allele of *PHF21A* might not be sufficient for effective repression of *SCN3A* transcription. This is consistent with correct dosage of *PHF21A* being important for its function. In support of this hypothesis, ChIP results showed a significant reduction of LSD1 binding to the *SCN3A* promoter in DGAP012 and MCN1762 cells compared to normal controls. These findings indicate that *SCN3A* may be repressed by the LSD1 co-repressor complex through a mechanism similar to that reported in HeLa cells<sup>212,228</sup> and that *PHF21A* might be similarly required for LSD1 promoter occupancy in lymphoblasts. This functional disruption of *PHF21A* in DGAP012 and MCN1762 lymphoblasts is likely to be mirrored by alterations in gene regulation in many tissues, including the central nervous system.

## 6.5 Discussion

PSS is a contiguous gene syndrome involving ID and CFA, along with other distinctive features, including eye abnormalities (severe myopia, nystagmus, strabismus), skeletal anomalies (small hands and feet, tapering fingers), multiple exostoses and enlarged parietal foramina.<sup>191,192,207,230-232</sup> The challenging aspect of positional cloning in such contiguous gene syndromes is to assign individual phenotypes to individual genes in the deleted region. As the size and location of contiguous gene syndrome deletions may vary from individual to individual, comparison of overlapping regions to define a minimal candidate region associated with a particular phenotype has often preceded candidate gene analysis to identify the associated gene.<sup>101,173</sup> This strategy and linkage analysis, respectively, led to two genes implicated in the pathogenesis of PSS, *ALX4*, associated with enlarged parietal foramina,<sup>233</sup> and *EXT2*, responsible for multiple exostoses,<sup>197</sup> but these approaches have not identified the gene(s) responsible for the ID and CFA

phenotypes. Attempts to identify the gene(s) underlying the latter PSS phenotypes have been hampered by the relatively large size of the minimal candidate interval (~2.1 Mb).<sup>192</sup> Both categories of clinical feature are individually relatively common, as ID affects ~2%-3% of humans<sup>204</sup> and CFA is present in ~1/3 of human congenital defects.<sup>203</sup> In many cases, these phenotypes present together, suggesting a common underlying etiology. Our reinterpretation of the PSS candidate region for ID and CFA to ~1.1 Mb based upon published subjects without ID and CFA,<sup>205-207</sup> and the identification of two independent translocation subjects with breakpoints at the proximal end of this region, suggest that disruption of a single gene, *PHF21A*, is responsible for both ID and CFA. This finding is supported by a third translocation case from the literature.<sup>213,214</sup>

Haploinsufficiency of *PHF21A* is the probable cause of ID and CFA in all cases that we studied, as we detected no additional mutation of *PHF21A* on the non-translocated alleles in DGAP012 and MCN1762, and no clear differences in critical ID and CFA features of these subjects with PSS subjects with interstitial deletion. In MCN1762, *PHF21A* was the only gene disrupted while in DGAP012, *ELAVL1* was also truncated. *ELAVL1* encodes a protein that contains three RNA-binding domains and binds *cis*-acting AU-rich elements. It destabilizes mRNAs and thereby regulates gene expression.<sup>234,235</sup> As both subjects show comparable phenotypes, it is unlikely that *ELAVL1* haploinsufficiency or fusion proteins contribute strongly to the critical PSS-like phenotypes, although they may be responsible for unusual features in DGAP012, such as possible retinal dystrophy and digitized thumb. The third translocation subject was only examined as a very young girl and displayed some phenotypes (bilateral superior coloboma, foveal hypoplasia, inferior cerebellar hypoplasia, ID) thought to be suggestive of Gillespie syndrome (partial aniridia, cerebellar ataxia, ID). Gillespie syndrome is known to be genetically heterogeneous, caused in some cases by lesions in *PAX6* (MIM 607108). In this subject, it is likely that the disruption of *PHF21A* contributed to the ID and craniofacial dysmorphism in common with the other two translocation subjects. The disparate phenotypes may then have been contributed by disruption of *ARHGAP6* or to an independent lesion elsewhere. Notably, mice homozygous for a targeted null mutation of *ARHGAP6* do not exhibit any detected phenotypic or behavioral abnormalities.<sup>236</sup> A strict *PHF21A* dosage requirement for normal craniofacial and neurodevelopment is also supported by our findings in the zebrafish system where

*phf21a* suppression produced abnormalities in head/face and jaw development along with increased neuronal apoptosis. Importantly, these deficiencies were rescued by the human protein PHF21A, indicating a conserved developmental function.

The discovery that translocations disrupting *PHF21A* are associated with abnormal craniofacial and intellectual development adds to the evidence that regulation of gene expression through chromatin modifications is crucial to both processes. To date, three genes, *NSD1* (MIM 606681), *KDM5C* (MIM 314690), *PHF8* (MIM 300560), encoding proteins that have PHD domains and bind methylated histone tails have been found in syndromes with ID and distinctive craniofacial features. *NSD1* protein, deficient in Sotos syndrome, specifically binds methylated H3K4 and H3K9 via domains PHD1, PHD4, PHD5, and PHD6,<sup>237</sup> *SMCX/JARID1C* protein (encoded by *KDM5C*) of nonsyndromic XLID binds histone H3K9me3 via its PHD1 domain,<sup>238</sup> and *PHF8* of Siderius-Hamel syndrome binds histone H3K4me3 via its PHD domain.<sup>239</sup> At least the latter two act as demethylases, targeting H3K4me2 and H3K4me3 in the case of *SMCX/JARID1C*,<sup>238</sup> and H3K9me1/2 and H4K20me1 in the case of *PHF8*.<sup>239,240</sup>

Unlike the above proteins, *PHF21A* is neither a methylase nor a demethylase, but instead specifically binds histone H3K4 when it is not methylated.<sup>212</sup> This suggests that both recognition of the unmodified state of histone tails and binding of proteins to methylated histone tails are critical for maintaining the appropriate balance and control of particular chromatin modifications to support normal intellectual and craniofacial development. While the PSS-associated ID and CFA appear to be due to haploinsufficiency of *PHF21A*, it has not been possible to ascertain and screen a large series of non-translocation subjects with comparable phenotypes for *PHF21A* mutations, so we cannot state with certainty that missense, nonsense, splicing or other mutations in *PHF21A* would lead to the same ID and CFA phenotypes. Indeed, it is conceivable that other types of genetic lesions in *PHF21A* could actually be associated with other developmental phenotypes, and we were able to identify 200 individuals with ID and/or CFA but without the full constellation of phenotypes exhibited by our translocation subjects and to perform a mutation screen of *PHF21A*. We did not detect any truncating or missense mutations that could aid in structure-function experiments, such as those that have implicated particular PHD domains in *NSD1* in binding to their methylated targets. This is not surprising given the frequency of these two major phenotypes and

58

the extent of genetic heterogeneity underlying each of them, but more extensive mutation analysis of subjects with various manifestations of ID, CFA and additional phenotypes seen in our translocation subjects may prove valuable to understanding the functional domains of PHF21A.

PHF21A (BHC80) is known to participate in the six-subunit BRAF-HDAC Complex (BHC) which also comprises BRAF35 (MIM 605535), HDAC1 (MIM 601241), HDAC2 (MIM 605164), CoREST (MIM 607675), and LSD1 (BHC110), where the latter is a histone demethylase that targets H3K4me2.<sup>228</sup> This complex interacts with the promoters of genes, such as synapsin and sodium channel genes, to mediate repression of these neuron-specific genes through the *cis*-regulatory element known as repressor element 1 or neural restrictive silencer (RE1/NRS).<sup>241</sup> Specific binding of PHF21A to H3K4me0 is required for optimal LSD1 promotor occupancy *in vivo* and for LSD1 mediated gene repression.<sup>212</sup> Our data showing derepression of the neural gene *SCN3A* in lymphoblasts from the translocation subjects as a consequence of reduced protein levels of PHF21A are consistent with this role for the protein. Repression of neuronal-specific genes is of fundamental importance in the development of both neuronal and non-neuronal tissues,<sup>241</sup> so the failure of this particular function in the translocation subjects might have contributed to their ID and CFA.

Another interesting XLID candidate, *ZMYM3* (MIM 300061), encodes a zinc finger protein predominantly expressed in brain which is a component of transcriptional co-repressor complexes that also contain LSD1 (BHC110) and HDAC2.<sup>242,243</sup> The 5' untranslated region of *ZMYM3* is disrupted by a presumably balanced t(X;13) in a female with ID and preferential inactivation of the normal X chromosome.<sup>244</sup> In addition, Kleefstra syndrome, characterized by ID and CFA comparable to PSS, has been associated with disruption of *EHMT1*, encoding euchromatin histone methyltransferase 1, which acts as a methyl transferase to modify H3K9 and has been reported as a component of the E2F6 transcription repressor complex and of a CtBP repressor complex that also contains LSD1.<sup>69,70,245,246</sup> The parallels between *PHF21A*, *ZMYM3* and *EHMT1* suggest that other X-linked and autosomal loci underlying ID and/or CFA may encode proteins that participate in complexes involving LSD1 or potentially other demethylases or methyltransferases. Our finding that decreased dosage for *PHF21A*, a histone binding protein that interacts with and is required for the histone demethylase

activity of LSD1, leads to both ID and CFA provides the proof of principle for investigation of other regulators of histone modification as genetic factors in ID and/or CFA. Indeed, the recent finding of haploinsufficiency for *ARID1B* (MIM 614556), encoding a component of an E3 ubiquitin ligase that functions with chromatin-remodeling complex SWI/SNF,<sup>247-251</sup> suggests that genes involved with other aspects of chromatin modification may also contribute to ID and CFA and that ultimately, human mutations affecting both regulatory and enzymatic components of histone modification complexes may represent important tools for delineating the features of chromatin regulation that are critical for normal craniofacial and neurological development and cognitive function.

## 7. Outlook and Conclusion

Identification of genes and types of mutations in monogenic disorders will provide substantial insight into the pathogenic mechanisms leading to delayed puberty, intellectual disability, craniofacial anomalies and thereby to identify potential targets for therapeutic intervention in the future. Our research seeks to further our understanding of human biology through the identification of genes and characterization of their naturally occurring mutations implicated in various human developmental disorders. Overall, the proposed studies will increase our understanding of gene's involvement in human sexual and intellectual development. It is likely that *FGFR1*, *CHD7*, and *WDR11* are central to sexual maturation and pubertal development and that *PHF21A* is critical for intellectual and craniofacial development.

Our finding that decreased dosage for *PHF21A*, a histone binding protein that interacts with and is required for the histone demethylase activity of LSD1, leads to both ID and CFA provides the proof of principle for investigation of other regulators of histone modification as genetic factors in ID and/or CFA. Indeed, the recent finding of haploinsufficiency for *ARID1B*, encoding a component of an E3 ubiquitin ligase that functions with chromatin-remodeling complex SWI/SNF,<sup>247-251</sup> suggests that genes involved in other aspects of chromatin modification may also contribute to ID and CFA and that ultimately, human mutations affecting both regulatory and enzymatic components of histone modification complexes may represent important tools for delineating the features of chromatin regulation that are critical for normal craniofacial and neurological development and cognitive function.

These studies clearly show the instrumental role of patients with *de novo* balanced translocations and the candidate gene approach in the disease gene discovery. The wealth of balanced translocation cases in MCN and DGAP and other consortiums could serve as a valuable resource and important starting point for the disease gene discovery projects that presents one of the major scientific challenges of the future.

## **8. Acknowledgements**

To the many colleagues who have contributed to this thesis directly and indirectly, including those whose names I have unwittingly forgotten, I owe a debt impossible to repay in words.

My sincere thanks are due to Prof. Hans-Hilger Ropers with admiration and gratitude for his important contributions in human genetics. To my supervisor Dr. Vera Kalscheuer who supported me from the beginning to the end for her patience and professional guidance; to the researchers at the time of my stay at Max-Planck-Institute of Molecular Genetics in Berlin for their hospitality; to Sabine Kübart for her immeasurable kindness and encouragement; to Dr. Hans Gerd Nothwang and Dr. Thomas Haaf for their expertise of Human Molecular Genetics; to Dr. Lawrence Layman for his unwavering faith in me; to my parents and two brothers for their emotional support; and to Jeong-Hi and my two children for more than I can say.

## 9. Literature

1. Wolfsberg, T.G., McEntyre, J. & Schuler, G.D. Guide to the draft human genome. *Nature* **409**, 824-6 (2001).
2. Chiang, A.P. *et al.* Homozygosity mapping with SNP arrays identifies TRIM32, an E3 ubiquitin ligase, as a Bardet-Biedl syndrome gene (BBS11). *Proc Natl Acad Sci U S A* **103**, 6287-92 (2006).
3. Swanson, C.P. *Cytology and cytogenetics*, (Prentice-Hall, Englewood Cliffs, N. J., 1957).
4. Gersen, S.L. & Keagle, M.B. *The principles of clinical cytogenetics*, xiii, 596 p. (Humana Press, Totowa, N.J., 2005).
5. Therman, E. & Susman, M. *Human chromosomes : structure, behavior, and effects*, xvii, 376 p. (Springer-Verlag, New York, 1993).
6. Gardner, R.J.M. & Sutherland, G.R. *Chromosome abnormalities and genetic counseling*, xviii, 577 p. (Oxford University Press, Oxford ; New York, 2004).
7. Bishop, A. & Cooke, P. *Human chromosomes; an illustrated introduction to human cytogenetics [by] Audrey Bishop and Patricia Cooke*, 56 p. (Heinemann, London., 1966).
8. Edelmann, L. *et al.* AT-rich palindromes mediate the constitutional t(11;22) translocation. *Am J Hum Genet* **68**, 1-13 (2001).
9. Kurahashi, H. & Emanuel, B.S. Long AT-rich palindromes and the constitutional t(11;22) breakpoint. *Hum Mol Genet* **10**, 2605-17 (2001).
10. Breg, W.R., Miller, D.A., Allderdice, P.W. & Miller, O.J. Identification of translocation chromosomes by quinacrine fluorescence. *Am J Dis Child* **123**, 561-4 (1972).
11. Jacobs, P.A. Correlation between euploid structural chromosome rearrangements and mental subnormality in humans. *Nature* **249**, 164-5 (1974).
12. Koenig, M. *et al.* Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell* **50**, 509-17 (1987).
13. Boyd, Y. *et al.* Muscular dystrophy in girls with X;autosome translocations. *J Med Genet* **23**, 484-90 (1986).
14. Lindenbaum, R.H., Clarke, G., Patel, C., Moncrieff, M. & Hughes, J.T. Muscular dystrophy in an X; 1 translocation female suggests that Duchenne locus is on X chromosome short arm. *J Med Genet* **16**, 389-92 (1979).
15. Burghes, A.H. *et al.* A cDNA clone from the Duchenne/Becker muscular dystrophy gene. *Nature* **328**, 434-7 (1987).
16. Collins, F.S. Positional cloning moves from perditional to traditional. *Nat Genet* **9**, 347-50 (1995).
17. Fryns, J.P., Kleczkowska, A., Kubien, E. & Van Den Berghe, H. On the excess of mental retardation and/or congenital malformations in apparently balanced reciprocal translocations. A critical review of the Leuven data 1966-1991. *Genet Couns* **2**, 185-94 (1991).
18. Baptista, J. *et al.* Breakpoint mapping and array CGH in translocations: comparison of a phenotypically normal and an abnormal cohort. *Am J Hum Genet* **82**, 927-36 (2008).
19. Warburton, D. De novo balanced chromosome rearrangements and extra marker chromosomes identified at prenatal diagnosis: clinical significance and distribution of breakpoints. *Am J Hum Genet* **49**, 995-1013 (1991).

20. Clifford, K., Rai, R., Watson, H. & Regan, L. An informative protocol for the investigation of recurrent miscarriage: preliminary experience of 500 consecutive cases. *Hum Reprod* **9**, 1328-32 (1994).
21. Raymond, F.L., Whibley, A., Stratton, M.R. & Gecz, J. Lessons learnt from large-scale exon re-sequencing of the X chromosome. *Hum Mol Genet* **18**, R60-4 (2009).
22. Gregg, C., Zhang, J., Butler, J.E., Haig, D. & Dulac, C. Sex-specific parent-of-origin allelic expression in the mouse brain. *Science* **329**, 682-5 (2010).
23. Tonon, G. *et al.* t(11;19)(q21;p13) translocation in mucoepidermoid carcinoma creates a novel fusion product that disrupts a Notch signaling pathway. *Nat Genet* **33**, 208-13 (2003).
24. Kim, H.G. *et al.* WDR11, a WD protein that interacts with transcription factor EMX1, is mutated in idiopathic hypogonadotropic hypogonadism and Kallmann syndrome. *Am J Hum Genet* **87**, 465-79 (2010).
25. Schinzel, A. *et al.* Kallmann syndrome in a boy with a t(1;10) translocation detected by reverse chromosome painting. *J Med Genet* **32**, 957-61 (1995).
26. Cox, J.J., Holden, S.T., Dee, S., Burbridge, J.I. & Raymond, F.L. Identification of a 650 kb duplication at the X chromosome breakpoint in a patient with 46,X,t(X;8)(q28;q12) and non-syndromic mental retardation. *J Med Genet* **40**, 169-74 (2003).
27. Gribble, S.M. *et al.* The complex nature of constitutional de novo apparently balanced translocations in patients presenting with abnormal phenotypes. *J Med Genet* **42**, 8-16 (2005).
28. De Gregori, M. *et al.* Cryptic deletions are a common finding in "balanced" reciprocal and complex chromosome rearrangements: a study of 59 cases. *J Med Genet* (2007).
29. Wakamatsu, N. *et al.* Mutations in SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. *Nat Genet* **27**, 369-70 (2001).
30. Beysen, D. *et al.* Deletions involving long-range conserved nongenic sequences upstream and downstream of FOXL2 as a novel disease-causing mechanism in blepharophimosis syndrome. *Am J Hum Genet* **77**, 205-18 (2005).
31. Fernandez, B.A., Siegel-Bartelt, J., Herbrick, J.A., Teshima, I. & Scherer, S.W. Holoprosencephaly and cleidocranial dysplasia in a patient due to two position-effect mutations: case report and review of the literature. *Clin Genet* **68**, 349-59 (2005).
32. Bakkaloglu, B. *et al.* Molecular cytogenetic analysis and resequencing of contactin associated protein-like 2 in autism spectrum disorders. *Am J Hum Genet* **82**, 165-73 (2008).
33. Bhalla, K. *et al.* Alterations in CDH15 and KIRREL3 in patients with mild to severe intellectual disability. *Am J Hum Genet* **83**, 703-13 (2008).
34. Wilmot, P.L., Shapiro, L.R. & Casamassima, A.C. Disomic balanced reciprocal translocation. *Clin Genet* **38**, 126-7 (1990).
35. Martinet, D. *et al.* Fetus with two identical reciprocal translocations: description of a rare complication of consanguinity. *Am J Med Genet A* **140**, 769-74 (2006).
36. Zaki, M. *et al.* Identification of a novel recessive RELN mutation using a homozygous balanced reciprocal translocation. *Am J Med Genet A* **143A**, 939-44 (2007).
37. Baala, L. *et al.* Homozygous silencing of T-box transcription factor EOMES leads to microcephaly with polymicrogyria and corpus callosum agenesis. *Nat Genet* **39**, 454-6 (2007).

38. Schneider, E. *et al.* Homozygous disruption of PDZD7 by reciprocal translocation in a consanguineous family: a new member of the Usher syndrome protein interactome causing congenital hearing impairment. *Hum Mol Genet* **18**, 655-66 (2009).
39. Leschot, N.J., von den Velden, J., Marinkovic-Ilsen, A., Darling, S.M. & Nijenhuis, L.E. Homozygosity for a Y/22 chromosome translocation: t(Y;22) (q12;p12/13). *Clin Genet* **29**, 251-7 (1986).
40. Alley, T.L. *et al.* Identification of a yeast artificial chromosome clone spanning a translocation breakpoint at 7q32.1 in a Smith-Lemli-Opitz syndrome patient. *Am J Hum Genet* **56**, 1411-6 (1995).
41. Hearn, T. *et al.* Mutation of ALMS1, a large gene with a tandem repeat encoding 47 amino acids, causes Alstrom syndrome. *Nat Genet* **31**, 79-83 (2002).
42. Kuechler, A. *et al.* An unbalanced translocation unmasks a recessive mutation in the follicle-stimulating hormone receptor (FSHR) gene and causes FSH resistance. *Eur J Hum Genet* **18**, 656-61 (2010).
43. Crisponi, L. *et al.* The putative forkhead transcription factor FOXL2 is mutated in blepharophimosis/ptosis/epicanthus inversus syndrome. *Nat Genet* **27**, 159-166 (2001).
44. Rose, C.S., Patel, P., Reardon, W., Malcolm, S. & Winter, R.M. The TWIST gene, although not disrupted in Saethre-Chotzen patients with apparently balanced translocations of 7p21, is mutated in familial and sporadic cases. *Hum Mol Genet* **6**, 1369-73 (1997).
45. Ishikawa-Brush, Y. *et al.* Autism and multiple exostoses associated with an X;8 translocation occurring within the GRPR gene and 3' to the SDC2 gene. *Hum Mol Genet* **6**, 1241-50 (1997).
46. Tarpey, P.S. *et al.* Mutations in the gene encoding the Sigma 2 subunit of the adaptor protein 1 complex, AP1S2, cause X-linked mental retardation. *Am J Hum Genet* **79**, 1119-24 (2006).
47. Crisponi, L. *et al.* FOXL2 inactivation by a translocation 171 kb away: analysis of 500 kb of chromosome 3 for candidate long-range regulatory sequences. *Genomics* **83**, 757-64 (2004).
48. Belloni, E. *et al.* Identification of Sonic hedgehog as a candidate gene responsible for holoprosencephaly. *Nat Genet* **14**, 353-6 (1996).
49. Ton, C.C. *et al.* Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. *Cell* **67**, 1059-74 (1991).
50. Foster, J.W. *et al.* Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature* **372**, 525-30 (1994).
51. Velagaleti, G.V. *et al.* Position effects due to chromosome breakpoints that map approximately 900 Kb upstream and approximately 1.3 Mb downstream of SOX9 in two patients with campomelic dysplasia. *Am J Hum Genet* **76**, 652-62 (2005).
52. Bedell, M.A., Jenkins, N.A. & Copeland, N.G. Good genes in bad neighbourhoods. *Nat Genet* **12**, 229-32 (1996).
53. Dlugaszewska, B. *et al.* Breakpoints around the HOXD cluster result in various limb malformations. *J Med Genet* **43**, 111-8 (2006).
54. Jensen, L.R. *et al.* Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation. *Am J Hum Genet* **76**, 227-36 (2005).
55. Shoubridge, C. *et al.* Mutations in the guanine nucleotide exchange factor gene IQSEC2 cause nonsyndromic intellectual disability. *Nat Genet* **42**, 486-8 (2010).

56. Deardorff, M.A. *et al.* Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of cornelia de Lange syndrome with predominant mental retardation. *Am J Hum Genet* **80**, 485-94 (2007).
57. Froyen, G. *et al.* Submicroscopic duplications of the hydroxysteroid dehydrogenase HSD17B10 and the E3 ubiquitin ligase HUWE1 are associated with mental retardation. *Am J Hum Genet* **82**, 432-43 (2008).
58. Laumonnier, F. *et al.* Mutations in PHF8 are associated with X linked mental retardation and cleft lip/cleft palate. *J Med Genet* **42**, 780-6 (2005).
59. Hertz, J.M. *et al.* Early onset, non-progressive, mild cerebellar ataxia co-segregating with a familial balanced translocation t(8;20)(p22;q13). *J Med Genet* **41**, e25 (2004).
60. Kim, H.G. *et al.* Disruption of neurexin 1 associated with autism spectrum disorder. *Am J Hum Genet* **82**, 199-207 (2008).
61. Genuardi, M. *et al.* Split hand/split foot anomaly in a family segregating a balanced translocation with breakpoint on 7q22.1. *Am J Med Genet* **47**, 823-31 (1993).
62. Moller, R.S. *et al.* Truncation of the Down syndrome candidate gene DYRK1A in two unrelated patients with microcephaly. *Am J Hum Genet* **82**, 1165-70 (2008).
63. Kalscheuer, V.M. *et al.* Disruption of the serine/threonine kinase 9 gene causes severe X-linked infantile spasms and mental retardation. *Am J Hum Genet* **72**, 1401-11 (2003).
64. Kalscheuer, V.M. *et al.* Mutations in autism susceptibility candidate 2 (AUTS2) in patients with mental retardation. *Hum Genet* **121**, 501-9 (2007).
65. Krantz, I.D. *et al.* Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of *Drosophila melanogaster* Nipped-B. *Nat Genet* **36**, 631-5 (2004).
66. Nopola-Hemmi, J. *et al.* Two translocations of chromosome 15q associated with dyslexia. *J Med Genet* **37**, 771-5 (2000).
67. Taipale, M. *et al.* A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in brain. *Proc Natl Acad Sci U S A* **100**, 11553-8 (2003).
68. Fernandez, T. *et al.* Disruption of contactin 4 (CNTN4) results in developmental delay and other features of 3p deletion syndrome. *Am J Hum Genet* **74**, 1286-93 (2004).
69. Kleefstra, T. *et al.* Loss-of-function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome. *Am J Hum Genet* **79**, 370-7 (2006).
70. Kleefstra, T. *et al.* Disruption of the gene Euchromatin Histone Methyl Transferase1 (Eu-HMTase1) is associated with the 9q34 subtelomeric deletion syndrome. *J Med Genet* **42**, 299-306 (2005).
71. Boland, E. *et al.* Mapping of deletion and translocation breakpoints in 1q44 implicates the serine/threonine kinase AKT3 in postnatal microcephaly and agenesis of the corpus callosum. *Am J Hum Genet* **81**, 292-303 (2007).
72. Bonaglia, M.C. *et al.* Disruption of the ProSAP2 gene in a t(12;22)(q24.1;q13.3) is associated with the 22q13.3 deletion syndrome. *Am J Hum Genet* **69**, 261-8 (2001).
73. Thienpont, B. *et al.* Haploinsufficiency of TAB2 causes congenital heart defects in humans. *Am J Hum Genet* **86**, 839-49 (2010).
74. Nishimura, D.Y. *et al.* The forkhead transcription factor gene FKHL7 is responsible for glaucoma phenotypes which map to 6p25. *Nat Genet* **19**, 140-7 (1998).
75. Tonkin, E.T., Wang, T.J., Lisgo, S., Bamshad, M.J. & Strachan, T. NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. *Nat Genet* **36**, 636-41 (2004).

76. Laumonnier, F. *et al.* Transcription factor SOX3 is involved in X-linked mental retardation with growth hormone deficiency. *Am J Hum Genet* **71**, 1450-5 (2002).
77. Fantes, J. *et al.* Mutations in SOX2 cause anophthalmia. *Nat Genet* **33**, 461-3 (2003).
78. Kishino, T., Lalande, M. & Wagstaff, J. UBE3A/E6-AP mutations cause Angelman syndrome. *Nat Genet* **15**, 70-3 (1997).
79. Conroy, J.M. *et al.* Balanced translocation 46,XY,t(2;15)(q37.2;q11.2) associated with atypical Prader-Willi syndrome. *Am J Hum Genet* **61**, 388-94 (1997).
80. Ray, P.N. *et al.* Cloning of the breakpoint of an X;21 translocation associated with Duchenne muscular dystrophy. *Nature* **318**, 672-5 (1985).
81. Verellen-Dumoulin, C. *et al.* Expression of an X-linked muscular dystrophy in a female due to translocation involving Xp21 and non-random inactivation of the normal X chromosome. *Hum Genet* **67**, 115-9 (1984).
82. Gusella, J.F. *et al.* A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* **306**, 234-8 (1983).
83. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell* **72**, 971-83 (1993).
84. Bodrug, S.E. *et al.* Prenatal identification of a girl with a t(X;4)(p21;q35) translocation: molecular characterisation, paternal origin, and association with muscular dystrophy. *J Med Genet* **27**, 426-32 (1990).
85. Bodrug, S.E., Holden, J.J., Ray, P.N. & Worton, R.G. Molecular analysis of X-autosome translocations in females with Duchenne muscular dystrophy. *EMBO J* **10**, 3931-9 (1991).
86. Robinson, D.O. *et al.* The parental origin of de novo X-autosome translocations in females with Duchenne muscular dystrophy revealed by M27 beta methylation analysis. *Genet Res* **56**, 135-40 (1990).
87. Kean, V.M. *et al.* Paternal inheritance of translocation chromosomes in a t(X;21) patient with X linked muscular dystrophy. *J Med Genet* **23**, 491-3 (1986).
88. Hagemeyer, A., Hoovers, J., Smit, E.M. & Bootsma, D. Replication pattern of the X chromosomes in three X/autosomal translocations. *Cytogenet Cell Genet* **18**, 333-48 (1977).
89. Jones, C. *et al.* Bilateral retinoblastoma in a male patient with an X; 13 translocation: evidence for silencing of the RB1 gene by the spreading of X inactivation. *Am J Hum Genet* **60**, 1558-62 (1997).
90. Kumar, D. & Weatherall, D.J. *Genomics and clinical medicine*, xix, 651 p. (Oxford University Press, Oxford ; New York, 2008).
91. Mattei, M.G., Mattei, J.F., Ayme, S. & Giraud, F. X-autosome translocations: cytogenetic characteristics and their consequences. *Hum Genet* **61**, 295-309 (1982).
92. Waters, J.J., Campbell, P.L., Crocker, A.J. & Campbell, C.M. Phenotypic effects of balanced X-autosome translocations in females: a retrospective survey of 104 cases reported from UK laboratories. *Hum Genet* **108**, 318-27 (2001).
93. Vervoort, V.S. *et al.* AGTR2 mutations in X-linked mental retardation. *Science* **296**, 2401-3 (2002).
94. Leisti, J.T., Kaback, M.M. & Rimoin, D.L. Human X-autosome translocations: differential inactivation of the X chromosome in a kindred with an X-9 translocation. *Am J Hum Genet* **27**, 441-53 (1975).
95. Kim, H.G., Bhagavath, B. & Layman, L.C. Clinical Manifestations of Impaired GnRH Neuron Development and Function. *Neurosignals* **16**, 165-182 (2008).

96. Bhagavath, B. *et al.* Clinical and molecular characterization of a large sample of patients with hypogonadotropic hypogonadism. *Fertil Steril* **85**, 706-13 (2006).
97. Knudson, A.G. *Genetics and disease*, ix, 294 p. (McGraw-Hill, New York,, 1965).
98. Franco, B. *et al.* A gene deleted in Kallmann's syndrome shares homology with neural cell adhesion and axonal path-finding molecules. *Nature* **353**, 529-36 (1991).
99. Oliveira, L.M. *et al.* The importance of autosomal genes in Kallmann syndrome: genotype-phenotype correlations and neuroendocrine characteristics. *J Clin Endocrinol Metab* **86**, 1532-8. (2001).
100. Kim, H.G. *et al.* Hypogonadotropic hypogonadism and cleft lip and palate caused by a balanced translocation producing haploinsufficiency for FGFR1. *J Med Genet* **42**, 666-72 (2005).
101. Dode, C. *et al.* Loss-of-function mutations in FGFR1 cause autosomal dominant Kallmann syndrome. *Nat Genet* **33**, 463-5 (2003).
102. Legouis, R. *et al.* The candidate gene for the X-linked Kallmann syndrome encodes a protein related to adhesion molecules. *Cell* **67**, 423-35 (1991).
103. Best, L.G., Wasdahl, W.A., Larson, L.M. & Sturlaugson, J. Chromosome abnormality in Kallmann syndrome. *Am J Med Genet* **35**, 306-9 (1990).
104. Kikuchi, I. *et al.* Chromosomal translocation t(13;16) in a patient with idiopathic hypogonadotropic hypogonadism. *Intern Med* **32**, 465-7 (1993).
105. Brioude, F. *et al.* Non-syndromic congenital hypogonadotropic hypogonadism: clinical presentation and genotype-phenotype relationships. *Eur J Endocrinol* **162**, 835-51 (2010).
106. Pitteloud, N. *et al.* Mutations in fibroblast growth factor receptor 1 cause both Kallmann syndrome and normosmic idiopathic hypogonadotropic hypogonadism. *Proc Natl Acad Sci U S A* **103**, 6281-6 (2006).
107. de Roux, N. *et al.* Hypogonadotropic hypogonadism due to loss of function of the KiSS1-derived peptide receptor GPR54. *Proc Natl Acad Sci U S A* **100**, 10972-6 (2003).
108. Seminara, S.B. *et al.* The GPR54 gene as a regulator of puberty. *N Engl J Med* **349**, 1614-27 (2003).
109. Topaloglu, A.K. *et al.* TAC3 and TACR3 mutations in familial hypogonadotropic hypogonadism reveal a key role for Neurokinin B in the central control of reproduction. *Nat Genet* (2008).
110. de Roux, N. *et al.* A family with hypogonadotropic hypogonadism and mutations in the gonadotropin-releasing hormone receptor. *N Engl J Med* **337**, 1597-1602 (1997).
111. Layman, L.C. *et al.* Mutations in gonadotropin-releasing hormone receptor gene cause hypogonadotropic hypogonadism. *Nat Genet* **18**, 14-5 (1998).
112. Miura, K., Acierno, J.S., Jr. & Seminara, S.B. Characterization of the human nasal embryonic LHRH factor gene, NELF, and a mutation screening among 65 patients with idiopathic hypogonadotropic hypogonadism (IHH). *J Hum Genet* **49**, 265-8 (2004).
113. Kim, H.G. *et al.* Mutations in CHD7, encoding a chromatin-remodeling protein, cause idiopathic hypogonadotropic hypogonadism and Kallmann syndrome. *Am J Hum Genet* **83**, 511-9 (2008).
114. Falardeau, J. *et al.* Decreased FGF8 signaling causes deficiency of gonadotropin-releasing hormone in humans and mice. *J Clin Invest* **118**, 2822-31 (2008).
115. Bouligand, J. *et al.* Isolated familial hypogonadotropic hypogonadism and a GNRH1 mutation. *N Engl J Med* **360**, 2742-8 (2009).
116. Chan, Y.M. *et al.* GNRH1 mutations in patients with idiopathic hypogonadotropic hypogonadism. *Proc Natl Acad Sci U S A* **106**, 11703-8 (2009).

117. Dode, C. *et al.* Kallmann syndrome: mutations in the genes encoding prokineticin-2 and prokineticin receptor-2. *PLoS Genet* **2**, e175 (2006).
118. Pitteloud, N. *et al.* Digenic mutations account for variable phenotypes in idiopathic hypogonadotropic hypogonadism. *J Clin Invest* **117**, 457-63 (2007).
119. Cole, L.W. *et al.* Mutations in prokineticin 2 and prokineticin receptor 2 genes in human gonadotrophin-releasing hormone deficiency: molecular genetics and clinical spectrum. *J Clin Endocrinol Metab* **93**, 3551-9 (2008).
120. Quaynor, S.D. *et al.* The prevalence of digenic mutations in patients with normosmic hypogonadotropic hypogonadism and Kallmann syndrome. *Fertil Steril* (2011).
121. Sykiotis, G.P. *et al.* Oligogenic basis of isolated gonadotropin-releasing hormone deficiency. *Proc Natl Acad Sci U S A* **107**, 15140-4 (2010).
122. Bhagavath, B. *et al.* The prevalence of gonadotropin-releasing hormone receptor mutations in a large cohort of patients with hypogonadotropic hypogonadism. *Fertil Steril* **84**, 951-7 (2005).
123. Silveira, L.F., Trarbach, E.B. & Latronico, A.C. Genetics basis for GnRH-dependent pubertal disorders in humans. *Mol Cell Endocrinol* **324**, 30-8 (2010).
124. Falardeau, J. *et al.* Decreased FGF8 signaling causes deficiency of gonadotropin-releasing hormone in humans and mice. *J Clin Invest* **118**, 2822-31 (2008).
125. Crowley, W.F., Jr., Filicori, M., Spratt, D.I. & Santoro, N.F. The physiology of gonadotropin-releasing hormone (GnRH) secretion in men and women. *Recent Prog Horm Res* **41**, 473-531 (1985).
126. Vermeulen, S. *et al.* Kallmann syndrome in a patient with congenital spherocytosis and an interstitial 8p11.2 deletion. *Am J Med Genet* **108**, 315-8 (2002).
127. Anderson, J., Burns, H.D., Enriquez-Harris, P., Wilkie, A.O. & Heath, J.K. Apert syndrome mutations in fibroblast growth factor receptor 2 exhibit increased affinity for FGF ligand. *Hum Mol Genet* **7**, 1475-83 (1998).
128. Van Dissel-Emiliani, F.M., De Boer-Brouwer, M. & De Rooij, D.G. Effect of fibroblast growth factor-2 on Sertoli cells and gonocytes in coculture during the perinatal period. *Endocrinology* **137**, 647-54 (1996).
129. Wilkie, A.O. & Morriss-Kay, G.M. Genetics of craniofacial development and malformation. *Nat Rev Genet* **2**, 458-68 (2001).
130. Bosman, E.A. *et al.* Multiple mutations in mouse Chd7 provide models for CHARGE syndrome. *Hum Mol Genet* **14**, 3463-76 (2005).
131. Hurd, E.A. *et al.* Loss of Chd7 function in gene-trapped reporter mice is embryonic lethal and associated with severe defects in multiple developing tissues. *Mamm Genome* **18**, 94-104 (2007).
132. Sanlaville, D. *et al.* Phenotypic spectrum of CHARGE syndrome in fetuses with CHD7 truncating mutations correlates with expression during human development. *J Med Genet* **43**, 211-217 (2006).
133. Wheeler, P.G., Quigley, C.A., Sadeghi-Nejad, A. & Weaver, D.D. Hypogonadism and CHARGE association. *Am J Med Genet* **94**, 228-31 (2000).
134. Chalouhi, C. *et al.* Olfactory evaluation in children: application to the CHARGE syndrome. *Pediatrics* **116**, e81-8 (2005).
135. Pinto, G. *et al.* CHARGE syndrome includes hypogonadotropic hypogonadism and abnormal olfactory bulb development. *J Clin Endocrinol Metab* **90**, 5621-6 (2005).
136. Ogata, T. *et al.* Kallmann syndrome phenotype in a female patient with CHARGE syndrome and CHD7 mutation. *Endocr J* **53**, 741-3 (2006).

137. Mutoh, A., Sasagawa, I., Tateno, T., Sawamura, T. & Nakada, T. Long arm deletion of chromosome 10 in a boy with monorchidism. *Scand J Urol Nephrol* **33**, 77-8 (1999).
138. Suzuki, Y., Sasagawa, I., Nakada, T. & Onmura, Y. Bilateral cryptorchidism associated with terminal deletion of 10q. in *Urol Int* 1999/02/06 edn Vol. 61 186-7 (1998).
139. Leonard, N.J., Harley, F.L. & Lin, C.C. Terminal deletion of chromosome 10q at band 26.1: follow-up in an adolescent male with high-output renal failure from congenital obstructive uropathy. *Am J Med Genet* **86**, 115-7 (1999).
140. Baccetti, B. *et al.* 10, 15 reciprocal translocation in an infertile man: ultrastructural and fluorescence in-situ hybridization sperm study: case report. *Hum Reprod* **18**, 2302-8 (2003).
141. Ogata, T. *et al.* Genetic evidence for a novel gene(s) involved in urogenital development on 10q26. *Kidney Int* **58**, 2281-90 (2000).
142. Lukusa, T. & Fryns, J.P. Pure distal monosomy 10q26 in a patient displaying clinical features of Prader-Willi syndrome during infancy and distinct behavioural phenotype in adolescence. *Genet Couns* **11**, 119-26 (2000).
143. Bofinger, M.K. *et al.* A familial MCA/MR syndrome due to translocation t(10;16) (q26;p13.1): report of six cases. *Am J Med Genet* **38**, 1-8 (1991).
144. Higgins, A.W. *et al.* Characterization of apparently balanced chromosomal rearrangements from the developmental genome anatomy project. *Am J Hum Genet* **82**, 712-22 (2008).
145. Carson-Walter, E.B. *et al.* Cell surface tumor endothelial markers are conserved in mice and humans. *Cancer Res* **61**, 6649-55 (2001).
146. Chen, H., Ishii, A., Wong, W.K., Chen, L.B. & Lo, S.H. Molecular characterization of human tensin. *Biochem J* **351 Pt 2**, 403-11 (2000).
147. Plotnikov, A.N., Schlessinger, J., Hubbard, S.R. & Mohammadi, M. Structural basis for FGF receptor dimerization and activation. *Cell* **98**, 641-50 (1999).
148. Sato, N. *et al.* Clinical assessment and mutation analysis of Kallmann syndrome 1 (KAL1) and fibroblast growth factor receptor 1 (FGFR1, or KAL2) in five families and 18 sporadic patients. *J Clin Endocrinol Metab* **89**, 1079-88 (2004).
149. Hardelin, J.P. Kallmann syndrome: towards molecular pathogenesis. *Mol Cell Endocrinol* **179**, 75-81 (2001).
150. Cariboni, A. *et al.* The product of X-linked Kallmann's syndrome gene (KAL1) affects the migratory activity of gonadotropin-releasing hormone (GnRH)-producing neurons. *Hum Mol Genet* **13**, 2781-91 (2004).
151. Ornitz, D.M. FGFs, heparan sulfate and FGFRs: complex interactions essential for development. *BioEssays : news and reviews in molecular, cellular and developmental biology* **22**, 108-12 (2000).
152. Schlessinger, J. *et al.* Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol Cell* **6**, 743-50 (2000).
153. Dode, C. & Hardelin, J.P. Kallmann syndrome: fibroblast growth factor signaling insufficiency? *J Mol Med* **82**, 725-34 (2004).
154. Muenke, M. *et al.* A common mutation in the fibroblast growth factor receptor 1 gene in Pfeiffer syndrome. *Nat Genet* **8**, 269-74 (1994).
155. Roscioli, T. *et al.* Clinical findings in a patient with FGFR1 P252R mutation and comparison with the literature. *Am J Med Genet* **93**, 22-8 (2000).
156. Xiao, S. *et al.* FGFR1 is fused with a novel zinc-finger gene, ZNF198, in the t(8;13) leukaemia/lymphoma syndrome. *Nature genetics* **18**, 84-7 (1998).

157. Popovici, C. *et al.* Fibroblast growth factor receptor 1 is fused to FIM in stem-cell myeloproliferative disorder with t(8;13). *Proc Natl Acad Sci U S A* **95**, 5712-7 (1998).
158. Guasch, G. *et al.* FGFR1 is fused to the centrosome-associated protein CEP110 in the 8p12 stem cell myeloproliferative disorder with t(8;9)(p12;q33). *Blood* **95**, 1788-96 (2000).
159. Demiroglu, A. *et al.* The t(8;22) in chronic myeloid leukemia fuses BCR to FGFR1: transforming activity and specific inhibition of FGFR1 fusion proteins. *Blood* **98**, 3778-83 (2001).
160. Radovick, S. *et al.* Migratory arrest of gonadotropin-releasing hormone neurons in transgenic mice. *Proc Natl Acad Sci U S A* **88**, 3402-6 (1991).
161. Mellon, P.L. *et al.* Immortalization of hypothalamic GnRH neurons by genetically targeted tumorigenesis. *Neuron* **5**, 1-10 (1990).
162. Marfella, C.G. & Imbalzano, A.N. The Chd family of chromatin remodelers. *Mutat Res* **618**, 30-40 (2007).
163. Bannister, A.J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120-4 (2001).
164. Flanagan, J.F. *et al.* Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* **438**, 1181-5 (2005).
165. Jongmans, M.C. *et al.* Familial CHARGE syndrome and the CHD7 gene: a recurrent missense mutation, intrafamilial recurrence and variability. *Am J Med Genet A* **146A**, 43-50 (2008).
166. Delahaye, A. *et al.* Familial CHARGE syndrome because of CHD7 mutation: clinical intra- and interfamilial variability. *Clin Genet* **72**, 112-21 (2007).
167. Thoma, N.H. *et al.* Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54. *Nat Struct Mol Biol* **12**, 350-6 (2005).
168. Pitteloud, N. *et al.* Loss-of-function mutation in the prokineticin 2 gene causes Kallmann syndrome and normosmic idiopathic hypogonadotropic hypogonadism. *Proc Natl Acad Sci U S A* **104**, 17447-52 (2007).
169. Schwarting, G.A., Wierman, M.E. & Tobet, S.A. Gonadotropin-releasing hormone neuronal migration. *Semin Reprod Med* **25**, 305-12 (2007).
170. Bouazoune, K. *et al.* The dMi-2 chromodomains are DNA binding modules important for ATP-dependent nucleosome mobilization. *Embo J* **21**, 2430-40 (2002).
171. Ng, P.C. & Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* **31**, 3812-4 (2003).
172. Blake, K.D., Salem-Hartshorne, N., Daoud, M.A. & Gradstein, J. Adolescent and adult issues in CHARGE syndrome. *Clin Pediatr (Phila)* **44**, 151-9 (2005).
173. Vissers, L.E. *et al.* Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet* **36**, 955-7 (2004).
174. Chernova, O.B. *et al.* A novel member of the WD-repeat gene family, WDR11, maps to the 10q26 region and is disrupted by a chromosome translocation in human glioblastoma cells. *Oncogene* **20**, 5378-92 (2001).
175. Ojeda, S.R. *et al.* Gene networks and the neuroendocrine regulation of puberty. *Mol Cell Endocrinol* **324**, 3-11 (2010).
176. Parent, A.S. *et al.* Gene expression profiling of hypothalamic hamartomas: a search for genes associated with central precocious puberty. *Horm Res* **69**, 114-23 (2008).
177. Lee, J.H. *et al.* KiSS-1, a novel human malignant melanoma metastasis-suppressor gene. *J Natl Cancer Inst* **88**, 1731-7 (1996).

178. Lee, J.H. & Welch, D.R. Identification of highly expressed genes in metastasis-suppressed chromosome 6/human malignant melanoma hybrid cells using subtractive hybridization and differential display. *Int J Cancer* **71**, 1035-44 (1997).
179. Lee, J.H. & Welch, D.R. Suppression of metastasis in human breast carcinoma MDA-MB-435 cells after transfection with the metastasis suppressor gene, KiSS-1. *Cancer Res* **57**, 2384-7 (1997).
180. Ohtaki, T. *et al.* Metastasis suppressor gene KiSS-1 encodes peptide ligand of a G-protein-coupled receptor. *Nature* **411**, 613-7 (2001).
181. Philipps, D.L. *et al.* The dual bromodomain and WD repeat-containing mouse protein BRWD1 is required for normal spermiogenesis and the oocyte-embryo transition. *Dev Biol* **317**, 72-82 (2008).
182. Field, M. *et al.* Mutations in the BRWD3 gene cause X-linked mental retardation associated with macrocephaly. *Am J Hum Genet* **81**, 367-74 (2007).
183. Higa, L.A. & Zhang, H. Stealing the spotlight: CUL4-DDB1 ubiquitin ligase docks WD40-repeat proteins to destroy. *Cell Div* **2**, 5 (2007).
184. Kudryashov, D.S. *et al.* The crystal structure of a cross-linked actin dimer suggests a detailed molecular interface in F-actin. *Proc Natl Acad Sci U S A* **102**, 13105-10 (2005).
185. Bishop, K.M., Garel, S., Nakagawa, Y., Rubenstein, J.L. & O'Leary, D.D. Emx1 and Emx2 cooperate to regulate cortical size, lamination, neuronal differentiation, development of cortical efferents, and thalamocortical pathfinding. *J Comp Neurol* **457**, 345-60 (2003).
186. Lichtneckert, R., Nobs, L. & Reichert, H. Empty spiracles is required for the development of olfactory projection neuron circuitry in Drosophila. *Development* **135**, 2415-24 (2008).
187. Theil, T., Alvarez-Bolado, G., Walter, A. & Ruther, U. Gli3 is required for Emx gene expression during dorsal telencephalon development. *Development* **126**, 3561-71 (1999).
188. Tole, S., Ragsdale, C.W. & Grove, E.A. Dorsoventral patterning of the telencephalon is disrupted in the mouse mutant extra-toes(J). *Dev Biol* **217**, 254-65 (2000).
189. Villavicencio, E.H., Walterhouse, D.O. & Iannaccone, P.M. The sonic hedgehog-patched-gli pathway in human development and disease. *Am J Hum Genet* **67**, 1047-54 (2000).
190. Shaffer, L.G., Hecht, J.T., Ledbetter, D.H. & Greenberg, F. Familial interstitial deletion 11(p11.2p12) associated with parietal foramina, brachymicrocephaly, and mental retardation. *Am J Med Genet* **45**, 581-3 (1993).
191. Potocki, L. & Shaffer, L.G. Interstitial deletion of 11(p11.2p12): a newly described contiguous gene deletion syndrome involving the gene for hereditary multiple exostoses (EXT2). *Am J Med Genet* **62**, 319-25 (1996).
192. Wakui, K. *et al.* Construction of a natural panel of 11p11.2 deletions and further delineation of the critical region involved in Potocki-Shaffer syndrome. *Eur J Hum Genet* **13**, 528-40 (2005).
193. Chen, G., Zhang, D., Feng, G., Liu, W. & He, L. A novel locus for parietal foramina maps to chromosome 4q21-q23. *J Hum Genet* **48**, 420-4 (2003).
194. Mavrogiannis, L.A. *et al.* Haploinsufficiency of the human homeobox gene ALX4 causes skull ossification defects. *Nat Genet* **27**, 17-8 (2001).
195. Wu, Y.Q. *et al.* Haploinsufficiency of ALX4 as a potential cause of parietal foramina in the 11p11.2 contiguous gene-deletion syndrome. *Am J Hum Genet* **67**, 1327-32 (2000).
196. Joshi, P.A., Chang, H. & Hamel, P.A. Loss of Alx4, a stromally-restricted homeodomain protein, impairs mammary epithelial morphogenesis. *Dev Biol* **297**, 284-94 (2006).

197. Stickens, D. *et al.* The EXT2 multiple exostoses gene defines a family of putative tumour suppressor genes. *Nat Genet* **14**, 25-32 (1996).
198. Wuyts, W. *et al.* Positional cloning of a gene involved in hereditary multiple exostoses. *Hum Mol Genet* **5**, 1547-57 (1996).
199. Raskind, W.H., Conrad, E.U., Chansky, H. & Matsushita, M. Loss of heterozygosity in chondrosarcomas for markers linked to hereditary multiple exostoses loci on chromosomes 8 and 11. *Am J Hum Genet* **56**, 1132-9 (1995).
200. Hecht, J.T. *et al.* Hereditary multiple exostosis and chondrosarcoma: linkage to chromosome II and loss of heterozygosity for EXT-linked markers on chromosomes II and 8. *Am J Hum Genet* **56**, 1125-31 (1995).
201. McCormick, C., Duncan, G., Goutsos, K.T. & Tufaro, F. The putative tumor suppressors EXT1 and EXT2 form a stable complex that accumulates in the Golgi apparatus and catalyzes the synthesis of heparan sulfate. *Proc Natl Acad Sci U S A* **97**, 668-73 (2000).
202. Lind, T., Tufaro, F., McCormick, C., Lindahl, U. & Lidholt, K. The putative tumor suppressors EXT1 and EXT2 are glycosyltransferases required for the biosynthesis of heparan sulfate. *J Biol Chem* **273**, 26265-8 (1998).
203. Tassabehji, M. *et al.* GTF2IRD1 in craniofacial development of humans and mice. *Science* **310**, 1184-7 (2005).
204. Gecz, J. The molecular basis of intellectual disability: novel genes with naturally occurring mutations causing altered gene expression in the brain. *Front Biosci* **9**, 1-7 (2004).
205. Hall, C.R., Wu, Y., Shaffer, L.G. & Hecht, J.T. Familial case of Potocki-Shaffer syndrome associated with microdeletion of EXT2 and ALX4. *Clin Genet* **60**, 356-9 (2001).
206. Mavrogiannis, L.A. *et al.* Enlarged parietal foramina caused by mutations in the homeobox genes ALX4 and MSX2: from genotype to phenotype. *Eur J Hum Genet* **14**, 151-8 (2006).
207. Wuyts, W. *et al.* Molecular and clinical examination of an Italian DEFECT11 family. *Eur J Hum Genet* **7**, 579-84 (1999).
208. O'Driscoll, M., Dobyns, W.B., van Hagen, J.M. & Jeggo, P.A. Cellular and clinical impact of haploinsufficiency for genes involved in ATR signaling. *Am J Hum Genet* **81**, 77-86 (2007).
209. Cheung, V.G. *et al.* Integration of cytogenetic landmarks into the draft sequence of the human genome. *Nature* **409**, 953-8 (2001).
210. Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A. & Lukyanov, S.A. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res* **23**, 1087-8 (1995).
211. Talkowski, M.E. *et al.* Next-generation sequencing strategies enable routine detection of balanced chromosome rearrangements for clinical diagnostics and genetic research. *Am J Hum Genet* **88**, 469-81 (2011).
212. Lan, F. *et al.* Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature* **448**, 718-22 (2007).
213. Dollfus, H. *et al.* Gillespie syndrome phenotype with a t(X;11)(p22.32;p12) de novo translocation. *Am J Ophthalmol* **125**, 397-9 (1998).
214. Fantes, J.A. *et al.* FISH mapping of de novo apparently balanced chromosome rearrangements identifies characteristics associated with phenotypic abnormality. *Am J Hum Genet* **82**, 916-26 (2008).

215. Romeike, B.F. & Wuyts, W. Proximal chromosome 11p contiguous gene deletion syndrome phenotype: case report and review of the literature. *Clin Neuropathol* **26**, 1-11 (2007).
216. Ferreira, M.A. *et al.* Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet* **40**, 1056-8 (2008).
217. O'Donovan, M.C. *et al.* Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet* **40**, 1053-5 (2008).
218. Purcell, S.M. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748-52 (2009).
219. Smith, E.N. *et al.* Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatr* **14**, 755-63 (2009).
220. Sullivan, P.F. *et al.* Genomewide association for schizophrenia in the CATIE study: results of stage 1. *Mol Psychiatr* **13**, 570-84 (2008).
221. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455**, 237-41 (2008).
222. Cooper, G.M. *et al.* A copy number variation morbidity map of developmental delay. *Nat Genet* **43**, 838-46 (2011).
223. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061-73 (2010).
224. Iwase, S. *et al.* A component of BRAF-HDAC complex, BHC80, is required for neonatal survival in mice. *FEBS Lett* **580**, 3129-35 (2006).
225. Kim, C.H. *et al.* Repressor activity of Headless/Tcf3 is essential for vertebrate head formation. *Nature* **407**, 913-6 (2000).
226. Choi, J. *et al.* FoxH1 negatively modulates flk1 gene expression and vascular formation in zebrafish. *Dev Biol* **304**, 735-44 (2007).
227. Schilling, T.F. *et al.* Jaw and branchial arch mutants in zebrafish I: branchial arches. *Development* **123**, 329-44 (1996).
228. Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941-53 (2004).
229. Shi, Y.J. *et al.* Regulation of LSD1 histone demethylase activity by its associated factors. *Mol Cell* **19**, 857-64 (2005).
230. Bartsch, O. *et al.* Delineation of a contiguous gene syndrome with multiple exostoses, enlarged parietal foramina, craniofacial dysostosis, and mental retardation, caused by deletions in the short arm of chromosome 11. *Am J Hum Genet* **58**, 734-42 (1996).
231. Wuyts, W. *et al.* Proximal 11p deletion syndrome (P11pDS): additional evaluation of the clinical and molecular aspects. *Eur J Hum Genet* **12**, 400-6 (2004).
232. Yamamoto, T., Akaboshi, S., Ninomiya, H. & Nanba, E. DEFECT 11 syndrome associated with agenesis of the corpus callosum. *J Med Genet* **38**, E5 (2001).
233. Wuyts, W. *et al.* The ALX4 homeobox gene is mutated in patients with ossification defects of the skull (foramina parietalia permagna, OMIM 168500). *J Med Genet* **37**, 916-20 (2000).
234. Ma, W.J. & Furneaux, H. Localization of the human HuR gene to chromosome 19p13.2. *Hum Genet* **99**, 32-3 (1997).
235. Ma, W.J., Cheng, S., Campbell, C., Wright, A. & Furneaux, H. Cloning and characterization of HuR, a ubiquitously expressed Elav-like protein. *J Biol Chem* **271**, 8144-51 (1996).

236. Prakash, S.K. *et al.* Functional analysis of ARHGAP6, a novel GTPase-activating protein for RhoA. *Hum Mol Genet* **9**, 477-88 (2000).
237. Pasillas, M.P., Shah, M. & Kamps, M.P. Nsd1 PHD domains bind methylated H3K4 and H3K9 using interactions disrupted by point mutations in human Sotos syndrome. *Hum Mutat* (2010).
238. Iwase, S. *et al.* The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. *Cell* **128**, 1077-88 (2007).
239. Feng, W., Yonezawa, M., Ye, J., Jenuwein, T. & Grummt, I. PHF8 activates transcription of rRNA genes through H3K4me3 binding and H3K9me1/2 demethylation. *Nat Struct Mol Biol* **17**, 445-50 (2010).
240. Qi, H.H. *et al.* Histone H4K20/H3K9 demethylase PHF8 regulates zebrafish brain and craniofacial development. *Nature* **466**, 503-7 (2010).
241. Hakimi, M.A. *et al.* A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. *Proc Natl Acad Sci U S A* **99**, 7420-5 (2002).
242. Hakimi, M.A., Dong, Y., Lane, W.S., Speicher, D.W. & Shiekhattar, R. A candidate X-linked mental retardation gene is a component of a new family of histone deacetylase-containing complexes. *J Biol Chem* **278**, 7234-9 (2003).
243. Scheer, M.P. *et al.* DXS6673E encodes a predominantly nuclear protein, and its mouse ortholog DXHXS6673E is alternatively spliced in a developmental- and tissue-specific manner. *Genomics* **63**, 123-32 (2000).
244. van der Maarel, S.M. *et al.* Cloning and characterization of DXS6673E, a candidate gene for X-linked mental retardation in Xq13.1. *Hum Mol Genet* **5**, 887-97 (1996).
245. Ogawa, H., Ishiguro, K., Gaubatz, S., Livingston, D.M. & Nakatani, Y. A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. *Science* **296**, 1132-6 (2002).
246. Shi, Y. *et al.* Coordinated histone modifications mediated by a CtBP co-repressor complex. *Nature* **422**, 735-8 (2003).
247. Halgren, C. *et al.* Corpus callosum abnormalities, intellectual disability, speech impairment, and autism in patients with haploinsufficiency of ARID1B. *Clin Genet* (2011).
248. Li, X.S., Trojer, P., Matsumura, T., Treisman, J.E. & Tanese, N. Mammalian SWI/SNF--a subunit BAF250/ARID1 is an E3 ubiquitin ligase that targets histone H2B. *Mol Cell Biol* **30**, 1673-88 (2010).
249. Hoyer, J. *et al.* Haploinsufficiency of ARID1B, a Member of the SWI/SNF-A Chromatin-Remodeling Complex, Is a Frequent Cause of Intellectual Disability. *Am J Hum Genet* **90**, 565-572 (2012).
250. Santen, G.W. *et al.* Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat Genet* **44**, 379-80 (2012).
251. Tsurusaki, Y. *et al.* Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet* **44**, 376-8 (2012).

## 10. Summary

We took advantage of the unique opportunity to locate genes of developmental importance provided by apparently balanced chromosomal rearrangements associated with phenotypic abnormalities. By positional cloning at or near the breakpoints, we aimed to identify the crucial disease genes whose functions were disrupted or dysregulated by chromosomal rearrangement.

This thesis describes the positional cloning of disease genes for Kallmann and Potocki-Shaffer syndromes from the breakpoint mapping to the genomic, bioinformatics, molecular and functional analyses, supporting the conclusions that *FGFR1* and *WDR11* cause Kallmann syndrome and *PHF21A* causes intellectual disability (ID) and craniofacial anomalies (CFA) in Potocki-Shaffer syndrome. Additionally, a candidate gene approach identified *CHD7* as a new Kallmann syndrome gene.

Specifically, we show that *FGFR1* is truncated by a translocation breakpoint at 8p11.2 and that haploinsufficiency is the likely underlying mechanism of Kallmann syndrome in this patient with 46,XY,t(7;8)(p12.3;p11.2)dn. By defining the chromosomal breakpoint of t(10;12)(q26.12;q13.11)dn from a subject with Kallmann syndrome and scanning genes in its vicinity in unrelated hypogonadal subjects, we have identified *WDR11* at 10q26.12 as another gene involved in human puberty. We discovered that *WDR11* interacts with *EMX1*, a homeodomain transcription factor involved in the development of olfactory neurons, and that missense mutations reduce or abolish this interaction. By candidate gene approach, we show that *CHD7* is mutated in patients with Kallmann syndrome, which represents a milder allelic variant of CHARGE syndrome.

Finally, through the characterization of two independent subjects with balanced translocations involving 11p11.2 and supportive comparative deletion mapping of Potocki-Shaffer syndrome patients with different phenotypes, we have discovered that the ID and CFA phenotypes of Potocki-Shaffer syndrome are both caused by haploinsufficiency of a single gene, *PHF21A*, at 11p11.2.

The research presented in this thesis underscores the instrumental role of constitutional balanced chromosomal translocations in the identification of monogenic disease genes by breakpoint mapping.

## 11. Zusammenfassung

Die systematische Untersuchung von Patienten mit balancierten Chromosomenveränderungen ist eine erfolgversprechende Strategie zur Identifizierung neuer, bisher unbekannter Krankheitsgene. Die hier vorliegende Arbeit beschreibt die Positionsklonierung von Genen für das Kallmann- und das Potocki-Shaffer-Syndrom durch Kartierung der Bruchpunkte und anschließende genomische, bioinformatische, molekulare und funktionelle Analysen. Diese Untersuchungen haben Mutationen in den Genen *FGFR1* und *WDR11* als molekulare Ursachen des Kallmann-Syndroms wahrscheinlich gemacht und *CHD7* als neues Gen für das Kallmann-Syndrom identifiziert. Darüberhinaus konnten wir nachweisen, dass Mutationen im *PHF21A*-Gen zur geistigen Behinderung mit kraniofazialen Auffälligkeiten führen, die für das Potocki-Shaffer-Syndrom charakteristisch sind.

Bei einem Träger einer *de novo*-Translokation zwischen den Chromosomen 7 und 8 (46,XY,t(7;8)(p12.3;p11.2)) konnten wir nachweisen, dass das *FGFR1*-Gen durch den Bruchpunkt im kurzen Arm von Chromosom 8 durchtrennt wird. Dies spricht dafür, dass eine Form des Kallmann-Syndroms durch Haploinsuffizienz von *FGFR1* verursacht wird.

Durch Untersuchung einer anderen, ebenfalls mit Kallmann-Syndrom assoziierten *de novo*-Translokation mit Bruchpunkten in den langen Armen der Chromosomen 10 und 12 (t(10;12)(q26.12;q13.11) und anschließende Suche nach Mutationen im Bereich 10q26.12 bei nicht verwandten Patienten haben wir *WDR11* als ein weiteres Gen für Kallmann-Syndrom identifiziert. Wir konnten zeigen, dass das *WDR11*-Protein mit dem Homeobox-Transkriptionsfaktor *EMX1* interagiert. Missense-Mutationen im *WDR11*-Gen können diese Interaktion schwächen oder sogar zu einem vollständigen Verlust der Interaktion führen. *EMX1* spielt eine wichtige Rolle in der Entwicklung von olfaktorischen Neuronen.

Durch gezielte Suche nach Mutationen in aussichtsreichen Kandidatengenen gelang es uns ausserdem, Defekte des *CHD7*-Gens als weitere Ursache des Kallmann-Syndroms ausfindig zu machen.

Im Rahmen dieser Doktorarbeit haben wir schliesslich auch die molekulare Ursache des Potocki-Shaffer-Syndroms aufgeklärt. Durch Kartierung von Translokationsbruchpunkten im kurzen Arm von Chromosom 11 (11p11.2) bei zwei nicht verwandten Patienten konnten wir zeigen, dass dieses Syndrom durch Inaktivierung des PHF21A-Gens verursacht wird, und dass die für dieses Syndrom charakteristische geistige Behinderung mit kraniofazialen Auffälligkeiten durch Haploinsuffizienz des PHF21A-Gens erklärt werden kann.

Die im Rahmen dieser Arbeit erzielten Ergebnisse sind ein weiterer Beleg für die Bedeutung von krankheitsassoziierten balanzierten Translokationen als sichtbare Brücke zwischen Genotyp und Phänotyp und als Schlüssel für die molekulare Aufklärung monogener Krankheiten.

## 12. Appendix

## 12.1 Paper I

**Hyung-Goo Kim**, Steven R. Herrick, Emma Lemyre, Shotaro Kishikawa, Joseph A. Salisz, Stephanie Seminara, Marcy E. MacDonald, Gail A. P. Bruns, Cynthia C. Morton, Brad J. Quade, James F. Gusella\*

\*Corresponding author

***Hypogonadotropic hypogonadism and cleft lip and palate caused by a balanced translocation producing haploinsufficiency for FGFR1***

Journal of Medical Genetics 2005 Aug;42(8):666-72.

## LETTER TO JMG

# Hypogonadotropic hypogonadism and cleft lip and palate caused by a balanced translocation producing haploinsufficiency for *FGFR1*

HG Kim, S R Herrick, E Lemyre, S Kishikawa, J A Salisz, S Seminara, M E MacDonald, G A P Bruns, C C Morton, B J Quade, J F Gusella

*J Med Genet* 2005;42:666–672. doi: 10.1136/jmg.2004.026989

We have established the Developmental Genome Anatomy Project (DGAP; [//dgap.harvard.edu](http://dgap.harvard.edu)) to take advantage of the unique opportunity to locate genes of developmental importance provided by apparently balanced chromosomal rearrangements associated with phenotypic abnormalities. By positional cloning at or near the breakpoints, we aim to identify the crucial disease genes whose functions have been disrupted by rearrangement.<sup>1</sup> Kallmann's syndrome (KS) is a developmental disorder characterised by anosmia resulting from agenesis of the olfactory lobes and hypogonadism secondary to deficiency of hypothalamic gonadotropin releasing hormone (GnRH). Its prevalence has been estimated at 1/10 000 in males and 1/50 000 in females. In a minority of cases there are inactivating mutations of *KALI*, an X linked gene encoding a putative adhesion molecule thought to mediate embryonic neuronal migration.<sup>2–3</sup> Constitutional autosomal chromosome translocations associated with KS have been reported, but the disrupted genes have not been identified.<sup>4–6</sup>

We have studied a white male subject with a de novo balanced translocation between chromosomes 7, in band p12.3, and 8, in band p11.2 (fig 1A), who was diagnosed on clinical examination to have hypogonadotropic hypogonadism (infantile testes), azoospermia, and cleft lip and palate, without frank anosmia. As a KS patient with a microdeletion involving the same 8p11.2 region had been reported, we sought to identify the chromosome 8 gene disrupted in this reciprocal translocation as a likely candidate for the cause of autosomal KS as well as of isolated hypogonadotropic hypogonadism.<sup>7</sup> While this breakpoint in *FGFR1* was being characterised, Dodé *et al* identified *FGFR1* mutations in several patients, establishing that disruption of *FGFR1* can cause autosomal dominant KS.<sup>8</sup>

## METHODS

This study was approved by the Institutional Review Board of Partners Healthcare Inc, encompassing both the Massachusetts General Hospital and the Brigham and Women's Hospital.

## Case report

The subject is a white man who was aged 24 years at the time of initial diagnosis. He had a history of cleft palate, corrected by surgery. He had no outstanding medical problems other than delayed sexual development and a feminine sounding voice. He had his growth spurt at age 18–19 years, developed sparse armpit hair at age 20, and penile hair at 16–17, but no penile or testicular enlargement. He displayed child-like facial hair, sparse axillary adult appearing hair, and prepubertal chest hair. Based on the presence of cleft palate and hypogonadism, a tentative diagnosis of Kallmann's syndrome was reached, though the subject did not complain

## Key points

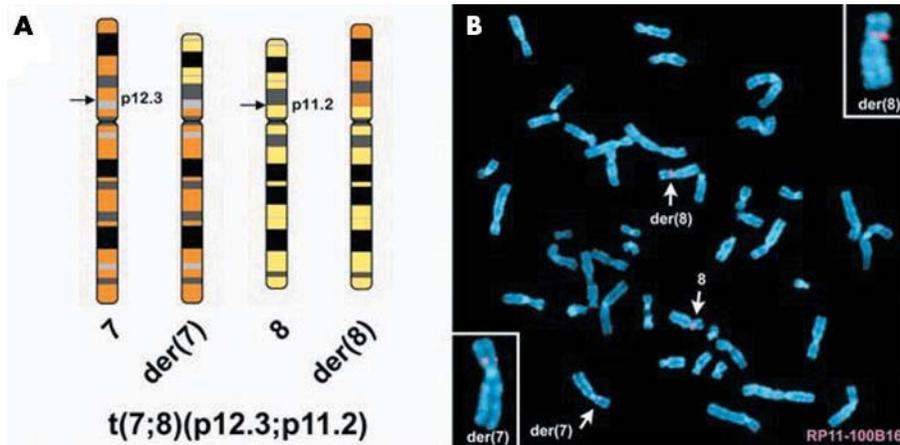
- Kallmann's syndrome (KS), characterised by hypogonadotropic hypogonadism and anosmia, can be caused by inactivating mutations of the X linked *KALI* gene, but these mutations account for less than 15% of KS patients. The remaining cases, as well as cases of hypogonadotropic hypogonadism without anosmia, are believed to be caused by mutations at two or more autosomal loci, including a segment of 8p heterozygous for a microdeletion in one KS patient.
- Recently, mutation in *FGFR1*, the 8p gene encoding fibroblast growth factor receptor 1, has been shown to cause autosomal dominant KS. We report positional cloning of the genomic breakpoints of the balanced reciprocal translocation t(7;8)(p12.3;p11.2) from a male patient with hypogonadotropic hypogonadism and cleft lip and palate. The translocation disrupts *FGFR1* (MIM 136350) between exons 2 and 3 and predicts a novel fusion gene product.
- Although various *FGFR1* translocations producing fusion proteins have been reported as causes of myeloproliferative disorders, this is the first case in which a constitutional *FGFR1* translocation is associated with a developmental disorder.

of anosmia. He was prescribed a regimen of testosterone injections, which successfully induced secondary sexual characteristics. At the age of 31, he was seen by a different physician for azoospermia and infertility, and cytogenetic analysis was ordered for the possibility of Klinefelter's syndrome. The analysis revealed an apparently balanced chromosomal translocation with the karyotype, 46,XY,t(7;8)(p12.3;p11.2). Informed consent for the generation of a lymphoblastoid cell line was obtained in accordance with institutional policies.<sup>9</sup>

## Fluorescent in situ hybridisation analysis

Breakpoint mapping on chromosome 8 was initiated using clones placed on the cytogenetic map by fluorescent in situ hybridisation (FISH) analysis and on the sequence map by sequence tagged sites.<sup>10</sup> Metaphase chromosomes from the patient cell line were prepared for analysis by GTG banding or FISH using standard protocols. Briefly, clones for FISH were

**Abbreviations:** FGF, fibroblast growth factor; FISH, fluorescent in situ hybridisation; KS, Kallmann's syndrome; SSCP, single strand conformation polymorphism; UCSC, University of California Santa Cruz



**Figure 1** Fluorescent in situ hybridisation (FISH) mapping of the chromosome 8 breakpoint. (A) Ideogram illustrating the balanced  $t(7;8)(p12.3;p11.2)$  in the patient. (B) FISH mapping with RP11-100B16, labelled with SpectrumOrange, resulted in hybridisation to the normal chromosome 8, and the der(8) and der(7) derivative chromosomes. The insets present the derivative chromosomes at higher magnification.

selected using genome maps provided by the National Center for Biotechnology Information and the University of California Santa Cruz (UCSC) Genomics Bioinformatics Group.<sup>10–11</sup> Bacterial artificial chromosome (BAC) clones were obtained from CITB-D and RP11 libraries (Invitrogen, San Diego, California, and the Children's Hospital of Oakland Research Institute) and directly labelled with SpectrumOrange or Green-dUTP (Vysis) by nick translation. Hybridisations were carried out according to manufacturers' protocols. Metaphase chromosomes were counterstained with 4,6-diamino-2-phenylindole-dihydrochloride (DAPI), and at least 10 metaphases were analysed using a Zeiss Axioskop microscope. Images were captured with the CytoVision system (Applied Imaging, San José, California, USA). The karyotype, 46,XY,t(7;8)(p12.3;p11.2), was reconfirmed by GTG banding before breakpoint mapping by FISH.

### Mapping and cloning of breakpoints

Southern blot analysis of patient lymphoblast genomic DNA with probes D011-A, D011-B, and D011-C to search for altered restriction fragments was carried out using standard protocols. For each lane, 10  $\mu$ g of genomic DNA from the patient and control were digested with an appropriate restriction enzyme. Fragments were separated on a 1.0% agarose gel and transferred to Hybond-N membrane (Amersham, Arlington Heights, Illinois, USA). Filters were ultraviolet cross linked, baked at 80°C, and hybridised with probes labelled with <sup>32</sup>P-dCTP by random priming. Hybridisation of labelled fragments was done in the presence of excess herring sperm competitor DNA, and hybridised membranes were washed at 60°C with 0.15 M NaCl/0.015 M sodium citrate/0.1% sodium dodecyl sulphate (SDS) for 30 minutes. Autoradiography took place for 16 hours at –70°C using two intensifying screens. Three hybridisation probes were amplified by the following primer sets:

D011-A: 5'CTGTCAGGGTTCCATCACC3'+5'CCTAGAAACC TCCGTGTGC3'; D011-B: 5'GTGGCTCTGTTCTATCCCTC3'+5'CAACAGTCATGGGAACCATC3'; D011-C: 5'GCACCTAGAG CCTGTAATAG3'+5'TGTCCAAGTCTCTCCTCGGA3'.

A 1.6 kb *Bam*HI junction fragment from der(8) was amplified by suppression polymerase chain reaction (PCR) using the following primer sets: 5'CCTAATACGACTCAC TATAGG3'+5'GCAATGCACTGTAAACACATG3'; 5'CTATAGG GCTCGAGCGGC3'+5'CCTAGAGCCTGTAATAGTAA3'.<sup>12</sup> Then, the der(7) junction fragment was amplified by nested PCR using the primer sets: 5'GGATCATTAGAGGGATTTCGAA3'+

5'GCAAGCTGTGCTGGAAGCA3'; 5'CCAGCTTCACAGGTG TTTC3'+5'CCAGCATTGAAGAGGGAGT3'.

### Fusion transcript amplification

Total RNA was isolated from patient and control lymphoblastoid cell lines with the RNeasy Mini Kit (Qiagen, Valencia, California, USA). Reverse transcription of total RNA (1  $\mu$ g) was undertaken by using either random hexanucleotide priming and Superscript II (Gibco BRL, Gaithersburg, Maryland, USA) or the SMART-PCR cDNA synthesis kit (Clontech, Palo Alto, California, USA) according to the protocols provided. In each experiment, DNA contamination was excluded by the absence of a PCR product in the sample without reverse transcriptase, amplified under the same conditions as the reverse transcribed RNA sample. Nested PCR was carried out using *Pfu* polymerase (Gibco BRL) with the following primer sets, annealing at 56°C for 30 seconds with an extension for one minute 40 seconds: *TENS1-FGFR1*: 5'CTGAGAAAGCCCTCAGTGTC3'+5'CAAG ATCTGGACATAAGGCAGG3', 5'GGCAGAGCAGCTACTCC ACA3'+5'GTCACCTGTACACCTTACACATGAACTC3'; *FGFR1-TENS1*: 5'CCTCTTGCGGCCACAGGC3'+5'CCTTCAACATGGC GATGG3', 5'GCAGCGCGCGGAG3'+5'CCTTGTACCAGAAGT GGAAGTG3'.

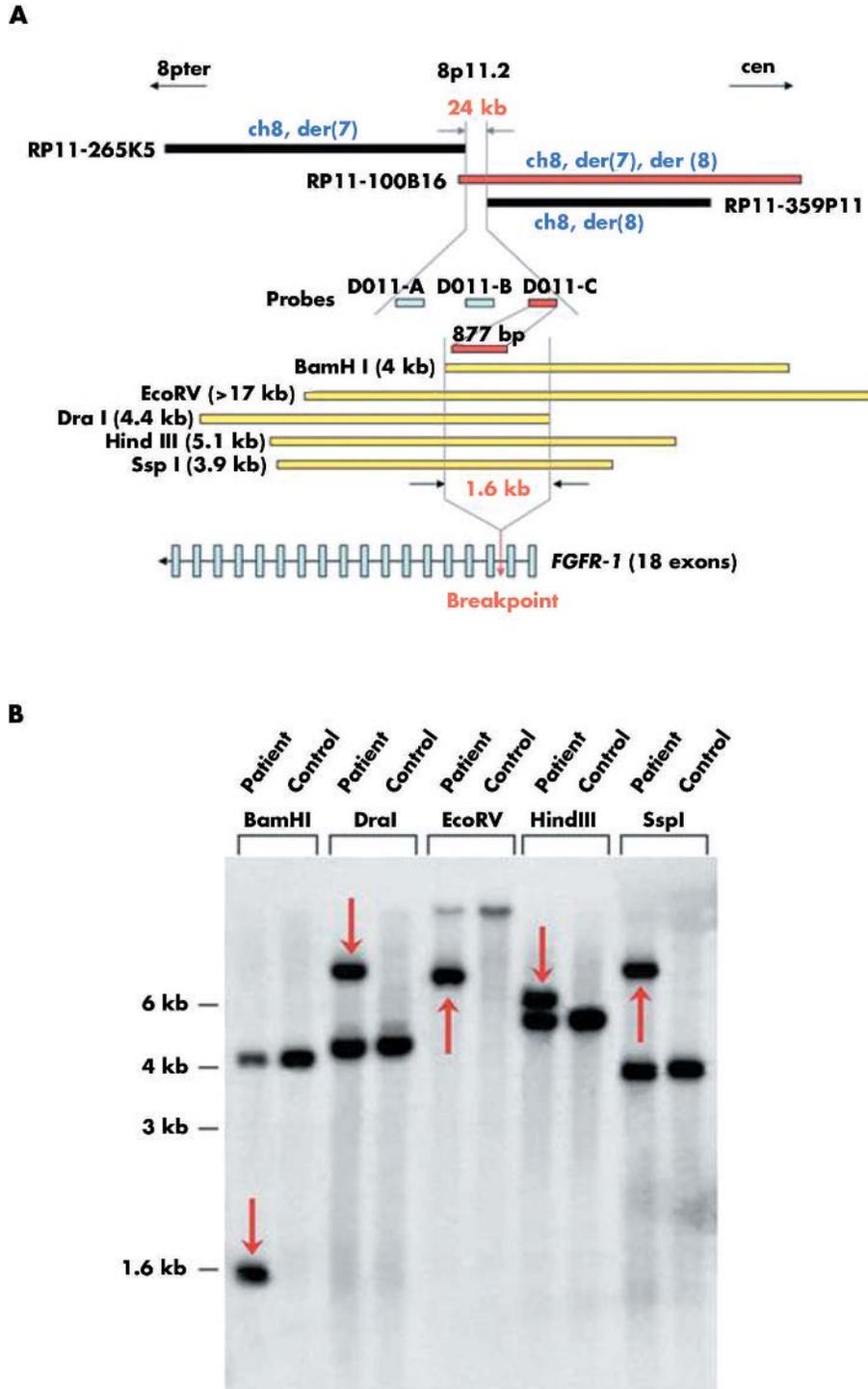
### Mutation analysis

Mutation analysis of the second allele of *FGFR1* was done by single strand conformation polymorphism (SSCP). In all, 24 genomic fragments including the entire coding region, UTR, and intron–exon boundaries were amplified from 18 exons of *FGFR1* by PCR with [<sup>32</sup>P]-dCTP. Primers were designed to amplify genomic fragments with the size of 200 to 300 base pairs (bp) (primer sequences and amplification conditions are available on request). PCR products were applied on non-denaturing 8% glycerol gels with electrophoresis overnight at room temperature and 8W constant power. PCR products that displayed a banding pattern different from control samples were sequenced by an ABI Prism 377/XL DNA sequencer (Applied Biosystems, Foster City, California, USA).

## RESULTS

### Delineation of the breakpoint region on 8p11.2

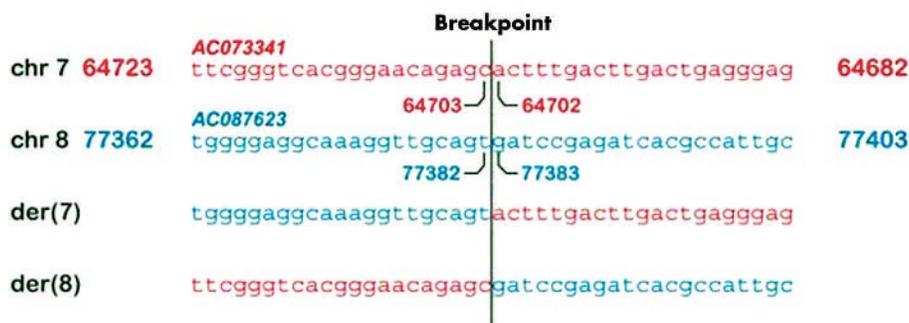
To identify the genes potentially disrupted in the patient, we first mapped the translocation breakpoints using FISH. Two BAC clones selected from the UCSC map as starting clones for FISH—GS1-211B7 and GS1-165D4 from 8p12 and 8p11.2—mapped telomeric and centromeric to the breakpoint,



**Figure 2** (A) Schematic presentation of positional cloning on chromosome 8. The 24 kb breakpoint region determined by fluorescent in situ hybridisation (FISH) was further narrowed to 1.6 kb by Southern blot hybridisation, which detected junction fragments with five different restriction enzymes. The breakpoint is located between exons 2 and 3 of *FGFR1*. The *FGFR1* gene is not to scale. (B) Southern blot hybridisation of genomic DNA from the translocation patient with 877 bp probe D011-C from intron 2 of *FGFR1*. Note the detection of five altered fragments caused by the translocation junction, generated by enzymes BamHI, DraI, EcoRV, HindIII, and SspI. C, genomic DNA from karyotypically normal control; P, patient genomic DNA.

respectively, showing that the chromosome 8 breakpoint was contained in a 12 Mb region. Subsequent experiments were carried out using BACs chosen from within this region to narrow the candidate region until a breakpoint crossing BAC clone was identified. Seventeen BACs were examined, of

which eight were proximal to the breakpoint and nine distal. The final BAC, RP11-100B16, hybridised to the normal chromosome 8, and both der(7) and der(8) chromosomes, indicating that it spans the translocation breakpoint (fig 1B). Additionally two BAC clones that partially



**Figure 3** Genomic DNA sequence at two breakpoints from two junction fragments. The breakpoint on chromosome 7 is located between nucleotides 64702 and 64703 of AC073341, while on chromosome 8 it occurs between nucleotides 77382 and 77383 of AC087623. Note no gain or loss of nucleotides at the two breakpoints of the derivative chromosomes, showing a perfectly balanced reciprocal translocation.

overlap with RP11-100B16—RP11-265K5 and RP11-359P11—mapped distal and proximal to the breakpoint, respectively (fig 2A). Based on the sequence of these BACs, the location of the breakpoint region was confined to ~24 kb of DNA.

### Southern blot hybridisation and cloning of the breakpoints on 8p11.2

To localise the breakpoint region in 8p11.2 further, three DNA fragments—D011-A, D011-B, and D011-C—were amplified by PCR from the narrowed 24 kb region (fig 2A) and used to probe patient DNA on genomic blots. D011-B and D011-C both detected altered restriction fragments due to the translocation (fig 2B), narrowing the breakpoint to 1557 bp on the BAC restriction map and suggesting that the breakpoint is between exons 2 and 3 of *FGFR1* isoform 1 (fig 2A). We cloned and sequenced junction fragments spanning the breakpoints from both derivative chromosomes, which revealed that the translocation is perfectly balanced, without the gain or loss of any sequence. The sequences of the breakpoint regions for the der(7) and der(8) chromosomes are given in fig 3.

### Delineation of the breakpoint region on 7p12.3

BAC clones RP11-183O1 from 7p22.1 and RP11-34J24 from 7p11.2 were used as starting clones for FISH, and mapped distal and proximal to the breakpoint, respectively, indicating that the chromosome 7 breakpoint was contained in a 49 Mb region. Using randomly selected BACs, the breakpoint region was narrowed to ~1.3 Mb, flanked by RP11-126K7 and RP11-271O10, which map telomeric and centromeric to the breakpoint, respectively. After the breakpoint was cloned and sequenced, on the basis of the chromosome 8 findings (see below) the junction sequence was found to be located in RP11-549I23. This was confirmed by FISH, showing three signals, one each on chromosome 7 and both derivative chromosomes (data not shown).

### *TENS1* in 7p12 and *FGFR1* in 8p11 are disrupted

The chromosome 7 breakpoint lies in intron 15 of *TENS1* (AF417489, 1445 amino acids, between nucleotides 64702 and 64703 of GenBank entry AC073341, 15665 bp downstream of exon 15), while the breakpoint on chromosome 8 is in intron 2 of *FGFR1* (NM\_000604, 822 amino acids, between nucleotides 77382 and 77383 of GenBank entry AC087623, 22414 bp downstream of exon 2). The chromosome 8 breakpoint maps within a SINE/Alu repetitive sequence while the breakpoint in *TENS1* occurs in unique intronic sequence with no apparent homology to the chromosome 8 breakpoint region.

### Fusion transcript amplification

The location of the translocation breakpoint predicts two putative reciprocal in-frame fusion transcripts *TENS1/FGFR1* and *FGFR1/TENS1* (fig 4A). On the derivative chromosome 7, exons 1–15 of *TENS1* are predicted to join with exons 3–18 of *FGFR1* isoform 1 and to form a 5891 bp *TENS1Δex16-26/FGFR1Δex1-2* transcript that spans 31 exons and encodes 1675 amino acids without frameshift, from the normal *TENS1* initiation codon (ATG) in exon 1 to the termination codon (TGA) in exon 31 (corresponding to the normal stop codon in *FGFR1* exon 18). The putative fusion protein would consist of the first 883 amino acids of *TENS1* joined to the final 792 amino acids of *FGFR1* (fig 4, panels A and B).

On the derivative chromosome 8, exons 1–2 of *FGFR1* are joined to exons 16–26 of *TENS1*, predicting a 2546 bp *FGFR1Δex3-18/TENS1Δex1-15* transcript that comprises 13 exons and encodes 592 amino acids. Again there is no frameshift, as the start codon and stop codon occur at exon 2 and exon 13, respectively, at the same positions as the corresponding exon 2 of *FGFR1* and exon 26 of *TENS1*. The predicted *FGFR1-TENS1* fusion protein would contain the first 30 amino acids of *FGFR1* followed by the final 562 amino acids of *TENS1* (fig 4, panels A and B).

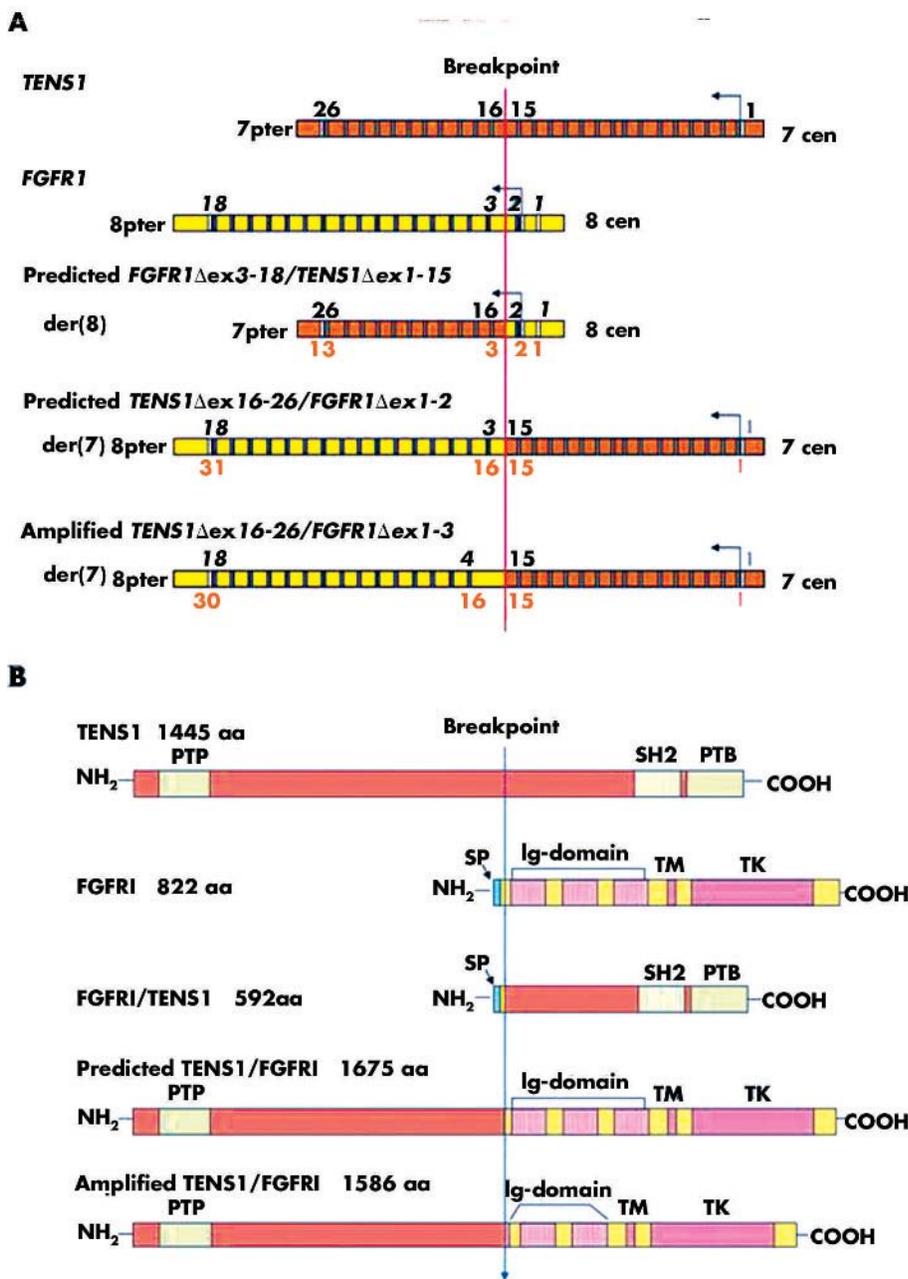
To establish whether either fusion transcript is expressed in a lymphoblastoid cell line from the patient, we carried out reverse transcriptase PCR (RT-PCR). Only the *TENS1-FGFR1* fusion transcript was detected (fig 5), but sequencing revealed the skipping of *FGFR1* exon 3, an alternative splicing pattern also seen in several native *FGFR1* encoded isoforms. The 5624 bp *TENS1Δex16-26/FGFR1Δex1-3* transcript encodes a fusion protein of 1586 amino acids comprising the first 883 amino acids of *TENS1* joined to the final 703 amino acids of *FGFR1* (fig 4, panels A and B).

### Mutation analysis of the non-translocated *FGFR1* allele

Mutation analysis of the second non-translocated *FGFR1* allele from the patient, done by SSCP and direct sequencing, identified only a heterozygous nucleotide difference, 345 C→T in exon 3, a known SNP (NCBI reference SNP ID: rs2915665) which does not alter the Ser amino acid encoded at this site. Thus the presence of the translocated allele of the *FGFR1* results in a disease phenotype without a corresponding coding sequence mutation in the alternate *FGFR1* allele.

### DISCUSSION

The *TENS1* locus encodes tensin-like SH2 domain-containing protein 1 (also known as tumour endothelial marker 6, tensin 3), a 1445 amino acid protein named for its similarity with tensin, an actin filament crosslinking protein found in focal adhesions.<sup>13 14</sup> *TENS1* protein contains a protein tyrosine

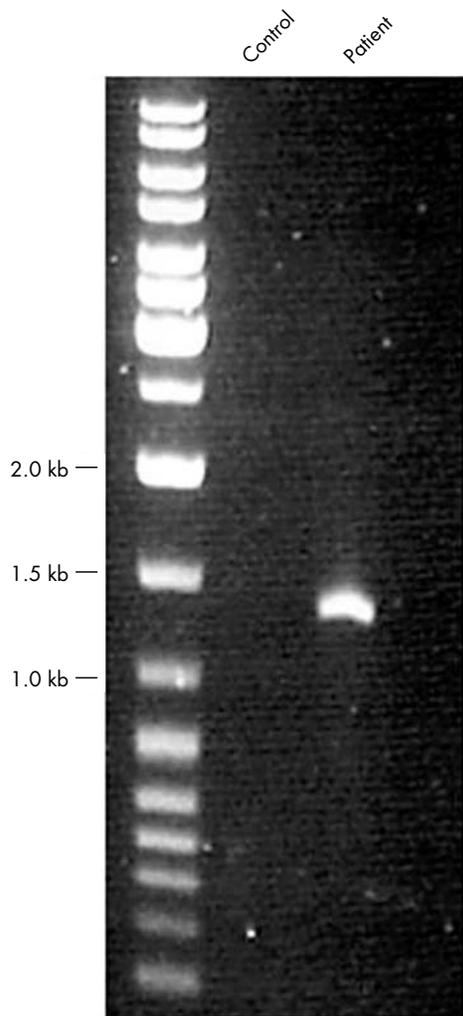


**Figure 4** (A) Disruption of genes *FGFR1* and *TENS1* by t(7;8), resulting in two in-frame fusion genes. White boxes indicate 5' and 3' untranslated regions of exons. *TENS1* and *FGFR1* coding exons are shown as grey boxes and blue boxes, respectively, numbered above the gene in the same colour, whereas exons of fusion genes are numbered in orange below the gene. Note that the stop codons of the two fusion genes are at the same exon locations as in the corresponding wild-type genes, as there is no frameshift. In the amplified *TENS1*-*FGFR1* fusion transcript, exon 3 of *FGFR1* is skipped. Notable exons are numbered and the direction of gene transcription is indicated by arrows. The sizes of exons and introns is not to scale. (B) *TENS1*, *FGFR1*, *FGFR1*-*TENS1*, and *TENS1*-*FGFR1* protein domains. In the amplified fusion protein *TENS1*-*FGFR1*, the transmembrane domain (TM), and tyrosine kinase domain (TK) of *FGFR1* are not affected by the translocation, but the signal peptide (SP) of *FGFR1* and the first immunoglobulin (Ig) like domain are absent. The translocation does not directly disrupt the amino terminal protein tyrosine phosphatase domain (PTP) or the carboxyl terminal Src homology 2 (SH2) and phosphotyrosine binding (PTB) domains of *TENS1*, but does segregate the coding sequences for these domains to different fusion transcripts.

phosphatase domain in the amino-terminal region and Src homology 2 (SH2) and phosphotyrosine binding domains near its carboxyl-terminus. *FGFR1* encodes several different isoforms of a transmembrane protein, the extracellular moiety of which interacts with fibroblast growth factors (FGFs), setting in motion a cascade of downstream signals, ultimately influencing mitogenesis and differentiation; it is characterised by two or three extracellular immunoglobulin-like loops (depending on inclusion of exon 3), a transmembrane domain, and an intracellular tyrosine kinase domain.

The predicted *TENS1*/*FGFR1* fusion protein lacks the *FGFR1* signal peptide and the first Ig-like domain, but contains Ig-like domains 2 and 3, which are sufficient for specific FGF binding, and an intact tyrosine kinase domain region, suggesting the potential for functionality (fig 3B).<sup>15</sup>

However, the related KS phenotype in a patient hemizygous for 8p due to the deletion region suggests that the translocation produces hypogonadotropic hypogonadism as a result of haploinsufficiency for *FGFR1*.<sup>7</sup> Consistent with this view, there was no evidence for an *FGFR1* mutation on the second allele in SSCP/sequence analysis in any of 18 exons and splice junctions. The apparent functional hemizygosity for *FGFR1* in the translocation patient probably reflects a failure to direct the *FGFR1* functional domains to the proper location in the plasma membrane owing to the absence of the appropriate signal peptide and the presence of the large *TENS1* moiety. While this work was being completed, Dodé *et al*, and subsequently Sato *et al*, reported several truncating and missense *FGFR1* mutations in KS patients, some with cleft lip and palate, consistent with the haploinsufficiency in



**Figure 5** Expression of the fusion gene *TENS1*<sup>ex16-26</sup>/*FGFR1*<sup>ex1-3</sup>. Nested RT-PCR was performed by using forward primers in exon 15 of *TENS1* and reverse primers in exon 7 of *FGFR1*. Reverse transcribed t(7;8) patient RNA, but not RNA from a normal individual (control), resulted in amplification of a fusion transcript of 1.4 kb smaller than the expected size of 1.7 kb. The sequence analysis of this amplified fusion gene from the patient confirmed the skipping of exon 3 of *FGFR1*.

the patient reported here.<sup>8 16</sup> However, the absence of frank anosmia in the current patient indicates that sufficient *FGFR1* function may have been maintained to prevent the degree of agenesis of the olfactory lobes typical in KS.

As the translocation patient displays no obvious phenotypes distinct from those seen in patients with KS associated *FGFR1* point mutations, the disruption of *TENS1* does not seem to contribute to the patient's abnormalities. This suggests that heterozygous inactivation of *TENS1* is without dramatic consequence, but the possibility that the predicted fusion proteins effectively provide normal *TENS1* function cannot be excluded.

The X linked form of KS is associated with inactivating mutation of the *KALI* gene, encoding anosmin 1, a secreted proteoglycan binding protein with similarities to neuronal cell adhesion molecules.<sup>3 17</sup> Anosmin 1 interacts with heparan and chondroitin sulphates to promote cell adhesion and neuronal outgrowth, and has been implicated in the migration of gonadotropin releasing hormone (GnRH) producing neurones and olfactory axonal fibres, though the receptor system through which it acts remains uncertain.<sup>18</sup> Notably, *FGFR1* activation by binding to FGF ligands involves

receptor dimerisation that also requires heparan sulphate proteoglycan binding.<sup>19</sup> Indeed, FGF2 ligand and *FGFR1* have been co-crystallised with heparin, and the structure of the complex defined.<sup>20</sup> The common association with heparan sulphates and the similar effects of *KALI* and *FGFR1* inactivating mutations support the suggestion that the *FGFR1* signalling pathway participates directly in mediating anosmin 1 function.<sup>8 21</sup>

The translocation patient reported here and the KS patients reported by others also support the view that haploinsufficiency for *FGFR1* is a cause of cleft lip and palate.<sup>7 8 16</sup> Interestingly, *FGFR1* gain of function mutations have previously been associated with the craniosynostosis of Pfeiffer's syndrome and in the Jackson-Weiss syndrome.<sup>22 23</sup> These syndromes can also be caused by mutations in *FGFR2*, which has also been associated with cleft palate in Apert's syndrome, indicating that the two receptors may function in the same signalling pathway.<sup>24</sup> This suggests that *FGFR2*, located at 10q26, may be an excellent candidate for an additional KS or idiopathic hypogonadotropic hypogonadism locus. Hence it would be of interest to determine whether *FGFR2* is disrupted by translocation in a KS patient with a de novo unbalanced der(1)t(1;10)(q44;q26).<sup>4</sup>

It is noteworthy that a variety of fusion proteins involving *FGFR1* underlie the 8p11 myeloproliferative syndrome (EMS)/stem cell leukaemia-lymphoma syndrome, presumably because of constitutive activity of the tyrosine kinase domain.<sup>25-29</sup> However, neither Pfeiffer's syndrome nor the Jackson-Weiss syndrome shows a myeloproliferation defect, suggesting that the gain of function in these cases is insufficient or inappropriate to transform target cells. The constitutional translocation reported here creates a predicted fusion protein that has not produced a myeloid disorder, despite being likely to mislocalise a portion of *FGFR1* containing the tyrosine kinase domain. This reinforces the view that both the fusion partner and the site of the breakpoint are likely to be critical in producing constitutive tyrosine kinase activity in a manner that leads to malignancy. This is the first demonstration that constitutional translocation of *FGFR1* can lead to abnormal development rather than to myeloid disorder, and provides a basis for more detailed structure-function comparison of the respective fusion proteins.

#### ELECTRONIC DATABASE INFORMATION

GenBank accession numbers: *FGFR1*, AC087623, NM\_000604; *TENS1*, AC073341, AF417489.  
dbSNP information: rs2915665.

#### ACKNOWLEDGEMENTS

We are indebted to Carolyne Rooryck and Robert E Eisenman for technical assistance, Amy Bosco and Heather L Ferguson for obtaining informed consent and clinical information, Wenqi Zeng and Jo-Chen Chou for technical advice, Tammy Gillis, Michelle Flores, and the MGH Genomics Core Facility for DNA sequence analysis and Francesca Puglisi for cell culture and genomic DNA extraction. This work was supported by USPHS grants GM061354 (Developmental Genome Anatomy Project) and HD28138.

#### Authors' affiliations

**HG Kim, S Kishikawa, M E MacDonald, J F Gusella**, Molecular Neurogenetics Unit, Center for Human Genetic Research, Massachusetts General Hospital/Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA

**S R Herrick, B J Quade**, Department of Pathology, Brigham and Women's Hospital/Harvard Medical School

**E Lemyre**, Medical Genetics Service, Hôpital Ste Justine, University of Montreal, Montreal, Canada

**J A Salisz**, West Shore Urology, Mercy Drive, Muskegon, Michigan, USA

**S Seminara**, Reproductive Endocrine Unit, Massachusetts General Hospital

**G A P Bruns**, Genetics Division, Children's Hospital/Department of Pediatrics, Harvard Medical School

**C C Morton**, Departments of Obstetrics, Gynecology and Reproductive Biology and Pathology, Brigham and Women's Hospital/Harvard Medical School

Conflicts of interest: none declared.

Correspondence to: Dr James F Gusella, Molecular Neurogenetics Unit, Center for Human Genetic Research, Massachusetts General Hospital/Department of Genetics, Harvard Medical School, CNY149-6214, 13th Street, Boston, Massachusetts 02129, USA; [gusella@helix.mgh.harvard.edu](mailto:gusella@helix.mgh.harvard.edu)

## REFERENCES

- Collins FS. Positional cloning moves from perditional to traditional. *Nat Genet* 1995;**9**:347-50.
- Oliveira LM, Seminara SB, Beranova M, Hayes FJ, Valkenburgh SB, Schipani E, Costa EM, Latronico AC, Crowley WF, Vallejo M. The importance of autosomal genes in Kallmann syndrome: genotype-phenotype correlations and neuroendocrine characteristics. *J Clin Endocrinol Metab* 2001;**86**:1532-8.
- Franco B, Guioli S, Pragliola A, Incerti B, Bardoni B, Tonlorenzi R, Carozzo R, Maestrini E, Pieretti M, Taillon-Miller P, Brown C, Willard H, Lawrence C, Persico M, Camerino G, Ballabio A. A gene deleted in Kallmann's syndrome shares homology with neural cell adhesion and axonal path-finding molecules. *Nature* 1991;**353**:529-36.
- Schinzl A, Lorda-Sanchez I, Binkert F, Carter NP, Bebb CE, Ferguson-Smith MA, Eiholzer U, Zachmann M, Robinson WP. Kallmann syndrome in a boy with a t(1;10) translocation detected by reverse chromosome painting. *J Med Genet* 1995;**32**:957-61.
- Kikuchi I, Nagamine M, Ueda A, Mihara K, Seitama M, Minoda M. Chromosomal translocation t(13;16) in a patient with idiopathic hypogonadotropic hypogonadism. *Intern Med* 1993;**32**:465-7.
- Best LG, Wasdahl WA, Larson LM, Sturlaugson J. Chromosome abnormality in Kallmann syndrome. *Am J Med Genet* 1990;**35**:306-9.
- Vermeulen S, Messiaen L, Scheir P, De Bie S, Speleman F, De Paepe A. Kallmann syndrome in a patient with congenital spherocytosis and an interstitial 8p11.2 deletion. *Am J Med Genet* 2002;**108**:315-18.
- Dode C, Levilliers J, Dupont JM, De Paepe A, Le Du N, Soussi-Yanicostas N, Coimbra RS, Delmagnani S, Compain-Nouaille S, Baverel F, Pecheux C, Le Tessier D, Cruaud C, Delpuch M, Speleman F, Vermeulen S, Amalfitano A, Bachelot Y, Bouchard P, Cabrol S, Carel JC, Delemarre-van de Waal H, Goulet-Salmon B, Kottler ML, Richard O, Sanchez-Franco F, Saura R, Young J, Petit C, Hardelin JP. Loss-of-function mutations in FGFR1 cause autosomal dominant Kallmann syndrome. *Nat Genet* 2003;**33**:463-5.
- Anderson MA, Gusella JF. Use of cyclosporin A in establishing Epstein-Barr virus-transformed human lymphoblastoid cell lines. *In Vitro* 1984;**20**:856-8.
- Cheung VG, Nowak N, Jang W, Kirsch IR, Zhao S, Chen XN, Furey TS, Kim UJ, Kuo WL, Olivier M, Conroy J, Kasprzyk A, Massa H, Yonescu R, Sait S, Thoren C, Snijders A, Lemyre E, Bailey JA, Bruzel A, Burrill WD, Clegg SM, Collins S, Dhani P, Friedman C, Han CS, Herrick S, Lee J, Ligon AH, Lowry S, Morley M, Narasimhan S, Osoegawa K, Peng Z, Plajzer-Frick I, Quade BJ, Scott D, Sirotkin K, Thorpe AA, Gray JW, Hudson J, Pinkel D, Ried T, Rowen L, Shen-Ong GL, Strausberg RL, Birney E, Callen DF, Cheng JF, Cox DR, Doggett NA, Carter NP, Eichler EE, Haussler D, Korenberg JR, Morton CC, Albertson D, Schuler G, de Jong PJ, Trask BJ. Integration of cytogenetic landmarks into the draft sequence of the human genome. *Nature* 2001;**409**:953-8.
- Karolchik D, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, Roskin KM, Schwartz M, Sugnet CW, Thomas DJ, Weber RJ, Haussler D, Kent WJ. The UCSC Genome Browser Database. *Nucleic Acids Res* 2003;**31**:51-4.
- Siebert PD, Chenchik A, Kellogg DE, Lukyanov KA, Lukyanov SA. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res* 1995;**23**:1087-8.
- Carson-Walter EB, Watkins DN, Nanda A, Vogelstein B, Kinzler KW, St Croix B. Cell surface tumor endothelial markers are conserved in mice and humans. *Cancer Res* 2001;**61**:6649-55.
- Chen H, Ishii A, Wong WK, Chen LB, Lo SH. Molecular characterization of human tensin. *Biochem J* 2000;**351**:403-11.
- Plotnikov AN, Schlessinger J, Hubbard SR, Mohammadi M. Structural basis for FGF receptor dimerization and activation. *Cell* 1999;**98**:641-50.
- Sato N, Katsumata N, Kagami M, Hasegawa T, Hori N, Kawakita S, Minowada S, Shimotsuka A, Shishiba Y, Yokozawa M, Yasuda T, Nagasaki K, Hasegawa D, Hasegawa Y, Tachibana K, Naiki Y, Horikawa R, Tanaka T, Ogata T. Clinical assessment and mutation analysis of Kallmann syndrome 1 (KAL1) and fibroblast growth factor receptor 1 (FGFR1, or KAL2) in five families and 18 sporadic patients. *J Clin Endocrinol Metab* 2004;**89**:1079-88.
- Hardelin JP. Kallmann syndrome: towards molecular pathogenesis. *Mol Cell Endocrinol* 2001;**179**:75-81.
- Cariboni A, Pimpinelli F, Colamarino S, Zaninetti R, Piccollella M, Rumio C, Piva F, Rugarli EI, Maggi R. The product of X-linked Kallmann's syndrome gene (KAL1) affects the migratory activity of gonadotropin-releasing hormone (GnRH)-producing neurons. *Hum Mol Genet* 2004;**13**:2781-91.
- Ornitz DM. FGFs, heparan sulfate and FGFRs: complex interactions essential for development. *Bioessays* 2000;**22**:108-12.
- Schlessinger J, Plotnikov AN, Ibrahim OA, Eliseenkova AV, Yeh BK, Yayon A, Linhardt RJ, Mohammadi M. Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol Cell* 2000;**6**:743-50.
- Dode C, Hardelin JP. Kallmann syndrome: fibroblast growth factor signaling inefficiency? *J Mol Med* 2004;**82**:725-34.
- Muenke M, Schell U, Hehr A, Robin NH, Losken HW, Schinzl A, Pulley LJ, Rutland P, Reardon W, Malcolm S, et al. A common mutation in the fibroblast growth factor receptor 1 gene in Pfeiffer syndrome. *Nat Genet* 1994;**8**:269-74.
- Roscioli T, Flanagan S, Kumar P, Masel J, Gattas M, Hyland VJ, Glass IA. Clinical findings in a patient with FGFR1 P252R mutation and comparison with the literature. *Am J Med Genet* 2000;**93**:22-8.
- Wilkie AO, Morriss-Kay GM. Genetics of craniofacial development and malformation. *Nat Rev Genet* 2001;**2**:458-68.
- Xiao S, Nalabolu SR, Aster JC, Ma J, Abruzzo L, Jaffe ES, Stone R, Weissman SM, Hudson TJ, Fletcher JA. FGFR1 is fused with a novel zinc-finger gene, ZNF198, in the t(8;13) leukaemia/lymphoma syndrome. *Nat Genet* 1998;**18**:84-7.
- Popovici C, Adelaide J, Ollendorff V, Chaffanet M, Guasch G, Jacrot M, Leroux D, Birnbaum D, Pebusque MJ. Fibroblast growth factor receptor 1 is fused to FIM in stem-cell myeloproliferative disorder with t(8;13). *Proc Natl Acad Sci USA* 1998;**95**:5712-17.
- Popovici C, Zhang B, Gregoire MJ, Jonveaux P, Lafage-Pochitaloff M, Birnbaum D, Pebusque MJ. The t(6;8)(q27;p11) translocation in a stem cell myeloproliferative disorder fuses a novel gene, FOP, to fibroblast growth factor receptor 1. *Blood* 1999;**93**:1381-9.
- Guasch G, Mack GJ, Popovici C, Dastugue N, Birnbaum D, Rattner JB, Pebusque MJ. FGFR1 is fused to the centrosome-associated protein CEP110 in the 8p12 stem cell myeloproliferative disorder with t(8;9)(p12;q33). *Blood* 2000;**95**:1788-96.
- Demiroglu A, Steer EJ, Heath C, Taylor K, Bentley M, Allen SL, Koduru P, Brody JP, Hawson G, Rodwell R, Doody ML, Carnicero F, Reiter A, Goldman JM, Melo JV, Cross NC. The t(8;22) in chronic myeloid leukemia fuses BCR to FGFR1: transforming activity and specific inhibition of FGFR1 fusion proteins. *Blood* 2001;**98**:3778-83.

## 12.2 Paper II

Hyung-Goo Kim, Ingo Kurth, Fei Lan, Irene Meliciani, Wolfgang Wenzel, Soo Hyun Eom, Gil Bu Kang, Georg Rosenberger, Mustafa Tekin, Metin Ozata, David P. Bick, Richard J. Sherins, Steven L. Walker, Yang Shi, James F. Gusella, and Lawrence C. Layman\*

\* Corresponding author

***Mutations in CHD7, Encoding a Chromatin-Remodeling Protein, Cause Idiopathic Hypogonadotropic Hypogonadism and Kallmann Syndrome***

American Journal of Human Genetics, 2008 Oct;83(4):511-9.

# Mutations in *CHD7*, Encoding a Chromatin-Remodeling Protein, Cause Idiopathic Hypogonadotropic Hypogonadism and Kallmann Syndrome

Hyung-Goo Kim,<sup>1,13</sup> Ingo Kurth,<sup>2</sup> Fei Lan,<sup>3,14</sup> Irene Melicani,<sup>4,15</sup> Wolfgang Wenzel,<sup>5</sup> Soo Hyun Eom,<sup>6</sup> Gil Bu Kang,<sup>6</sup> Georg Rosenberger,<sup>2</sup> Mustafa Tekin,<sup>7</sup> Metin Ozata,<sup>8</sup> David P. Bick,<sup>9</sup> Richard J. Sherins,<sup>10</sup> Steven L. Walker,<sup>11,12</sup> Yang Shi,<sup>3</sup> James F. Gusella,<sup>1</sup> and Lawrence C. Layman<sup>11,12,\*</sup>

CHARGE syndrome and Kallmann syndrome (KS) are two distinct developmental disorders sharing overlapping features of impaired olfaction and hypogonadism. KS is a genetically heterogeneous disorder consisting of idiopathic hypogonadotropic hypogonadism (IHH) and anosmia, and is most commonly due to *KAL1* or *FGFR1* mutations. CHARGE syndrome, a multisystem autosomal-dominant disorder, is caused by *CHD7* mutations. We hypothesized that *CHD7* would be involved in the pathogenesis of IHH and KS (IHH/KS) without the CHARGE phenotype and that IHH/KS represents a milder allelic variant of CHARGE syndrome. Mutation screening of the 37 protein-coding exons of *CHD7* was performed in 101 IHH/KS patients without a CHARGE phenotype. In an additional 96 IHH/KS patients, exons 6–10, encoding the conserved chromodomains, were sequenced. RT-PCR, SIFT, protein-structure analysis, and in situ hybridization were performed for additional supportive evidence. Seven heterozygous mutations, two splice and five missense, which were absent in  $\geq 180$  controls, were identified in three sporadic KS and four sporadic normosmic IHH patients. Three mutations affect chromodomains critical for proper *CHD7* function in chromatin remodeling and transcriptional regulation, whereas the other four affect conserved residues, suggesting that they are deleterious. *CHD7*'s role is further corroborated by specific expression in IHH/KS-relevant tissues and appropriate developmental expression. Sporadic *CHD7* mutations occur in 6% of IHH/KS patients. *CHD7* represents the first identified chromatin-remodeling protein with a role in human puberty and the second gene to cause both normosmic IHH and KS in humans. Our findings indicate that both normosmic IHH and KS are mild allelic variants of CHARGE syndrome and are caused by *CHD7* mutations.

Idiopathic hypogonadotropic hypogonadism (IHH, MIM 146110), one of the most commonly inherited forms of hypogonadism, results from deficient hypothalamic GnRH release or action.<sup>1</sup> IHH patients present with absent or impaired sexual development due to sex-steroid-hormone deficiency, low serum levels of the pituitary gonadotropins follicle-stimulating hormone (FSH) and luteinizing hormone (LH), and infertility.<sup>1</sup> Kallmann syndrome (KS, MIM 308700, 147950, 244200, 610628), which couples IHH with the inability to smell (anosmia), is due to impairment of the normal embryologic migration of GnRH and olfactory neurons from the olfactory placode region into the hypothalamus.<sup>1,2</sup> KS patients may have additional phenotypic abnormalities including cleft lip and palate, unilateral renal agenesis, dental agenesis, and neurologic abnormalities such as synkinesia and cerebellar dysfunc-

tion.<sup>2</sup> The molecular basis of IHH and KS (IHH/KS) has been identified for approximately 25%–30% of patients, with mutations in the *KAL1* (MIM 308700),<sup>3,4</sup> *FGFR1* (MIM 136350),<sup>5–7</sup> and *GNRHR* (MIM 138850)<sup>8,9</sup> genes being most common. A variety of other genes also cause IHH/KS in some patients, including *GPR54* (MIM 604161),<sup>10</sup> *NROB1* (MIM 300473),<sup>11</sup> *PROKR2* (MIM 607123),<sup>12</sup> *PROK2* (MIM 607002),<sup>12</sup> *LEP* (MIM 164160),<sup>2</sup> and *LEPR* (MIM 601007).<sup>2</sup> A digenic inheritance pattern has been reported in two cases.<sup>13</sup> To date, *KAL1* mutations cause only KS,<sup>2</sup> whereas *GNRHR* mutations are restricted to normosmic IHH.<sup>14</sup> Only mutations in *FGFR1* are known to cause both normosmic IHH and KS.

Heterozygous *CHD7* (chromodomain helicase DNA-binding protein 7, MIM 608892) mutations have been identified in ~60%–70% of patients with CHARGE syndrome

<sup>1</sup>Molecular Neurogenetics Unit, Center for Human Genetic Research, Massachusetts General Hospital, and Department of Genetics, Harvard Medical School, Boston, MA 02114, USA; <sup>2</sup>Institut für Humangenetik, Universitätsklinikum Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany; <sup>3</sup>Department of Pathology, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA; <sup>4</sup>Faculty of Science, University "La Sapienza," Piazzale Aldo Moro 5, 00185 Roma, Italy; <sup>5</sup>Forschungszentrum Karlsruhe, Institute for Nanotechnology, PO Box 3640, 76021 Karlsruhe, Germany; <sup>6</sup>Department of Life Science, Cell Dynamics Research Center, Gwangju Institute of Science and Technology, Gwangju 500-712, South Korea; <sup>7</sup>Division of Clinical Molecular Pathology and Genetics, Department of Pediatrics, Ankara University School of Medicine, Ankara, 06100 Turkey; <sup>8</sup>GATA Haydarpaşa Training Hospital, Department of Endocrinology, Istanbul, 34660 Turkey; <sup>9</sup>Division of Medical Genetics, Departments of Pediatrics and Obstetrics & Gynecology, Medical College of Wisconsin, Milwaukee, WI 53226, USA; <sup>10</sup>Director of Andrology, Columbia Fertility Associates, Washington, DC 20037, USA; <sup>11</sup>Section of Reproductive Endocrinology, Infertility & Genetics, Department of Obstetrics & Gynecology; Reproductive Medicine and Developmental Neurobiology Programs in the Institute of Molecular Medicine and Genetics; The Medical College of Georgia, 1120 15<sup>th</sup> Street, Augusta, GA 30912, USA; <sup>12</sup>Neuroscience Program, The Medical College of Georgia, 1120 15<sup>th</sup> Street, Augusta, GA 30912, USA

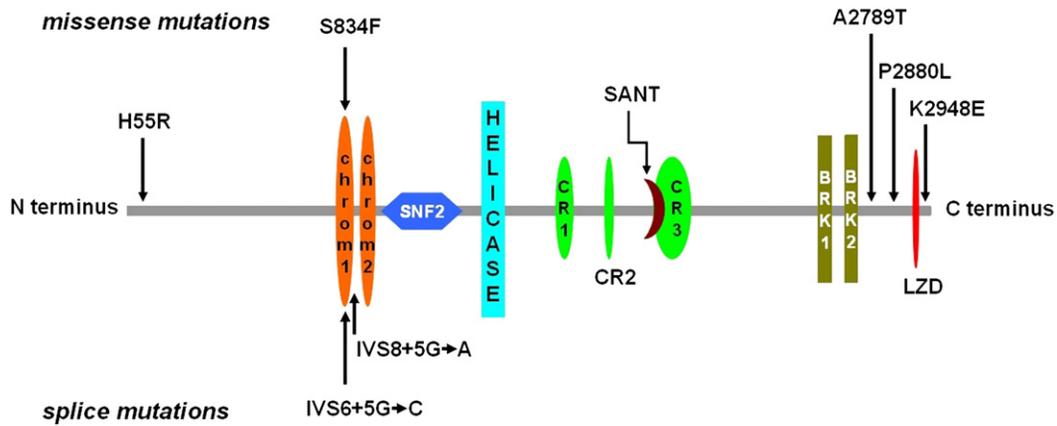
<sup>13</sup>Present address: Department of Obstetrics & Gynecology, Institute of Molecular Medicine and Genetics, The Medical College of Georgia, 1120 15<sup>th</sup> Street, Augusta, GA 30912, USA

<sup>14</sup>Present address: Constellation Pharmaceuticals, 148 Sidney Street, Cambridge, MA 02139, USA

<sup>15</sup>Present address: Forschungszentrum Karlsruhe, Institute for Nanotechnology, PO Box 3640, 76021 Karlsruhe, Germany

\*Correspondence: llayman@mccg.edu

DOI 10.1016/j.ajhg.2008.09.005. ©2008 by The American Society of Human Genetics. All rights reserved.



**Figure 1. CHD7 Domains and Positions of Mutations**

CHD7 structure with functional domains and the positions of five missense mutations and two splice-donor site mutations identified in IHH and KS patients are shown. The following abbreviations are used: chrom1, chromodomain 1; chrom2, chromodomain 2; SANT, SANT DNA binding domain; and LZD, leucine zipper domain. The function of CR1-CR3 and BRK domains are unknown. Three mutations affect chromodomains. Relative sizes and locations of domains are to scale.

(MIM 214800), a multisystem autosomal-dominant or sporadic disorder consisting of eye coloboma, heart defects, choanal atresia, retardation of growth and development, genito-urinary anomalies, and ear abnormalities (vestibular and auditory).<sup>15</sup> The large 2997 amino acid (AA) CHD7 protein contains two chromodomains at its N terminus, followed by centrally located SNF2 and helicase domains; three conserved region (CR) domains; a switching-defective protein 3, adaptor 2, nuclear receptor corepressor, transcription factor IIIB (SANT) domain; two Brahma and Kismet (BRK) domains; and, at the C terminus, a leucine-zipper domain, which we identified by using PSORTII prediction software (Figure 1A). The leucine-zipper domain is located in AA 2888–2909 (LAFNPFLSTMAPGLFYPSMFL). CHD7 belongs to a family of nine CHD proteins that have in common the ability to utilize ATP hydrolysis to alter nucleosome structure.<sup>16</sup> Chromodomains have been thought to mediate chromatin interactions and have been found to interact with DNA, RNA, and histone targets.<sup>17</sup>

CHD7 is expressed in the disease-associated organs of CHARGE syndrome, but also in KS-relevant tissues including the olfactory epithelium<sup>18</sup> and pituitary<sup>19</sup> in mice, as well as the olfactory nerve and bulb, hypothalamus, and pituitary in humans.<sup>20</sup> Prior to the identification of CHD7, hypogonadism was reported to be associated with CHARGE syndrome, including some patients old enough to be diagnosed with IHH, as defined below.<sup>21</sup> Several other studies implicate a possible connection between the KS phenotype and CHARGE syndrome, although it is important to emphasize that these patients with CHARGE syndrome are not yet of pubertal age. Anosmia or hyposmia has been identified in children with CHARGE syndrome,<sup>22</sup> as well as abnormal olfactory bulbs by magnetic resonance imaging (MRI).<sup>23</sup> Most boys studied had micropenis and/or cryptorchidism suggestive of IHH, although all subjects were prepubertal.<sup>23</sup> None of seven females under 12 years of age began puberty spontaneously, and hormonal data

suggested that IHH could ensue as they got older.<sup>23</sup> In fact, recently one female with CHARGE syndrome due to a CHD7 mutation and with some features of KS has been reported.<sup>24</sup> However, there has been no systematic evaluation of IHH/KS patients without a diagnosis of CHARGE syndrome for the presence of CHD7 mutations. We hypothesized that IHH/KS could be a milder allelic variant of CHARGE syndrome.

We first looked at *Chd7* mRNA expression in three different mouse GnRH neuronal cell lines, two migratory (GN11 and NLT)<sup>25</sup> and one postmigratory (GT1-7).<sup>26</sup> *Chd7*-mRNA expression was confirmed by cloning and sequencing of the PCR products (not shown). Next, RT-PCR on RNA extracted from KS-relevant rat tissues demonstrated expression in the olfactory bulb, medial basal hypothalamus, and pituitary (Figure S1 available online). Therefore, we undertook an extensive mutation analysis in IHH/KS patients without clinical features of CHARGE syndrome.

IHH was defined as the absence of puberty in females  $\geq$  17 yr old and in males  $\geq$  18 yr old with low serum gonadotropins as described previously.<sup>1</sup> Males had low serum testosterone ( $<$ 100 ng/dL, with normal being 300–1100), and females had hypoestrogenic amenorrhea, and remaining pituitary function and CNS imaging by MRI or CT were normal.<sup>1</sup> KS was defined as IHH and anosmia or hyposmia, as detailed by the University of Pennsylvania Smell Identification Test when possible or by patient history. This study was approved by the Human Assurance Committee (Medical College of Georgia), and signed written consent was obtained from all participants. Previously, mutation screening was performed in most patients for *KALI*, *GNRHR*, and *FGFR1* genes.<sup>14,27–30</sup>

DNA sequencing of the entire 37 protein-coding exons and splice junctions of the *CHD7* gene from 101 affected individuals (50 with KS and 51 with normosmic IHH) revealed six heterozygous mutations—one splice-donor site alteration and five missense mutations for a prevalence

**Table 1. CHD7 Mutations in Sporadic Patients with IHH/KS**

Patient	Gender and Phenotype	Geographic Origin	Exon or Intron	Nucleotide Change	Amino Acid Change	Confirmatory Method
TT20	Male, IHH; no other anomalies	Turkey	Intron 8	IVS8+5G→A	Premature termination	0/180 controls & 0/96 Turkish controls
C187	Female, KS, cleft lip and palate, hearing loss	USA	Intron 6	IVS6+5G→C	22 amino acid deletion	De novo, 0/180 controls, and 0/96 Turkish controls
C59	Male, IHH and cleft lip; cryptorchidism	USA	Exon 8	c.2501C→T	Ser834Phe	0/180 controls
C148	Male, KS; no other anomalies	USA	Exon 2	c.164A→G	His55Arg	0/180 controls
T47	Male, IHH, myopia; no other anomalies	Turkey	Exon 38	c.8365G→A	Ala2789Thr	0/180 controls and 0/96 Turkish controls
C137	Male, IHH; cryptorchidism	USA	Exon 38	c.8639C→T	Pro2880Leu	0/180 controls and 0/96 Turkish controls
C26	Male, KS; no other anomalies	USA	Exon 38	c.8842A→G	Lys2948Glu	0/180 controls

of 6% (Table 1; Figures 1 and 2; Figure S2). Another heterozygous splice-donor site mutation was identified in one of 96 additional IHH/KS patients screened for mutations only in exons 6–10. This gene region was selected for a focused mutation screen because these exons encode the highly conserved and important functional chromodomains (Figure 1). All cases with mutations were sporadic, and none of the probands had any affected family members or relatives. No identical nucleotide changes were identified in 180 controls or listed in the single-nucleotide polymorphism (SNP) database. In addition, the two *CHD7* mutations identified in probands of Turkish descent were also absent in 96 Turkish controls. Six new SNPs (Table S1, submitted into the SNP database in NCBI) in the *CHD7* were also detected. None of the patients with a *CHD7* mutation were known to possess a mutation in another gene involved in IHH/KS.

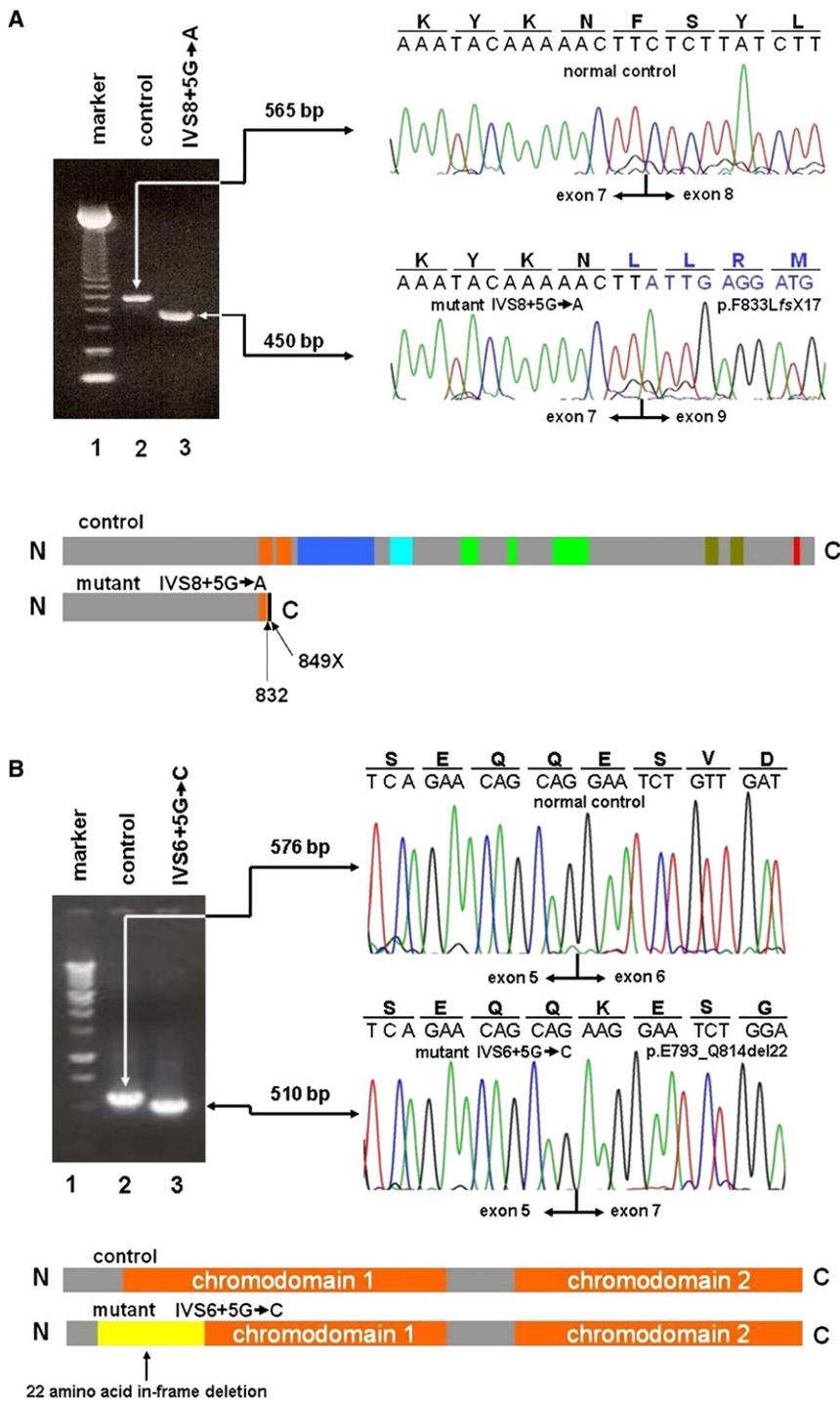
Both intronic mutations impaired mRNA splicing in lymphoblasts of affected individuals and are predicted to interfere with protein function. An IVS8+5G→A mutation identified in a normosmic IHH male without any other anomalies resulted in exon 8 skipping, as determined by RT-PCR of *CHD7* exons 4–10 from patient and control lymphoblastoid RNA, subsequent subcloning, and DNA sequencing. This mutant introduces 16 aberrant out-of-frame AA residues, causing a frameshift and subsequent premature termination codon at residue 849 located 49 bp downstream of the junction of exons 7 and 9. This results in a *CHD7* protein predicted to be truncated more than 70% of the C terminus (Figure 2A). This transcript is not a normal splice variant because it was absent in four control lymphoblastoid cell lines, and nor was the same G→A transition identified in 180 normal controls. Importantly, this mutation removes about half of the first and all of the second chromodomains, as well as other important domains of the protein, thereby predicting a nonfunctional protein.

In a KS female with mild sensorineural deafness and cleft lip and palate, a de novo heterozygous intronic transversion, IVS6+5G→C, was identified. The mutation was absent in both parents. For direct examination of whether

this variant affects splicing, *CHD7* exons 4–9 were similarly analyzed by RT-PCR. In addition to the normally spliced transcript, a transcript with reduced size demonstrated exon 6 skipping not observed in four controls (Figure 2B). This results in an in-frame deletion of 22 of 66 AAs of chromodomain 1. Chromodomains are evolutionarily conserved<sup>16</sup> and known to interact with histone tails.<sup>31,32</sup> Interestingly, this mutation has also been reported in a patient with CHARGE syndrome.<sup>33</sup> The patient with CHARGE syndrome has additional severe phenotypic findings, including the absence of earlobes, triangular concha, vestibular disturbance, autism-spectrum disorder, and mental retardation,<sup>33</sup> which are absent in our KS patient.

The five missense mutations affect highly conserved AA residues when compared with known *CHD7* orthologs (Figure 3A). Four of five point mutations were predicted to be deleterious by SIFT<sup>34</sup> (Ser834Phe, Lys2948Glu, Pro2880Leu, and His55Arg), whereas Ala2789Thr was tolerated. Ser834Phe, located in DNA-binding chromodomain 1 (UniProtKB/Swiss-Prot entry Q9P2D1; INTERPRO IPR000953),<sup>35,36</sup> a highly conserved sequence motif observed in a variety of animal and plant species, had the highest confidence measure. Further supportive evidence of the deleterious effect of Ser834Phe comes from the report of the same mutation in three patients from a family with a severe CHARGE syndrome phenotype.<sup>37</sup>

Structural models for the relevant domains suggest that Ala2789Thr and Pro2880Leu, located in the spacer sequence between the BRK2 and leucine-zipper regions, as well as Lys2948Glu, are also detrimental. All three AA residues are located in loop regions so that mutations of these residues will most likely affect structural and binding properties of the domains to their interaction partners (Figures 3B<sub>1</sub> and 3B<sub>2</sub>). The local secondary structure of the region around Pro2880 is a random coil, and the Pro2880Leu mutation induces helix formation in this region, predicting its deleterious effect (data not shown). The Ser834 residue is also located in a loop region, coordinating strongly with adjacent residues in the neighboring helix (Tyr881, context PDYV), so Phe834 will therefore strongly affect



## Figure 2. CHD7 Exon Skipping in Two Patients

(A) RT-PCR analysis of *CHD7* exon-skipping event in a patient with *IVS8+5G→A*. RT-PCR analysis confirms that the *IVS8+5G→A* mutation of *CHD7* causes aberrant exon 8 skipping in an IHH patient. Splicing patterns are compared between a normal control and the mutant by cloned cDNA sequencing. An expected *CHD7* product of 565 bp consisting of exons 4–10 is observed in the control (lane 2), whereas an abnormal product of 450 bp skipping exon 8 (115 bp) is observed in a patient with *IVS8+5G→A* (lane 3). Exon 9 nucleotide and out-of-frame AA sequence are depicted in blue. A 123 bp DNA marker is shown in the first lane.

The aberrant exclusion of exon 8 (115 bp) in a patient with *IVS8+5G→A* is predicted to introduce a frameshift in the coding region and a subsequent premature termination codon at 16 AAs downstream from exon 8 skipping. In the truncated *CHD7*, 16 out-of-frame AA residues generated by the frameshift are depicted as a black bar at the end. The functional domains from N terminus to C terminus are depicted in color as follows: orange, chromodomains 1 and 2; navy blue, SNF2; sky blue, helix; green, CR1, CR2, and CR3; deep green, BRK1 and BRK2; and red, leucine zipper domain. The SANT domain within CR3 was not depicted here.

(B) RT-PCR analysis of *CHD7* exon skipping event in a patient with *IVS6+5G→C*. RT-PCR analysis shows that mutation *IVS6+5G→C* of *CHD7* causes aberrant exon 6 skipping in a KS patient. Splicing patterns are compared between a normal control and the mutant by cloned cDNA sequencing. An expected *CHD7* product of 576 bp consisting of exons 4–9 is observed in the control (lane 2), whereas an abnormal product of 510 bp skipping exon 6 (66 bp) is observed in a patient with *IVS6+5G→C* (lane 3). A 1 kb DNA marker is shown in the first lane.

The abnormal exon 6 (66 bp) skipping in a patient with *IVS6+5G→C* causes a 66 bp in-frame deletion of 22 AAs. From chromodomain 1 comprising 66 AAs, 16 residues are deleted. Protein structure of the deleted region is depicted in yellow and chromodomains 1 and 2 in orange.

protein stability (Figure 3B<sub>3</sub>). The Ala2789T residue is conserved in seven of eight orthologs, and Lys2948 is relatively conserved (Figure 3A).

SIFT was also used to characterize the functional significance of a 22 amino acid (ESVDAEGPVVEKIMSSRSVKKQ)

in-frame deletion of exon 6 (*IVS6+5G→C* mutant) by predicting deleterious effects for almost all possible AA substitutions in this region (data not shown). The splice mutant *IVS8+5G→A* truncates the functional part of the protein starting at AA 810, including the region from AAs

920–1490. The latter is highly homologous to the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54 (PDB code: 1Z3I)<sup>38</sup> and therefore an essential DNA-binding domain of the wild-type. Collectively, involvement of the conserved chromodomain by three mutations, two of which have been identified in patients with CHARGE syndrome,<sup>33,37</sup> as well as SIFT AA conservation and protein structural analysis, indicates that these nucleotide alterations are pathogenic mutations.

To investigate developmental expression of *Chd7* further in the mouse embryo, in situ hybridization analysis was performed from embryonic day 10.5 (E10.5) to postnatal day 0 (P0). Already by E10.5 and E11, high *Chd7* expression was preferentially observed in the developing nervous system and its derivatives. The entire neuroepithelium was strongly labeled at that stage (Figures 4A–4C). At E10.5 and E11, high levels of expression were particularly noteworthy in the olfactory placode (Figures 4A and 4C), which gives rise to olfactory and GnRH neurons. With progressing differentiation, the label intensity declined and the signal became restricted to specific locations in the developing central and peripheral nervous systems. At E14, high signal intensity was seen in the olfactory epithelium, cochlea, anterior pituitary, and the spinal cord (Figures 4D and 4E). This pattern of expression is consistent with involvement of *Chd7* in the development of the olfactory pathway and the GnRH-positive neurons.<sup>39</sup> In the adult brain, intense *Chd7* expression was restricted to the granule cell layer of the cerebellum (Figure 4H), hippocampal formation (Figure 4I), and hypothalamus (Figures 4F and 4G). Signals were absent when the sense control was used.

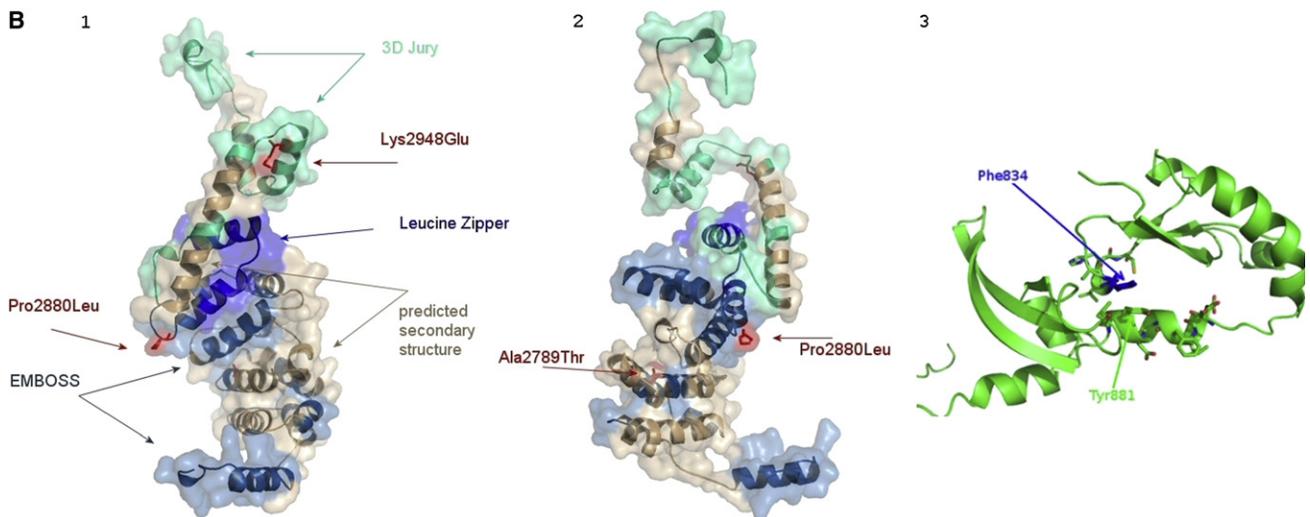
The molecular basis for 70%–75% of IHH/KS patients remains unknown.<sup>2</sup> To date, only *FGFR1* mutations have been reported to cause either normosmic IHH families or KS families.<sup>7</sup> Although a homozygous *PROKR2* deletion was seen in a single family comprising both normosmic and anosmic patients, this represents variable expressivity within the same family.<sup>40</sup> In our case, we have three unrelated probands with KS and four unrelated probands with IHH with *CHD7* mutations, demonstrating that *CHD7* is involved in either IHH or KS. We present new evidence for a role of CHD7 in the pathophysiology of both normosmic IHH and KS patients without a CHARGE phenotype. We first demonstrate *Chd7* mRNA expression in both migratory and postmigratory GnRH neuronal cell lines. We also document mRNA expression in the hypothalamus, pituitary, and olfactory bulb in the rat—all of which are IHH/KS-relevant organs. By in situ hybridization, we confirm mRNA expression at the appropriate time when GnRH and olfactory neurons migrate from the olfactory placode region to the hypothalamus. This begins on E10–E11 in the mouse and is virtually completed by E18.5.<sup>39</sup> In humans, GnRH migration begins at E5.5–E6.5 weeks and finishes 6–8 weeks postnatally.<sup>39</sup> Whether or not CHD7 affects this important developmental neuron migration requires future study.

Furthermore, we demonstrate sporadic heterozygous *CHD7* mutations, which were not present in  $\geq 180$  controls or the SNP database, in humans with both normosmic IHH and anosmic IHH (KS). Additionally, two mutations in Turkish patients were also absent in 96 Turkish controls (for a total of 276 controls). The prevalence of *CHD7* mutations of ~6% is similar to that of apparently sporadic *KAL1* mutations<sup>2</sup> and somewhat less than the 10% reported rate of *FGFR1* mutations in IHH/KS.<sup>7</sup> Two of our *CHD7* mutations alter mRNA splicing in lymphoblast RNA and predict deletions of chromodomains, whereas another missense mutation within chromodomain 1 affects a highly conserved Ser residue. The disruption of these important evolutionarily conserved chromodomains is highly likely to result in deleterious consequences,<sup>16</sup> because chromodomains are known to interact with histone tails.<sup>31,32</sup> Chromodomain deletion has also been reported to impair nucleosome binding and remodeling by CHD proteins,<sup>41</sup> indicating that their disruption will be detrimental to CHD7 function.

Further supportive evidence that our *CHD7* mutations are deleterious comes from SIFT analysis, which indicates that four of the five missense mutations (Ser834Phe, Lys2948Glu, Pro2880Leu, and His55Arg) involve highly conserved AA residues among known species and, therefore, are not likely to be tolerated by their observed substitutions.<sup>34</sup> These findings were also corroborated by protein structural analysis of the AA variants Ala2789Thr, Pro2880Leu, and Lys2948Glu, which were predicted to alter structural and binding properties of the domains. Taken together, both AA conservation and protein structural modeling provide additional support that these missense substitutions are highly likely to be deleterious mutations.

Importantly, our one IHH and one KS patient, who both lack the CHARGE phenotype, possess the same mutations (Ser834Phe and IVS6+5G→C) reported previously in patients with CHARGE syndrome,<sup>33,37</sup> further demonstrating the allelic relationship of both syndromes. The KS patient with the IVS6+5G→C mutation does not fulfill Blake's criteria for CHARGE syndrome,<sup>42</sup> although she does have hearing impairment and cleft lip and palate. This also indicates that the effects of modifying genes may determine whether the patient has the more severe CHARGE phenotype rather than the milder IHH/KS phenotype. Interestingly, our mutations have been localized to regions around four exons—2, 6, 8, and 38—suggesting the possibility of hotspots for IHH/KS mutation. Because of these findings and the absence of nonsense mutations, which often occur in CHARGE syndrome,<sup>15</sup> we provide the first convincing evidence that IHH/KS represents a milder allelic variant of CHARGE syndrome. Although its precise function is uncertain, CHD7 appears to be important in GnRH and olfactory neuron migration to their embryologic destination in the hypothalamus. CHD7 is the first chromatin-remodeling protein involved in normal puberty in humans and is the second gene (after *FGFR1*) identified that results in both normosmic IHH and KS.

<b>A</b>	<b>H55R</b>		<b>S834F</b>		
		*		*	
human	DQGFASLQ <b>FSLHHPST</b> NQNQTKLTHFDHY	72	human	IEEFYVKYKN <b>FSYLHCQWAS</b> IEDLEKDKR	851
chimpanzee	DQGFASLQ <b>FSLHHPST</b> NQNQTKLTHFDHY	116	chimpanzee	IEEFYVKYKN <b>FSYLHCQWAS</b> IEDLEKDKR	895
monkey	DQGFASLQ <b>FSLHHPST</b> NQNQTKLTHFDHY	72	monkey	IEEFYVKYKN <b>FSYLHCQWAS</b> VEDLEKDKR	851
mouse	DQGFASLQ <b>FSLHHPST</b> NQNQTKLTHFDHY	72	mouse	VEEFYVKYKN <b>FSYLHCQWAS</b> VEDLEKDKR	841
dog	DQGFASLQ <b>FSLHHPST</b> NQNQTKLTHFDHY	72	dog	VEEFYVKYKN <b>FSYLHCQWAS</b> VEDLEKDKR	850
cow	DQGFASLQ <b>FSLHHPST</b> NQNQTKLTHFDHY	72	cow	VEEFYVKYKN <b>FSYLHCQWAS</b> IEDLEKDKR	799
chicken	DQGFASLQ <b>FSLHHP</b> PAQNQAKLTHFDHY	72	chicken	IEEFYVKYKN <b>FSYLHCQWAS</b> VEELD <del>KDKR</del>	852
	<b>A2789T</b>		<b>P2880L</b>		
		*		*	
human	GLATAATAGGDA <b>KNPAAVLPMLPGMAGL</b>	2805	human	SANGSVGAAT <b>AP</b> -----AGLPS	2885
chimpanzee	GLATAATAGGDA <b>KNPAAVLPMLPGMAGL</b>	2849	chimpanzee	SANGSVGAAT <b>AP</b> -----AGLPS	2929
monkey	GLATAATAGGDA <b>KNPAAVLPMLPGMAGL</b>	2759	monkey	SANGSVGAAT <b>AP</b> -----AGLPS	2839
rat	GLATAATAGGDA <b>KGPAAVLPMLPGMAGL</b>	1380	rat	SANGSVGAAT <b>AP</b> -----AALPS	1460
mouse	GLATAATAGGDA <b>KSPAAVLPMLPGMAGL</b>	2794	mouse	SANGSVGAAT <b>AP</b> -----AALPS	2874
dog	GLATAAAAGGD <b>TKNPAAVLPMLPGVAGL</b>	2803	dog	SANGSVGAAT <b>AP</b> -----AGLPS	2883
cow	GLATAAAAGGD <b>AKNPAAVLPMLPGVAGL</b>	2746	cow	SANGSVGAAT <b>GE</b> -----AGLPS	2826
chicken	GLATAAAAGGD <b>AKNPAA</b> MLPLMLPGMAGL	2807	chicken	SANGSV <b>SAATA</b> ATTATATTTTTTTNTELEPT	2899
	<b>K2948E</b>				
		*			
human	AADKAEGG <b>PF</b> KDGETLEGSDAEESLDKTA	2966			
chimpanzee	AADKAEGG <b>PF</b> KDGETLEGSDAEESLDKTA	3010			
monkey	AADKAEGG <b>PF</b> KDGETLEGSDAEESLDKTA	2921			
rat	AADKAEGG <b>PF</b> KDGETLEGSDAEENLDKTT	1541			
mouse	AADKAEGG <b>PF</b> KDGETLEGSDAEENLDKTV	2955			
dog	APDKAEGG <b>AFQ</b> EEENLEGSDAEENLDKTA	2964			
cow	APDKAEGG <b>AFQ</b> EDEPLEGSDAEESLDKTA	2907			
chicken	ATDKTEGT <b>AKK</b> DEENLEGSDAEESLDKTA	2980			

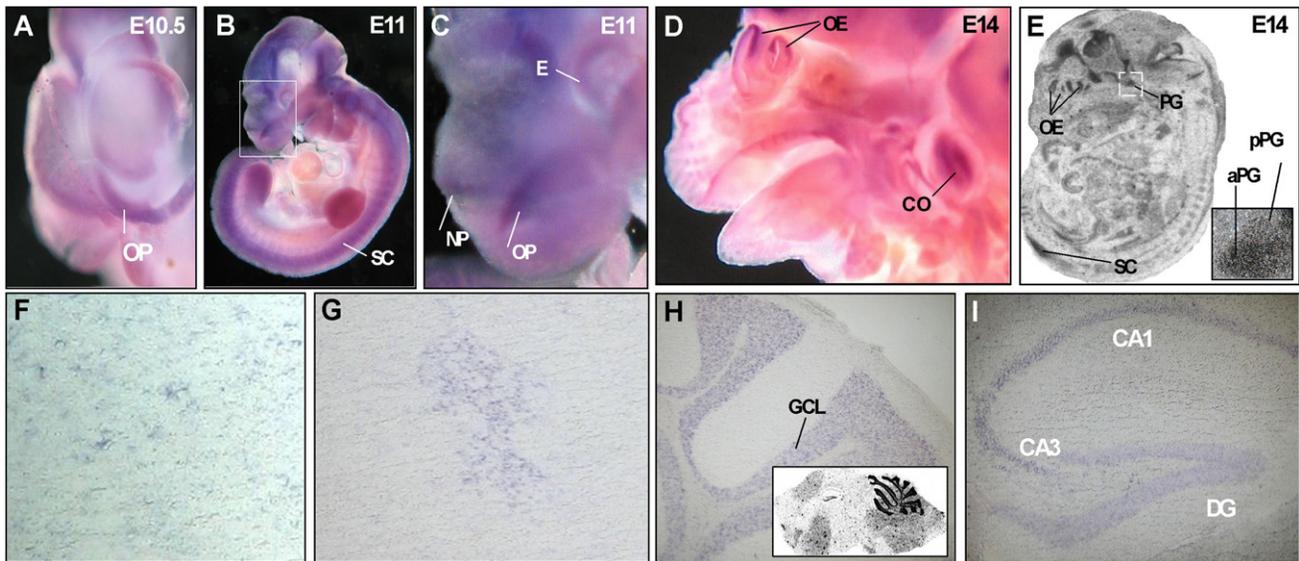


### Figure 3. Protein Sequence Alignment and Structure Modeling of CHD7

(A) Multiple protein-sequence alignment of CHD7 with its orthologs. The positions of residues affected by missense mutations in IHH/KS patients are marked by asterisks and bold letters in available CHD7 animal orthologs. H55 and S834 are evolutionarily fully conserved, whereas A2789 and P2880 are highly conserved. K2948 shows relative conservation. Human CHD7 N-terminal residues 1–1423 are missing in the predicted rat Chd7 protein (NP\_001101376). Blue shading represents the invariant residues that match the consensus exactly, and pink shading shows partial matching.

(B) CHD7 structure modeling.

(B<sub>1</sub> and B<sub>2</sub>) Shown are alternate views of the model of the 300 AA C-terminal region of CHD7 based on the 3DJURY model, which results from the alignment of this region with that of mdia1 gbd-fh3 in complex with rhoc (1z2c/B). Regions of the model that were predicted solely on the basis of homology in 3DJURY are shown in green, but additional information was used to substantiate the model. There is an additional long consecutive region where EMBOSS found substantial sequence similarity to the CHD7 sequence—these regions are marked in dark blue. Regions that, in addition to the alignment, agree in their secondary structure with the consensus prediction of PHYRE are shown in light brown. Finally, we found a region that, in addition to the alignment and secondary structure, exhibits a motif commensurate with the leucine zipper (bright blue). Sites and side chains of the three mutations Lys2948Glu, Pro2880Leu, and Ala2789Thr are shown in red for better visibility. All mutation sites lie in highly flexible loop regions. (B<sub>1</sub>) and (B<sub>2</sub>) show the same model rotated 180° around the vertical axis though the center of the molecule.



#### Figure 4. *Chd7* Expression during Murine Development

The following abbreviations are used: OP, olfactory placode; GCL, granule cell layer of the cerebellum; OE, olfactory epithelium; aPG, anterior pituitary gland; pPG, posterior pituitary gland; SC, spinal cord; CA1, CA1 region of the hippocampus; CA3, CA3 region of the hippocampus; DG, dentate gyrus; CO, cochlea; E, eye; and NP, nasal pit.

(A) DIG-labeled whole-mount in situ hybridization using a *Chd7* antisense probe shows high expression in the olfactory placode at E10.5. (B and C) Strong labeling of neuroepithelial structures is seen at E11, shown in higher magnification in (C).

(D and E) Expression of *Chd7* in the olfactory epithelium, developing cortex, cochlea, spinal cord, and anterior pituitary gland at E14. [35S]-UTP-labeled in situ hybridizations are shown in (E). Inset shows magnification of the pituitary gland.

(F–I) Expression of *Chd7* in the adult brain: Singular DIG-positive cells are found scattered within hypothalamic nuclei of the preoptic area (F). One cluster of cells within the medial preoptic area consistently stained positively for *Chd7* (G). DIG-labeled in situ hybridizations on 12  $\mu$ m cryosections show expression in the cerebellum (H). The inset in (H) shows an overview of the adult brain with a [35S]-UTP labeled *Chd7* antisense probe. The hippocampal region also stains positive for *Chd7* (I).

#### Supplemental Data

Supplemental Data include two figures and two tables and can be found with this article online at <http://www.ajhg.org/>.

#### Acknowledgments

We thank all affected individuals and their families for their cooperation. The authors also thank I. Hermans-Borgmeyer for unrestricted support with the in situ hybridizations and interpretation of the data; B. Dierkes for excellent technical assistance; G. Schnitzler for helpful discussions; E. Moutevelis for leucine-zipper-domain structure analysis; the Computational Science Center, KIST (Seoul, Korea), and the volunteers from POEM@HOME for computational resources; Y. Shen for RT-MLPA; S. West for zinc-finger prediction analysis; L. Chorich for her efforts in supervising and running the laboratory; and E. Shah for RT-PCR analysis.

We also acknowledge support to L.C.L. from National Institutes of Health (NIH) grants HD33004 and HD040287, as well as the Medical College of Georgia Research Institute (MCGRI), Dean

D.D. Miller, Institute of Molecular Medicine and Genetics (IM-MAG) director R. Yu, and ob/gyn chair A.A. Murphy at MCG. J.F.G. has support from NIH grant GM061354, W.W. from Deutsche Forschungsgemeinschaft (DFG) grant WE1863/10-2 and the DFG Center for Functional Nanostructures, and I.M. from the EU Leonardo program "Unipharm Graduates."

Received: July 29, 2008

Revised: September 8, 2008

Accepted: September 15, 2008

Published online: October 2, 2008

#### Web Resources

The URLs for data presented herein are as follows:

dbSNP, <http://www.ncbi.nlm.nih.gov/SNP/>

NCBI and GenBank, <http://www.ncbi.nlm.nih.gov/>

Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/>

PSORTII, <http://psort.hgc.jp/form2.html>

SIFT, <http://blocks.fhcr.org/sift/SIFT.html>

(B<sub>3</sub>) Shown is the model of the 155 AA region around Ser834Phe (AA 795–950). The side chain of Phe834 is highlighted in dark blue for better visibility. In addition to the secondary and tertiary structure, side chains of the AA in a 6 Å radius of Phe834 are shown in the standard color coding (carbon green, oxygen red, sulfur yellow). Of particular importance is the strong stacking interaction of Phe834 (blue) with the ring of Tyr881 (green) directly underneath.

## References

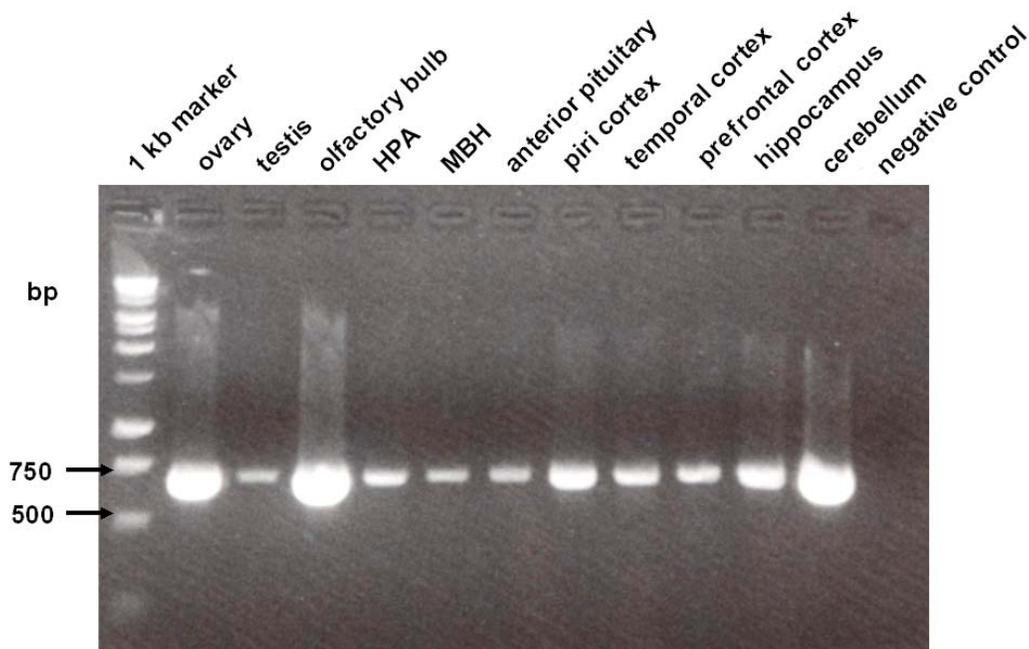
1. Bhagavath, B., Podolsky, R.H., Ozata, M., Bolu, E., Bick, D.P., Kulharya, A., Sherins, R.J., and Layman, L.C. (2006). Clinical and molecular characterization of a large sample of patients with hypogonadotropic hypogonadism. *Fertil. Steril.* **85**, 706–713.
2. Kim, H.G., Bhagavath, B., and Layman, L.C. (2008). Clinical Manifestations of Impaired GnRH Neuron Development and Function. *Neurosignals* **16**, 165–182.
3. Franco, B., Guioli, S., Pragliola, A., Incerti, B., Bardoni, B., Tonlorenzi, R., Carozzo, R., Maestrini, E., Pieretti, M., Tailon-Miller, P., et al. (1991). A gene deleted in Kallmann's syndrome shares homology with neural cell adhesion and axonal path-finding molecules. *Nature* **353**, 529–536.
4. Legouis, R., Hardelin, J.P., Levilliers, J., Claverie, J.M., Compain, S., Wunderle, V., Millasseau, P., Le Paslier, D., Cohen, D., Caterina, D., et al. (1991). The candidate gene for the X-linked Kallmann syndrome encodes a protein related to adhesion molecules. *Cell* **67**, 423–435.
5. Dode, C., Levilliers, J., Dupont, J.M., De Paepe, A., Le Du, N., Soussi-Yanicostas, N., Coimbra, R.S., Delmaghani, S., Compain-Nouaille, S., Baverel, F., et al. (2003). Loss-of-function mutations in FGFR1 cause autosomal dominant Kallmann syndrome. *Nat. Genet.* **33**, 463–465.
6. Kim, H.G., Herrick, S.R., Lemyre, E., Kishikawa, S., Salisz, J.A., Seminara, S., MacDonald, M.E., Bruns, G.A., Morton, C.C., Quade, B.J., et al. (2005). Hypogonadotropic hypogonadism and cleft lip and palate caused by a balanced translocation producing haploinsufficiency for FGFR1. *J. Med. Genet.* **42**, 666–672.
7. Pitteloud, N., Acierno, J.S., Jr., Meysing, A., Eliseenkova, A.V., Ma, J., Ibrahim, O.A., Metzger, D.L., Hayes, F.J., Dwyer, A.A., Hughes, V.A., et al. (2006). Mutations in fibroblast growth factor receptor 1 cause both Kallmann syndrome and normosmic idiopathic hypogonadotropic hypogonadism. *Proc. Natl. Acad. Sci. USA* **103**, 6281–6286.
8. de Roux, N., Young, J., Misrahi, M., Genet, R., Chanson, P., Schaison, G., and Milgrom, E. (1997). A family with hypogonadotropic hypogonadism and mutations in the gonadotropin-releasing hormone receptor. *N. Engl. J. Med.* **337**, 1597–1602.
9. Layman, L.C., Cohen, D.P., Jin, M., Xie, J., Li, Z., Reindollar, R.H., Bolbolan, S., Bick, D.P., Sherins, R.J., Duck, L.W., et al. (1998). Mutations in the gonadotropin-releasing hormone receptor gene cause hypogonadotropic hypogonadism. *Nat. Genet.* **18**, 14–15.
10. Seminara, S.B., Messenger, S., Chatzidaki, E.E., Thresher, R.R., Acierno, J.S. Jr., Shagoury, J.K., Bo-Abbas, Y., Kuohung, W., Schwino, K.M., Hendrick, A.G., et al. (2003). The GPR54 gene as a regulator of puberty. *N. Engl. J. Med.* **349**, 1614–1627.
11. Muscatelli, F., Strom, T.M., Walker, A.P., Zanaria, E., Recan, D., Meindl, A., Bardoni, B., Guioli, S., Zehetner, G., Rabl, W., et al. (1994). Mutations in the DAX-1 gene give rise to both X-linked adrenal hypoplasia congenita and hypogonadotropic hypogonadism. *Nature* **372**, 672–676.
12. Dode, C., Teixeira, L., Levilliers, J., Fouveaut, C., Bouchard, P., Kottler, M.L., Lespinasse, J., Lienhardt-Roussie, A., Mathieu, M., Moerman, A., et al. (2006). Kallmann syndrome: Mutations in the genes encoding prokineticin-2 and prokineticin receptor-2. *PLoS Genet* **2**, e175.
13. Pitteloud, N., Quinton, R., Pearce, S., Raivio, T., Acierno, J., Dwyer, A., Plummer, L., Hughes, V., Seminara, S., Cheng, Y.Z., et al. (2007). Digenic mutations account for variable phenotypes in idiopathic hypogonadotropic hypogonadism. *J. Clin. Invest.* **117**, 457–463.
14. Bhagavath, B., Ozata, M., Ozdemir, I.C., Bolu, E., Bick, D.P., Sherins, R.J., and Layman, L.C. (2005). The prevalence of gonadotropin-releasing hormone receptor mutations in a large cohort of patients with hypogonadotropic hypogonadism. *Fertil. Steril.* **84**, 951–957.
15. Vissers, L.E., van Ravenswaaij, C.M., Admiraal, R., Hurst, J.A., de Vries, B.B., Janssen, I.M., van der Vliet, W.A., Huys, E.H., de Jong, P.J., Hamel, B.C., et al. (2004). Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat. Genet.* **36**, 955–957.
16. Marfella, C.G., and Imbalzano, A.N. (2007). The Chd family of chromatin remodelers. *Mutat. Res.* **618**, 30–40.
17. Brehm, A., Tufteland, K.R., Aasland, R., and Becker, P.B. (2004). The many colours of chromodomains. *Bioessays* **26**, 133–140.
18. Bosman, E.A., Penn, A.C., Ambrose, J.C., Kettleborough, R., Stemple, D.L., and Steel, K.P. (2005). Multiple mutations in mouse Chd7 provide models for CHARGE syndrome. *Hum. Mol. Genet.* **14**, 3463–3476.
19. Hurd, E.A., Capers, P.L., Blauwkamp, M.N., Adams, M.E., Raphael, Y., Poucher, H.K., and Martin, D.M. (2007). Loss of Chd7 function in gene-trapped reporter mice is embryonic lethal and associated with severe defects in multiple developing tissues. *Mamm. Genome* **18**, 94–104.
20. Sanlaville, D., Etchevers, H.C., Gonzales, M., Martinovic, J., Clement-Ziza, M., Delezoide, A.L., Aubry, M.C., Pelet, A., Chemouny, S., Cruaud, C., et al. (2006). Phenotypic spectrum of CHARGE syndrome in fetuses with CHD7 truncating mutations correlates with expression during human development. *J. Med. Genet.* **43**, 211–217.
21. Wheeler, P.G., Quigley, C.A., Sadeghi-Nejad, A., and Weaver, D.D. (2000). Hypogonadism and CHARGE association. *Am. J. Med. Genet.* **94**, 228–231.
22. Chalouhi, C., Faulcon, P., Le Bihan, C., Hertz-Pannier, L., Bonfils, P., and Abadie, V. (2005). Olfactory evaluation in children: Application to the CHARGE syndrome. *Pediatrics* **116**, e81–e88.
23. Pinto, G., Abadie, V., Mesnage, R., Blustajn, J., Cabrol, S., Amiel, J., Hertz-Pannier, L., Bertrand, A.M., Lyonnet, S., Rappaport, R., et al. (2005). CHARGE syndrome includes hypogonadotropic hypogonadism and abnormal olfactory bulb development. *J. Clin. Endocrinol. Metab.* **90**, 5621–5626.
24. Ogata, T., Fujiwara, I., Ogawa, E., Sato, N., Udaka, T., and Kosaki, K. (2006). Kallmann syndrome phenotype in a female patient with CHARGE syndrome and CHD7 mutation. *Endocr. J.* **53**, 741–743.
25. Radovick, S., Wray, S., Lee, E., Nicols, D.K., Nakayama, Y., Weintraub, B.D., Westphal, H., Cutler, G.B., Jr., and Wondisford, F.E. (1991). Migratory arrest of gonadotropin-releasing hormone neurons in transgenic mice. *Proc. Natl. Acad. Sci. USA* **88**, 3402–3406.
26. Mellon, P.L., Windle, J.J., Goldsmith, P.C., Padula, C.A., Roberts, J.L., and Weiner, R.I. (1990). Immortalization of hypothalamic GnRH neurons by genetically targeted tumorigenesis. *Neuron* **5**, 1–10.
27. Bhagavath, B., Xu, N., Ozata, M., Rosenfield, R.L., Bick, D.P., Sherins, R.J., and Layman, L.C. (2007). KAL1 mutations are

- not a common cause of idiopathic hypogonadotrophic hypogonadism in humans. *Mol. Hum. Reprod.* *13*, 165–170.
28. Pedersen-White, J.R., Chorich, L.P., Bick, D.P., Sherins, R.J., and Layman, L.C. (2008). The prevalence of intragenic deletions in patients with idiopathic hypogonadotropic hypogonadism and Kallmann syndrome. *Mol. Hum. Reprod.* *14*, 367–370.
  29. Xu, N., Podolsky, R.H., Chudgar, P., Chorich, L.P., Liu, C., McDonough, P.G., Warrington, J.A., and Layman, L.C. (2005). Screening candidate genes for mutations in patients with hypogonadotropic hypogonadism using custom genome resequencing microarrays. *Am. J. Obstet. Gynecol.* *192*, 1274–1282.
  30. Xu, N., Qin, Y., Reindollar, R.H., Tho, S.P., McDonough, P.G., and Layman, L.C. (2007). A mutation in the fibroblast growth factor receptor 1 gene causes fully penetrant normosmic isolated hypogonadotropic hypogonadism. *J. Clin. Endocrinol. Metab.* *92*, 1155–1158.
  31. Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* *410*, 120–124.
  32. Flanagan, J.F., Mi, L.Z., Chruszcz, M., Cymborowski, M., Clines, K.L., Kim, Y., Minor, W., Rastinejad, F., and Khorasani-zadeh, S. (2005). Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* *438*, 1181–1185.
  33. Jongmans, M.C., Hoefsloot, L.H., van der Donk, K.P., Admiraal, R.J., Magee, A., van de Laar, I., Hendriks, Y., Verheij, J.B., Walpole, I., Brunner, H.G., et al. (2008). Familial CHARGE syndrome and the CHD7 gene: A recurrent missense mutation, intrafamilial recurrence and variability. *Am. J. Med. Genet. A.* *146A*, 43–50.
  34. Ng, P.C., and Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* *31*, 3812–3814.
  35. Mulder, N.J., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bradley, P., Bork, P., Bucher, P., Cerutti, L., et al. (2005). InterPro, progress and status in 2005. *Nucleic Acids Res.* *33*, D201–D205.
  36. Wu, C.H., Apweiler, R., Bairoch, A., Natale, D.A., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., et al. (2006). The Universal Protein Resource (UniProt): An expanding universe of protein information. *Nucleic Acids Res.* *34*, D187–D191.
  37. Delahaye, A., Sznajder, Y., Lyonnet, S., Elmaleh-Berges, M., Delpierre, I., Audollent, S., Wiener-Vacher, S., Mansbach, A.L., Amiel, J., Baumann, C., et al. (2007). Familial CHARGE syndrome because of CHD7 mutation: Clinical intra- and interfamilial variability. *Clin. Genet.* *72*, 112–121.
  38. Thoma, N.H., Czyzewski, B.K., Alexeev, A.A., Mazin, A.V., Kowalczykowski, S.C., and Pavletich, N.P. (2005). Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54. *Nat. Struct. Mol. Biol.* *12*, 350–356.
  39. Schwarting, G.A., Wierman, M.E., and Tobet, S.A. (2007). Gonadotropin-releasing hormone neuronal migration. *Semin. Reprod. Med.* *25*, 305–312.
  40. Pitteloud, N., Zhang, C., Pignatelli, D., Li, J.D., Raivio, T., Cole, L.W., Plummer, L., Jacobson-Dickman, E.E., Mellon, P.L., Zhou, Q.Y., et al. (2007). Loss-of-function mutation in the prokineticin 2 gene causes Kallmann syndrome and normosmic idiopathic hypogonadotropic hypogonadism. *Proc. Natl. Acad. Sci. USA* *104*, 17447–17452.
  41. Bouazoune, K., Mitterweger, A., Langst, G., Imhof, A., Akhtar, A., Becker, P.B., and Brehm, A. (2002). The dMi-2 chromodomains are DNA binding modules important for ATP-dependent nucleosome mobilization. *EMBO J.* *21*, 2430–2440.
  42. Blake, K.D., Salem-Hartshorne, N., Daoud, M.A., and Gradstein, J. (2005). Adolescent and adult issues in CHARGE syndrome. *Clin. Pediatr. (Phila.)* *44*, 151–159.

## Supplemental Data

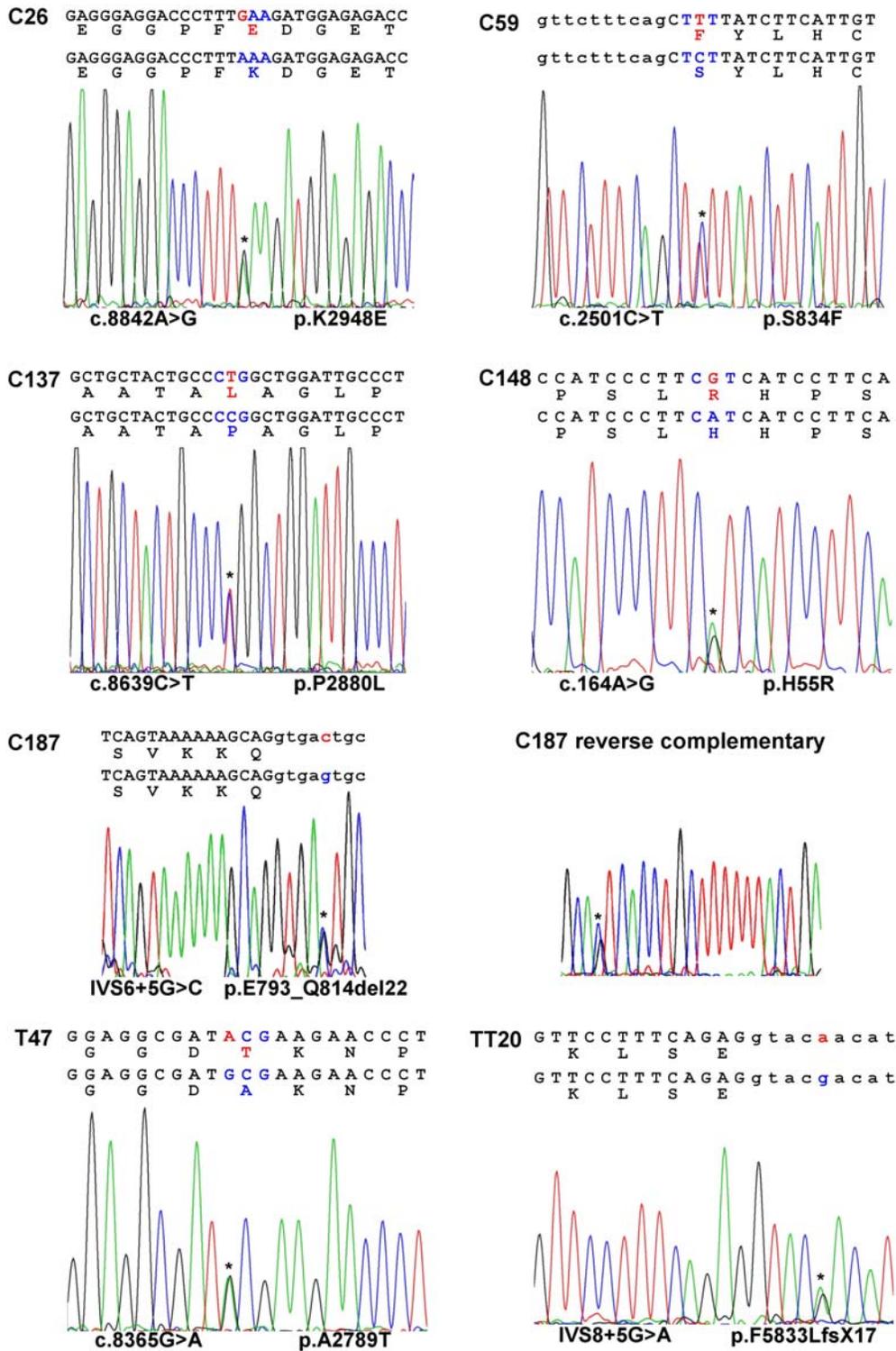
### Mutations in *CHD7*, Encoding a Chromatin-Remodeling Protein, Cause Idiopathic Hypogonadotropic Hypogonadism and Kallmann Syndrome

Hyung-Goo Kim, Ingo Kurth, Fei Lan, Irene Meliciani, Wolfgang Wenzel, Soo Hyun Eom, Gil Bu Kang, Georg Rosenberger, Mustafa Tekin, Metin Ozata, David P. Bick, Richard J. Sherins, Steven L. Walker, Yang Shi, James F. Gusella, and Lawrence C. Layman



**Figure S1 - Semi-quantitative RT-PCR analysis in rat tissues**

Total RNA isolated from 11 rat IHH/KS-relevant tissues was used for RT-PCR using rat-specific *Chd7* primers. 707bp amplicons were detected in all tissues after 35 cycles; in particular, expression levels were higher in the ovary, olfactory bulb and cerebellum. HPA: hypothalamus preoptic area; MBH: medial basal hypothalamus.



**Figure S2 - Heterozygous mutations in *CHD7***

Chromatograms show representative *CHD7* mutations in sporadic IHH/KS patients. Asterisks indicate positions of mutations. Above each are the mutant (top) and wild-type (bottom) nucleotide and AA sequences. Nucleotides and AA directly affected by

the mutations are shown in red in mutant and in blue in wild-type. Intronic and exonic nucleotide sequences are in lower and upper cases, respectively.

**Table S1 – SNPs of *CHD7* found in sporadic IHH/KS patients and normal controls**

Location	Nucleotide change	Amino-acid change	Frequency of second allele in patients	dbSNP or frequency of second allele in normal controls
exon 2	c.307T>A	Ser103Thr	3/206	8/360 (1) <sup>a</sup> ss99308621*
exon 2	c.1018A>G	Met340Val	1/206	ss68756179 3/150
exon 2	c.1397C>T	Ser466Leu	1/206	5/360 ss99308622*
exon 2	c.1467A>G	Gln489Gln	0/206	1/360 ss99308623*
intron 6	IVS6+38A>T	–	6/206	rs41272438 9/360 & 8/192 Turkish controls
intron 20	IVS20+36C>T	–	1/206	1/372 ss99308624*
intron 22	IVS22-4C>T	–	1/206	2/240 ss99308625*
intron 30	IVS30+8C>T	–	22/206	rs3763592
exon 31	c.6111C>T	Pro2037Pro	2/206	2/192 ss105110898*
exon 31	c.6135G>A	Pro2045Pro	8/206	rs6999971 8/192
exon 31	c.6276G>A	Glu2092Glu	4/206	rs2098096 12/192
exon 31	c.6282A>G	Gly2094Gly	0/206	rs41312172 1/192
exon 34	c.7356A>G	Thr2452Thr	13/206	rs2272727

<sup>a</sup> Total number of homozygotes among individuals screened

\* Newly identified SNPs

**Table S2 – Primers used in this study**

For semi-quantitative RT-PCR in rat tissues (rat <i>Chd7</i> ; NM_001107906)	
Forward primer (5'-3')	GAAGCGCTACACTGAAGATT TGGAGTTCAAGATTTCTG
Reverse primer (5'-3')	CTAAAATGCAGTTTCGCATGTTGTACCAATTGAAAAGTAGC
Amplicon size	707 bp
Annealing temperature	60 °C

RT-PCR analysis of <i>CHD7</i> exon skipping of IVS8+5G>A (human <i>CHD7</i> ; NM_017780)	
Forward primer (5'-3')	CCACCATCTC CTCCTCCTGA AGAAGATGAG
Reverse primer (5'-3')	GCTCCCACGTGCTGTCTTCATAAGGAAGTG
Amplicon size	565 bp and 450 bp
Annealing temperature	61 °C

RT-PCR analysis of <i>CHD7</i> exon skipping of IVS6+5G>C (human <i>CHD7</i> ; NM_017780)	
Forward primer (5'-3')	CAGAAGCAAG TGCTTTGAAG AAAAAGGTCA ACAAGGG
Reverse primer (5'-3')	GGTCATCTGTGCTACGTGCAAAGTCC
Amplicon size	576 bp and 510 bp
Annealing temperature	60 °C

Amplification of the <i>Chd7</i> probe for mouse <i>in situ</i> hybridization analysis (mouse <i>Chd7</i> ; NM_001081417)	
Forward primer (5'-3')	CAGGTGGCTGGAGGAGAACCC
Reverse primer (5'-3')	CTTACAGGGCCCTCCCTCGGCC
Amplicon size	836 bp
Annealing temperature	58 °C

### 12.3 Paper III

**Hyung-Goo Kim\***, Jang-Won Ahn, Ingo Kurth, Reinhard Ullmann, Hyun-Taek Kim, Anita Kulharya, Kyung-Soo Ha, Yasuhide Itokawa, Irene Meliciani, Wolfgang Wenzel, Deresa Lee, Georg Rosenberger, Metin Ozata, David P. Bick, Richard J. Sherins, Takahiro Nagase, Mustafa Tekin, Soo-Hyun Kim, Cheol-Hee Kim, Hans-Hilger Ropers, James F. Gusella, Vera Kalscheuer, Cheol Yong Choi, and Lawrence C. Layman\*

\*Corresponding author

***WDR11, a WD Protein that Interacts with Transcription Factor EMX1, Is Mutated in Idiopathic Hypogonadotropic Hypogonadism and Kallmann Syndrome***

American Journal of Human Genetics, 2010 Oct 8;87(4):465-479.

# WDR11, a WD Protein that Interacts with Transcription Factor EMX1, Is Mutated in Idiopathic Hypogonadotropic Hypogonadism and Kallmann Syndrome

Hyung-Goo Kim,<sup>1,4,15,\*</sup> Jang-Won Ahn,<sup>2</sup> Ingo Kurth,<sup>3,16</sup> Reinhard Ullmann,<sup>4</sup> Hyun-Taek Kim,<sup>5</sup> Anita Kulharya,<sup>6</sup> Kyung-Soo Ha,<sup>7</sup> Yasuhide Itokawa,<sup>8</sup> Irene Meliciani,<sup>9</sup> Wolfgang Wenzel,<sup>9</sup> Deresa Lee,<sup>2</sup> Georg Rosenberger,<sup>3</sup> Metin Ozata,<sup>10</sup> David P. Bick,<sup>11</sup> Richard J. Sherins,<sup>12</sup> Takahiro Nagase,<sup>8</sup> Mustafa Tekin,<sup>13,17</sup> Soo-Hyun Kim,<sup>14</sup> Cheol-Hee Kim,<sup>5</sup> Hans-Hilger Ropers,<sup>4</sup> James F. Gusella,<sup>15</sup> Vera Kalscheuer,<sup>4</sup> Cheol Yong Choi,<sup>2</sup> and Lawrence C. Layman<sup>1,\*</sup>

By defining the chromosomal breakpoint of a balanced t(10;12) translocation from a subject with Kallmann syndrome and scanning genes in its vicinity in unrelated hypogonadal subjects, we have identified *WDR11* as a gene involved in human puberty. We found six patients with a total of five different heterozygous *WDR11* missense mutations, including three alterations (A435T, R448Q, and H690Q) in WD domains important for  $\beta$  propeller formation and protein-protein interaction. In addition, we discovered that WDR11 interacts with EMX1, a homeodomain transcription factor involved in the development of olfactory neurons, and that missense alterations reduce or abolish this interaction. Our findings suggest that impaired pubertal development in these patients results from a deficiency of productive WDR11 protein interaction.

## Introduction

Human puberty is a dynamic process that initiates complex interactions of the hypothalamic-pituitary-gonadal axis, the purpose of which is to produce sex steroids for reproductive maturity and gametes for fertility. Any disruption of the development or regulation of this system, for which the hypothalamus serves as the master regulator through its pulsatile release of gonadotropin-releasing hormone (GnRH), can produce deleterious consequences for successful reproduction.<sup>1</sup> Patients with idiopathic hypogonadotropic hypogonadism (IHH, MIM 146110) show clinical signs and symptoms of GnRH deficiency: delayed puberty due to low sex steroid production along with low levels of serum gonadotropins. Patients with Kallmann syndrome (KS, MIM 308700, 147950, 244200, 610628, 612370, 612702) have IHH but also display an impaired sense of smell, thought to be due to the developmental failure of the migration of GnRH neurons along the olfactory axonal projections.

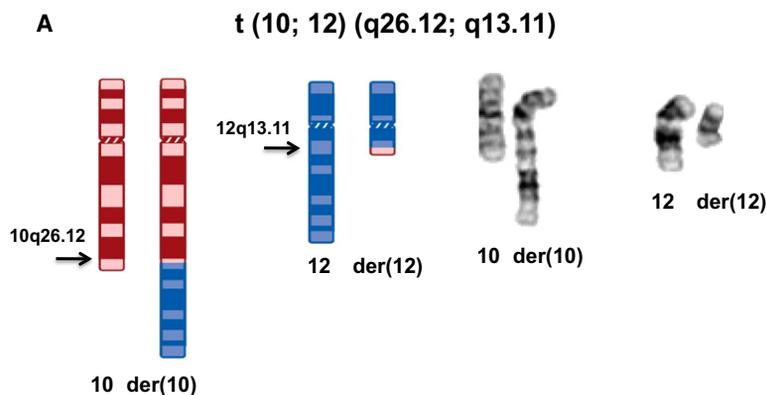
IHH/KS is one of the most common causes of hypogonadism and is genetically heterogeneous. Researchers

have used a variety of strategies to find IHH- and/or KS-causing mutations in a number of genes. Such strategies have included linkage analysis, deletion mapping, and candidate gene analysis. The discovery of a rare family with males displaying both X-linked KS and ichthyosis led to the identification of *KAL1* (MIM 308700) by positional cloning,<sup>2,3</sup> and characterization of deletions<sup>4</sup> and a balanced translocation<sup>5</sup> involving chromosome 8 facilitated the cloning of *FGFR1* (*KAL2*, MIM 136350), associated with both IHH and KS.<sup>6</sup> Positional cloning in consanguineous autosomal-recessive IHH families revealed *KISS1R* (MIM 604161) encoding GPR54,<sup>7,8</sup> *TAC3*<sup>9</sup> (MIM 162330), and *TACR3*<sup>9</sup> (MIM 162332) and candidate-gene approaches identified mutations in *GNRHR*<sup>10,11</sup> (MIM 138850), *NELF*<sup>12</sup> (MIM 608137), *CHD7*/*KAL5*<sup>13</sup> (MIM 608892), *FGF8* (*KAL6*)<sup>14</sup> (MIM 600483), and *GNRH1*<sup>15,16</sup> (MIM 152760); analogous mouse phenotypes pointed to *PROKR2* (*KAL3*)<sup>17</sup> (MIM 607123), and *PROK2* (*KAL4*)<sup>17</sup> (MIM 607002), as well as *KISS1R*.<sup>8</sup> Despite these significant advances in the past two decades, however, the genetic etiology remains unknown for about two-thirds of all IHH and KS patients.

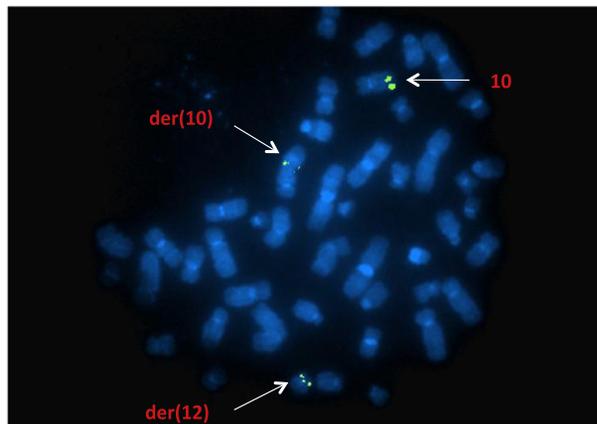
<sup>1</sup>Section of Reproductive Endocrinology, Infertility, and Genetics, Department of Obstetrics and Gynecology; Reproductive Medicine and Developmental Neurobiology Programs in the Institute of Molecular Medicine and Genetics; and Neuroscience Program, Medical College of Georgia, 1120 15<sup>th</sup> Street, Augusta, Georgia 30912, USA; <sup>2</sup>Department of Biological Science, Sungkyunkwan University, Suwon 440-746, Korea; <sup>3</sup>Institut für Humangenetik, Universitätsklinikum Hamburg-Eppendorf, Martinistrasse 52, Hamburg 20246, Germany; <sup>4</sup>Department of Human Molecular Genetics, Max Planck Institute for Molecular Genetics, Ihnestrasse 63-73, Berlin 14195, Germany; <sup>5</sup>Department of Biology and Graduate School of Analytical Science and Technology, Chungnam National University, Daejeon 305-764, Korea; <sup>6</sup>Departments of Pediatrics and Pathology, Medical College of Georgia, 1120 15<sup>th</sup> Street, Augusta, Georgia 30912, USA; <sup>7</sup>Medical College of Georgia Cancer Center, Augusta, Georgia, 30912 USA; <sup>8</sup>Department of Human Genome Research, Kazusa DNA Research Institute, 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan; <sup>9</sup>Forschungszentrum Karlsruhe, Institute for Nanotechnology, PO Box 3640, Karlsruhe, 76021 Germany; <sup>10</sup>Gulhane Military Medical Academy, Haydarpasa Training Hospital, Department of Endocrinology, Istanbul, 34660 Turkey; <sup>11</sup>Division of Medical Genetics, Departments of Pediatrics and Obstetrics and Gynecology, Medical College of Wisconsin, Milwaukee, Wisconsin 53226, USA; <sup>12</sup>Director of Andrology, Columbia Fertility Associates, Washington, DC 20037, USA; <sup>13</sup>Dr. John T. Macdonald Foundation Department of Human Genetics, Miller School of Medicine, University of Miami, Miami, Florida 33156, USA; <sup>14</sup>Division of Basic Medical Sciences, St. George's Medical School, University of London, London SW17 0RE United Kingdom; <sup>15</sup>Center for Human Genetic Research, Massachusetts General Hospital and Department of Genetics, Harvard Medical School, Boston, Massachusetts 02114, USA; <sup>16</sup>Jena University Hospital, Institute of Human Genetics, Jena 07743 Germany; <sup>17</sup>Division of Pediatric Genetics, Ankara University School of Medicine, Ankara 06610, Turkey

\*Correspondence: hkim@chgr.mgh.harvard.edu (H.-G.K.), llayman@mcg.edu (L.C.L.)

DOI 10.1016/j.ajhg.2010.08.018. ©2010 by The American Society of Human Genetics. All rights reserved.



**B**



To identify chromosomal rearrangements in IHH and KS as a basis for disease gene identification, we previously karyotyped 76 patients and found one male KS patient with chromosome translocation reported as: 46,XY, t(10;12)(q26.3;q13.1)dn (Figure 1A).<sup>18</sup> The chromosome 10q26 region has been associated previously with abnormal male genital development resulting from interstitial or terminal deletions as well as a balanced translocation. Abnormal development has included cryptorchidism,<sup>19,20</sup> small testes,<sup>21</sup> sperm defects and infertility,<sup>22</sup> micropenis and hypogonadism,<sup>23</sup> hypogonitalism,<sup>24</sup> and sparse sexual hair.<sup>25</sup> All of these phenotypic features overlap with characteristics of IHH and KS. Importantly, a trisomy involving 10q26 has been reported in a KS patient with an unbalanced chromosome translocation,<sup>26</sup> suggesting the presence of a new KS gene, which can involve dysregulation as a result of either reduced or increased dosage, in this critical region. Therefore, we postulated that a single disease gene disrupted or dysregulated by a position effect<sup>27</sup> is located at or near the 10q26 breakpoint in our KS patient with a balanced translocation.

## Material and Methods

### Patients

IHH was diagnosed in males who were 18 years old or younger and had delayed puberty, testosterone levels <100 ng/dl (normal is

## Figure 1. Balanced Chromosome Translocation and FISH Mapping of the Chromosome 10 Breakpoint

(A) Ideogram and composite chromosomes illustrating the balanced t(10;12)(q26.12;q13.11)dn (revised in this paper from the original karyotypic assessment, t(10;12)(q26.3;q13.1)dn) in the KS patient.

(B) FISH mapping with BAC clone RP11-254K03, labeled with SpectrumGreen to metaphase spreads of the KS patient, resulted in hybridization to the normal chromosome 10, as well as the der(10) and der(12) chromosomes, indicating that the translocation breakpoint of chromosome 10 is located within the sequence of this BAC clone.

300–1100 ng/dl), and low or normal serum gonadotropins. In females, IHH was defined as primary amenorrhea, nearly always with absent breast development at  $\geq 17$  years of age and low estradiol (<30 pg/ml).<sup>18,28</sup> All patients had normal pituitary function, including normal thyroid-stimulating hormone, thyroxine, cortisol, and prolactin. No pituitary tumor was present by radiographic imaging. Complete IHH or KS is a more severe phenotype defined as the complete lack of puberty with absent breast development (Tanner 1) in females and testis size  $\leq 3$  ml bilaterally in males. Incomplete IHH or KS was defined as partial breast development in females and testis size  $\geq 4$  ml bilaterally in males.<sup>18</sup>

Olfaction was either tested with the University of Pennsylvania Smell Identification Test when available or documented by history. Lymphoblastoid cell lines were generated from patients, and DNA, RNA, and/or protein was extracted by standard methods as described previously.<sup>18</sup> All patients signed an informed consent approved by the Human Assurance Committee of the Medical College of Georgia.

### Tissue Culture and Lymphoblastoid Cell Lines

Epstein-Barr-virus-immortalized lymphoblastoid cell lines used in this study were established from peripheral blood lymphocytes of individuals. Once established, these cell lines were cultured in RPMI 1640 (Mediatech) supplemented with 10% FCS, 2 mM L-glutamine, and 0.017 mg/ml benzylpenicillin (CLS) and grown at 37°C with 5% CO<sub>2</sub>.

### SNP Oligonucleotide Microarray Analysis

The Affymetrix Human Mapping 500K Array Set (Affymetrix, Santa Clara, CA) is comprised of two arrays, each capable of genotyping on average 250,000 SNPs (approximately 262,000 for the Nsp arrays and 238,000 for the Sty arrays). The mapping 500K array set has a mean spacing of 5.8 kb. Specimens were assayed with both arrays. The 500K assay was performed according to the manufacturer's protocol, beginning with 250 ng DNA. Ninety  $\mu$ g of PCR product were fragmented and labeled. Hybridization was carried out in the Affymetrix GeneChip hybridization oven 640. Posthybridization washing and processing were performed with the Affymetrix GeneChip Fluidics Station 450. The arrays were then scanned with the Affymetrix GeneChip Scanner 3000 7G. Image processing was performed with GCOS 1.4, and genotypes were called with GTYPE 4.1 software and default settings.

Detection of copy-number changes was performed in the Chromosome Copy Number Analysis Tool (CNAT) version 4.0, with a reference set of 42 normal males. Data from the separate Mapping 250K arrays were virtually combined after normalization but before smoothing in CNAT 4.0.

### Fluorescence In Situ Hybridization Analysis

The clones were labeled with either Spectrum Orange or Spectrum Green direct-labeled dUTP via the nick-translation Labeling Kit (Abbott Molecular). The procedure involved incubation of the DNA extracted from the clone with dATP, dCTP, dGTP, and fluorescent dUTP in the presence of DNA polymerase I and DNase I at 15°C for 8–16 hr. Heating the mixture in a 70°C water bath for 10 min stopped the reaction.

The slides from the harvested lymphoblastoid cell suspension were prepared according to standard procedures. The cells on the slides were denatured in 70% formamide and 2× SSC at 70°C for 3 min and dehydrated serially in cold ethanol at increasing concentrations up to absolute ethanol. The fluorescently labeled probe mix was then applied to the slides, which were covered with a coverslip and hybridized overnight in a moist chamber at 37°C. After overnight hybridization, the slides were washed in a solution of 4× SSC and 0.3% NP 40 at 72°C for 2 min and then washed with 2× SSC and 0.1% NP40 at room temperature for 2 min. The metaphases were counterstained with 0.4 µg/ml DAPI (4,6-diamidino-2-phenylindole) and analyzed under a fluorescence microscope. The images were photographed with a CCD camera with an Applied Imaging system.

### Mouse In Situ Hybridization Analysis

Two primer pairs were used for amplification of two independent *Wdr11* probes spanning nucleotides 641–1035 and 2892–3492 of the murine *Wdr11* transcript (NM\_172255.3). The primers used are listed in Table S4. Amplicons were ligated to a Topo-TA vector (Invitrogen) and subcloned into pBluescript via EcoRI and XbaI (restriction sites are underlined in Table S4). The probes were labeled with [ $\alpha$ -<sup>35</sup>S]UTP for hybridization on sections or with Digoxigenin<sup>29</sup> for whole-mount in situ hybridization as well as for cryosections. Murine embryos were frozen on solid CO<sub>2</sub>, and 10 µm sections were prepared on a cryostat. Adult mouse brains were sectioned at 15 µm. Sections hybridized with the [ $\alpha$ -<sup>35</sup>S]UTP probe were exposed to Kodak Biomax MR film for 3 days. No specific signals were detected with the respective sense probes. Whole-mount in situ hybridization with Digoxigenin-UTP-labeled probes was performed on total mouse embryos with the in situ probe spanning nucleotides 2892–3492. The day of plug was not counted for specification of embryonic stages.

### Isolation of Zebrafish *wdr11* Gene and Whole-Mount In Situ Hybridization

The zebrafish *wdr11* gene (XM\_682139.3) was isolated from the 24 hpf zebrafish cDNA library by RT-PCR with a BamHI-linked forward primer and an NcoI-linked reverse primer (Table S4; restriction sites are underlined), subcloned into the pGEM T-easy vector, and then sequenced by automatic sequencer. To examine the spatiotemporal expression patterns of *wdr11*, whole-mount in situ hybridization was performed as previously described.<sup>30</sup> Antisense digoxigenin-labeled RNA probes for *wdr11* and *emx1*<sup>31</sup> were produced with a DIG-RNA labeling kit (Roche, Germany) according to the manufacturer's instructions.

### Mutation Analysis

Genomic DNA extracted from 201 unrelated IHH and/or KS patients was amplified with *WDR11* primer pairs flanking the 29 exons and splice junction sites by nested PCR. Oligo 6.0 or Primer 3 was used for primer design, and conditions were optimized for each PCR as described previously.<sup>11,13</sup> Nested-PCR products were electrophoresed on agarose gels, precipitated via ethanol, quantified, and sequenced with the Big Dye Terminator Kit (Applied Biosystems; Foster City, CA).<sup>32</sup> Sequencing reactions were purified with Centrosep columns, lyophilized, and placed on an ABI 3730 automated DNA sequencer (Applied Biosystems; Foster City, CA). All exons were sequenced at least twice; mutations were confirmed 3–4 times by independent PCR. Any putative mutation was sequenced in >420 healthy white controls and >400 healthy Turkish controls. The CodonCode Aligner Program (CodonCode Corporation; Dedham, MA) was used for analysis of the sequencing data.

### Semiquantitative RT-PCR in Rat Tissue

Total RNA was extracted from lymphoblastoid cell lines and tissues with TRI Reagent (Molecular Research Center) and treated with DNase I (QIAGEN). We primed 100 ng of RNA with gene-specific primers by using Superscript III one-step RT-PCR kit (Invitrogen) and Mastercycler Gradient (Eppendorf). Typical conditions were as follows: 60°C for 30 min followed by 94°C for 2 min (one cycle), then 94°C for 15 s followed by 60°C for 30 s and 68°C for 50 s (30 cycles). Then 68°C for 5 min followed by cooling to 4°C. For semiquantitative RT-PCR analysis in rat tissues (Figure S1), we used two primers to amplify nucleotides 2651–3357 of the rat *WDR11* transcript (XM\_219377.5) (Table S4).

### Yeast Two-hybrid Screening

For bait construction, a DNA fragment encoding amino acids 1–830 of WDR11 was subcloned into the BamHI and Sall sites of pGBKT7 (Clontech). Approximately 5 × 10<sup>6</sup> transformants from a mouse-brain cDNA library (Clontech) were screened in the AH109 yeast strain. The positive colonies were confirmed with  $\beta$ -galactosidase colony lift assays. The yeast plasmids were isolated and transformed into *Escherichia coli* DH5 $\alpha$  cells. We cotransformed the isolated yeast DNA with a bait plasmid into AH109 to verify the interactions between WDR11 and the interacting proteins.

### Preparation of *WDR11* Mutant Constructs

A pFN21AA1351 clone that included *WDR11* open-reading frame (ORF) sequence with the GenBank accession number of AB385454 was used for construction of the *WDR11* mutant. DNA fragments containing a nucleotide substitution were amplified by two-step PCRs in which pFN21AA1351 served as a template; two separated DNA fragments were amplified by the first PCR with one of the mutant primers containing a nucleotide substitution (mR1 or mF1 series) and the corresponding opposite outside primer (F01, F06, R01 or R02), then two resultant amplified PCR products were combined by the second PCR with the outside primers and cloned into pUC118 vector. The cloned mutant DNA fragment was cut with EcoNI for m1, m2, m3, and m4 constructs, and the mutant EcoNI fragment was inserted between EcoNI sites of the *WDR11* ORF. For m5 and m6, the cloned DNA fragment was cut with SwaI and PvuII, and the 1422 bp SwaI-PvuII fragment of *WDR11* was replaced with the mutant fragment. The sequences of primers used for the construction are listed in Table S4.

## Plasmid Construction

The full-length human *WDR11* and *EMX1* clones were purchased from Openbiosystems (clone ID 34306203 and 5260039, respectively). The full ORFs of *WDR11* and *EMX1* were amplified from purchased cDNA clones by PCR with specific primers, and the DNA fragment was inserted into BamHI/NotI and EcoRI/XhoI sites of pEntr3C, respectively. The Myc-*WDR11*, HA-*EMX1*, and GST-*EMX1* plasmids were constructed by Gateway Technology (Invitrogen) with pEntr-*WDR11* and pEntr-*EMX1*.

## Coimmunoprecipitation and Immunoblotting

Coimmunoprecipitation was performed after the lysis of  $2 \times 10^7$  cells with lysis buffer (50 mM HEPES, 150 mM NaCl, 10% glycerol, 1% Nonidet P-40, and 1 mM EDTA). After incubation on ice for 10 min and centrifugation for 10 min at 4°C, equal volumes of protein were incubated overnight with antibody and protein A G-Sepharose beads at 4°C on a rotating wheel. The beads were washed three times with lysis buffer. The whole-cell lysates and immunoprecipitates were separated by SDS-PAGE and then transferred onto polyvinylidene difluoride membranes. The membranes were immunoblotted with anti-Myc (Invitrogen) or anti-HA (Invitrogen) antibodies.

## In Vitro Pull-Down Assays

The GST-fusion plasmids were expressed in *E. coli* BL21(DE3) cells and were purified with glutathione-Sepharose beads according to the instructions of the manufacturer. The Myc-EMX1 proteins were synthesized via the coupled TNT in vitro translation system (Promega, Madison, WI). We performed pull-down assays by incubating equal amounts of GST or GST-WDR11 fusion proteins immobilized onto glutathione-Sepharose beads with in vitro translated Myc-Emx1 and washing the beads three times with phosphate-buffered saline and 0.5% Triton X-100. After washing, the bound proteins were resolved by SDS-PAGE and detected via immunoblotting with anti-Myc antibody.

## Immunohistochemistry

U2OS cells were transfected with expression plasmids encoding HA-EMX1 and GFP-WDR11. Twenty-four hours after transfection, cells were treated with leptomycin B (10 ng/ml) for 12 hr before being fixed with 4% paraformaldehyde for 5 min and incubated with anti-HA antibody. Fluorescence microscopy was conducted with a Zeiss Axioskop 2 microscope; excitation wavelengths of 543 nm (rhodamine red) and 488 nm (GFP) were used. The acquired images were processed with Adobe Photoshop. For subcellular localization for FCNB4-hTERT cells, approximately 2 ml of cell suspension was plated into each well of a microscopy cover slide, incubated for 24 hr in a 37°C CO<sub>2</sub> incubator, and then fixed with 4% paraformaldehyde (PFA) for 5 min at room temperature. Cells were washed three times with PBS and permeabilized with 0.1% Triton X-100 in PBS for 5 min. After two washes with PBS, nonspecific binding sites were blocked with 1% BSA (Sigma) in PBS for 5 min. Samples were then rinsed with PBS and incubated overnight at 4°C with primary antibody. After being washed with PBS, samples were incubated for 1 hr at room temperature with the appropriate Alexa Fluor dye (Molecular Probes)-coupled anti-rabbit 488 (green) or anti-mouse 589 (red) secondary antibodies. After three PBS washes, nuclei were counterstained with To-pro3 (1 nM). Cover slips were mounted with antifade vector shield (Molecular Probe) reagent and viewed with the Zeiss LSM 510 Meta Confocal Microscope.

## Exon Trapping

We used the exon-trapping system (Invitrogen GIBCO) to investigate the effect of the heterozygous c.2932A>C nucleotide variant (p.K978Q) in the acceptor splice site of exon 24 of *WDR11*. A genomic fragment of 2411 bp encompassing exons 23–25 along with flanking intronic sequences and a 1286 bp fragment comprising exon 25 along with flanking intronic sequences were amplified by PCR with two primer pairs (Table S4). The resulting PCR products were cloned into exon-trapping vector pSPL3 via restriction sites EcoRI (5' end) and BamHI (3' end). We verified sequences for all constructs to exclude PCR-induced mutations and to select constructs with or without the splice site variant c.2932A>C (p.K978Q). COS-7 cells were grown in 100 mm dishes in DMEM (Invitrogen) supplemented with 10% fetal calf serum (Invitrogen) in a 10% CO<sub>2</sub> atmosphere at 37°C. At approximately 60% confluency, cells were each transfected with 8 µg of exon-trapping construct DNA via Lipofectamine 2000 (Invitrogen). Cells were collected 30 hr after transfection, and cytoplasmic RNA was isolated with RNeasy (QIAGEN) according to the manufacturer's instructions. After transcription into cDNA via Omniscript Reverse Transcriptase (QIAGEN), two rounds of PCR with vector-specific oligonucleotides were performed. PCR products were subcloned into TA cloning vector pCR2.1-Topo (Invitrogen) and directly sequenced with the ABI BigDye Terminator Sequencing Kit (Applied Biosystems) and the automated capillary sequencer ABI 3130 (Applied Biosystems).

## Anti-WDR11 Antibody Generation

A new WDR11 polyclonal antibody was generated against the N-terminal residues 32–48 of NP\_060587.8 (QGLIAYGCHSLVV VIDS); an N-terminal cysteine was added for conjugation purposes. Rabbits were injected with Freund's adjuvant containing the peptide sequence. After the second bleed, affinity purification was performed. The peptide sequence used for raising the antibody was conserved in ten species (human, chimpanzee, cow, horse, panda, pig, dog, rat, mouse, and rabbit).

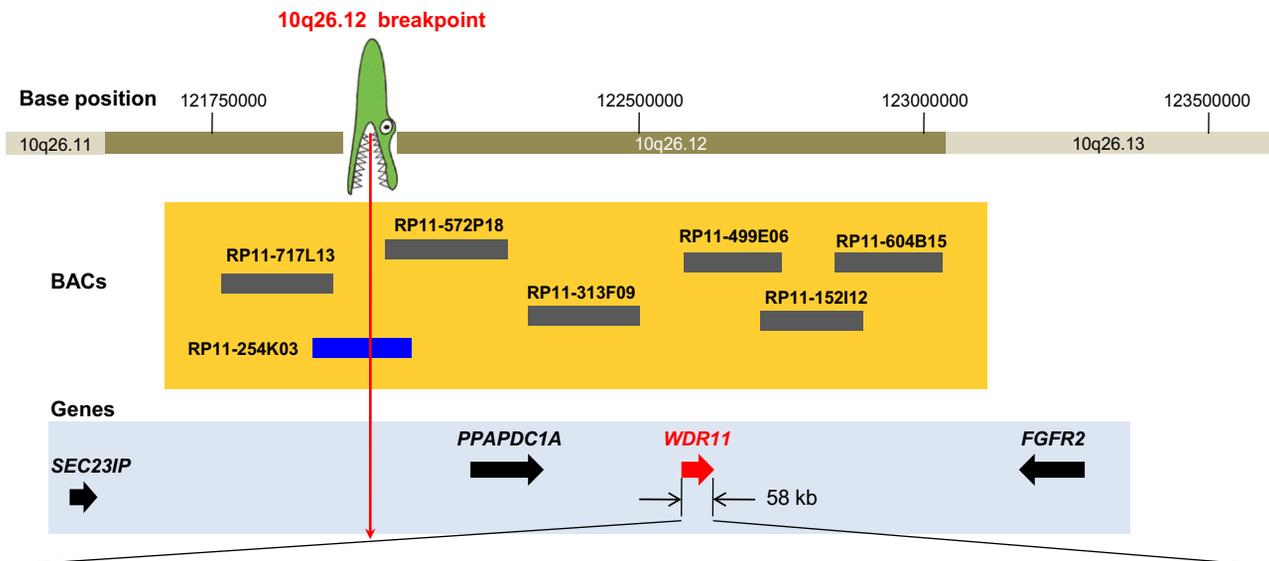
## Array Painting

Flow-sorting of metaphase chromosomes was performed as described previously.<sup>33</sup> Both derivative chromosomes were amplified with the GenomiPhi V2 DNA Amplification kit (GE Healthcare, Piscataway, NJ, USA) according to the manufacturer's protocol. One microgram of each amplification product was labeled with Cy3 or Cy5 via Agilent's Genomic DNA Enzymatic Labeling Kit Plus (Agilent). To each labeling reaction, adding 100 ng of genomic control DNA ensured proper placement of the grid after image analysis. A customized oligonucleotide array that represented chromosomal regions chr10:121,855,186–122,255,228 and chr12:45,828,906–46,226,893 (hg18) was designed by means of eArray (Agilent, Santa Clara, CA); there was an average oligonucleotide spacing of 200–300 bases. Hybridization and washing steps were performed as previously described.<sup>34</sup>

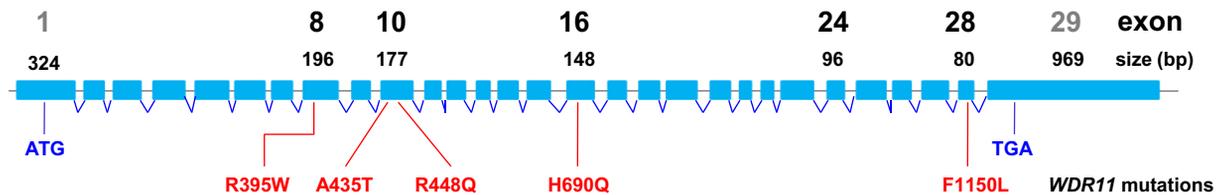
## Quantitative Real-Time RT-PCR (RT-qPCR) of *WDR11* in the t(10;12) Patient

cDNA from lymphoblastoid cells was used for real-time RT-qPCR with the StepOnePlus Real-Time PCR system (Applied Biosystems) and Power SYBR Green PCR Master Mix (Applied Biosystems) and two primer pairs (WDR11-1:3117forward+3213reverse; WDR11-2:3221forward+3370reverse; Table S4). RT-qPCR was performed in triplicate with same amount of cDNA, and *WDR11* levels were

## A Translocation breakpoint in 10q26.12



## B Exon structure of *WDR11*



**Figure 2. Genetic Mapping of the 10q26 Locus Involved in IHH and KS**

(A) The translocation breakpoint, depicted as a crocodile head, is located between *SEC23IP* and *PPAPDC1A*. Horizontal bars at 10q26.12 show BACs used for FISH mapping. The size and location of BACs are to scale. The blue bar depicts a BAC clone spanning the breakpoint. Four positional candidate genes located proximal and distal to the breakpoint are shown as arrows in a 2 Mb region.

(B) Exon and intron structure of the 58 kb gene *WDR11* (NM\_018117.11). Locations of human missense mutations are identified in sporadic IHH and KS patients. Notable exons are shown to scale as blue rectangles and are numbered with exon size. The sizes of introns are not to scale. Mutation F1150L was identified in two independent sporadic IHH patients.

normalized to *GAPDH* expression. *WDR11* levels in t(10;12) were normalized to those in a male control with StepOne Software (Applied Biosystems).

### Immunoblot Analysis of *WDR11* in the t(10;12) Patient

Cells were washed with PBS twice and lysed in RIPA buffer (PIERCE) containing complete EDTA-free protease inhibitor (Roche Diagnostics). Total cellular lysates were separated by SDS-polyacrylamide gel electrophoresis and transferred to nitrocellulose membranes. Membranes were incubated with primary antibody overnight at 4°C. Membranes were washed with TBS-T buffer three times and incubated with a secondary antibody conjugated to horseradish peroxidase for 2 hr at room temperature. After three washes with TBS-T buffer, the blots were visualized by an enhanced chemiluminescence method (GE Healthcare).

### Results

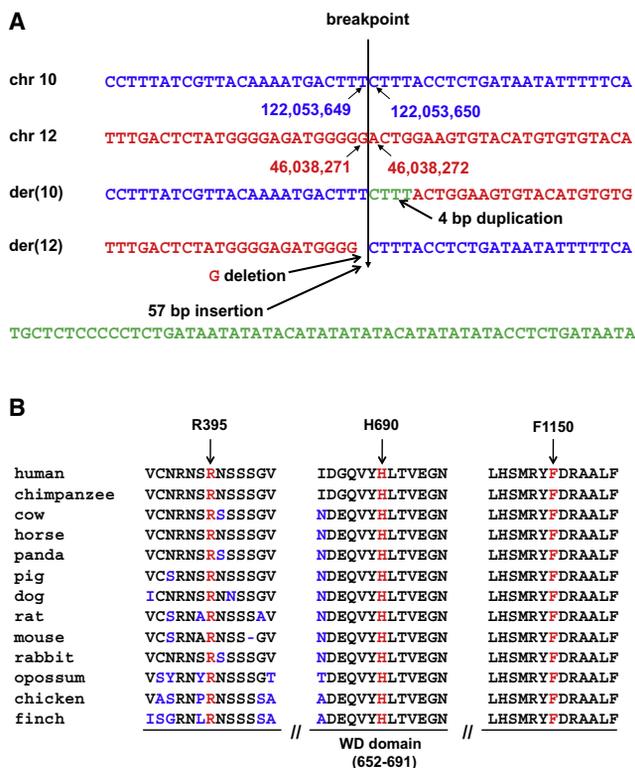
#### Delineation of the Breakpoint Region on 10q26

Affymetrix Human Mapping 500K Array SNP oligonucleotide microarray analysis (SOMA) of patient DNA was

consistent with a balanced translocation and excluded copy-number variation (CNV) as the cause of the IHH or KS phenotype (data not shown). To identify the potentially disrupted genes, we first mapped the translocation breakpoints by using fluorescence in situ hybridization (FISH). Two BAC clones, RP11-592D19 in 10q25.3 and RP11-91E2 in 10q26.3, mapped as centromeric and telomeric, respectively, to the chromosome 10 breakpoint and thus flanked it within a 16.4 Mb region. Sequential rounds of FISH with BACs within this region narrowed the candidate region until a breakpoint-crossing BAC clone was identified. Of nineteen BACs examined, three were proximal and fifteen were distal to the breakpoint, whereas RP11-254K03 hybridized to the normal chromosome 10 and both der(10) and der(12) chromosomes, indicating that it spans the translocation breakpoint, as shown in Figures 1B and 2A.

#### Delineation of the Breakpoint Region on 12q13

BAC clones RP11-88L2 from 12q12 and RP11-762I7 from 12q13.2, used as starting clones for FISH, mapped proximal



**Figure 3. Sequences of the Translocation Breakpoints and Protein Sequence Alignment of WDR11 Orthologs**

(A) Genomic DNA sequence at the breakpoints from the normal and derivative chromosomes. The breakpoint on chromosome 10 is located between nucleotides 122,053,649 and 122,053,650, whereas on chromosome 12 it occurs between nucleotides 46,038,271 and 46,038,272 (UCSC Genome Browser NCBI hg18). The junction sequence reveals a 4 bp duplication at the breakpoint of the der(10) chromosome and a 1 bp deletion and a 57 bp insertion at the junction on the der(12) chromosome.

(B) ClustalW multiple alignment of partial protein sequences of WDR11 orthologs. The positions of three residues affected by missense mutations of WDR11 are marked by arrows and red letters in the corresponding segments of the multiple alignment. The amino acid residues that differ from the sequence of the human WDR11 protein are indicated in blue, and the ninth WD domain is indicated under the panel. All three mutated residues are evolutionarily fully conserved in all 13 available WDR11 orthologs.

and distal, respectively, to the breakpoint, confining the chromosome 12 breakpoint within an 11.7 Mb region. FISH experiments with twenty BACs within the region revealed twelve proximal and seven distal to the breakpoint, as well as RP11-464D5, which hybridized to the normal chromosome 12 and both der(10) and der(12) chromosomes, indicating that it spans the translocation breakpoint (data not shown). Sequencing of the breakpoint cloned as described below confirmed localization of the junction sequence within RP11-464D5, as shown in Figure 3A.

### Refining the Breakpoint Regions by Array Painting and Cloning of Junction Fragments

After narrowing down both breakpoint regions by FISH and before finding the breakpoint spanning clones, we performed array painting by using a customized oligonucleo-

tide array that represented chromosomal regions chr10: 121,855,186–122,255,228 and chr12:45,828,906–46,226,893 (hg18) with an average oligonucleotide spacing of 200–300 bases to determine the breakpoint regions more precisely and eventually clone the junction fragments. Based upon the array-painting results, the breakpoint regions were refined to 2.7 kb (between nucleotide positions 122,053,148 and 122,055,901) in 10q26.12 and 4.2 kb (between nucleotide positions 46,034,656 and 46,038,863 in 12q13.11) (UCSC Genome Browser, NCBI Build 36/hg18 assembly), resulting in restatement of the previously cytogenetically assigned karyotype to 46,XY,t(10;12)(q26.12;q13.11)dn. Five forward primers (12q13.2-10kb-4208f, -4562f, -5338f, -5961f, and -7351f) designed from 10 kb containing the 4.2 kb breakpoint region at 12q13.11 and three reverse primers (10q26-10kb-3497r, -5710r, and -5750r) designed from 10 kb encompassing the 2.7 kb breakpoint region at 10q26.12 were paired in a long-range PCR (QIAGEN LongRange PCR Kit 250). Annealing was performed at 60°C for 30 s with an extension for 7 min. All fifteen PCR reactions produced amplicons with sizes from 1.7–7.1 kb (consistent with the interprimer distances on each chromosome) from the subject's DNA but not from that of a normal control male. Therefore, the breakpoints at 10q26.12 and 12q13.11 were narrowed to 590 bp and 2.2 kb, respectively, confirming the array-painting result. Finally, the junction fragment of ~1.8 kb from der(12) was amplified by normal PCR from one primer pair of 12q13.2-10kb-7351forward and 10q26-10kb-3497reverse and confirmed by sequencing and subsequent BLAST searching.

To amplify the junction fragment of der(10), we paired five forward primers (10q26-10kb -1851f, -2367f, -2801f, -2851f, and -2914f) designed from the 10 kb breakpoint region at 10q26 and five reverse primers (12q13.2-10kb-7630r, -8180r, -8836r, -9664r, and -9890r) from the 10 kb breakpoint region at 12q13.2 in a long-range PCR, which produced amplicons of 1–2.5 kb in combinations with two of the reverse primers (12q13.2-10kb-9664r and -9890r) but no amplicons with the other three (12q13.2-10kb-7630r, -8180r, and -8836r), further refining the breakpoint region. The 2 kb junction fragment of der(10) was amplified by primer set 10q26-10kb-2367f and 12q13.2-10kb-9890r. The partial sequences from chromosomes 10 and 12 of this junction fragment were confirmed by sequencing and BLAST searching. The breakpoint on chromosome 12 is located between nucleotides 46,038,271 and 46,038,272, whereas on chromosome 10 it occurs between nucleotides 122,053,649 and 122,053,650, 547 kb from the 5' end of *WDR11*. At the junction of the der(10) fragment, there was a CTTT duplication, and at the junction of der(12) fragment, there was a 1 bp G deletion and a 57 bp insertion (Figure 3A).

### Candidate Gene Expression Pattern in Rat

The translocation did not directly disrupt a gene on either chromosome 10 or 12, but on the basis of the common

chromosomal location associated with IHH and KS phenotypes, we hypothesized that 10q26 was more likely to harbor the causative gene; therefore, we did not study the 12q region further.

On 10q26.12, four candidate genes (*FGFR2*, *PPAPDC1A*, *SEC23IP*, and *WDR11*) were present in the vicinity of the breakpoint (Figure 2A). We examined their expression in the rat ovary, testis, olfactory bulb, hypothalamic preoptic area, medial basal hypothalamus, anterior pituitary, piriform cortex, temporal cortex, prefrontal cortex, hippocampus, and cerebellum using semiquantitative RT-PCR analysis. All four genes were expressed in all tissues tested; the highest *Wdr11* (XM\_219377.5) expression, in decreasing order, was in the ovary, olfactory bulb, and piriform cortex (Figure S1).

### Mutation Screening of Positional Candidate Genes

We screened 123 IHH and KS patients for *FGFR2* (MIM 176943) mutations by direct sequencing of all coding exons and splice junctions (exons 2–18 of NM\_022970.3 and the additional exon 8 of NM\_000141.4) and found only known polymorphisms rs755793, rs1047100, rs3135774, rs41293763, and rs2278202. Exons 1–7 of *PPAPDC1A* (NM\_001030059.1) were directly sequenced in 60 IHH or KS patients and were likewise negative for mutations other than reported polymorphisms rs10886691 and rs1047369 and additional synonymous variants F15F, L194L, and P270P. Exons 1–18 of *SEC23IP* (NM\_007190.2) were screened in 120 IHH or KS patients. Only known polymorphisms rs34157476, rs3859163, rs3740569, rs3740570, rs17099368, rs34824340, rs2475298, rs2271123, rs12771873, rs58111481, and rs34826964 were seen as variants. Additionally, heterozygous c.760G>T of *SEC23IP* (p. V254F) was identified in one KS female and two IHH males but not in a second KS female within a family, indicating a lack of segregation with IHH or KS and suggesting that it is a rare sequence variant (absent from 182 normal white controls). Furthermore, heterozygous c.1191+38A>C identified in a male IHH patient was inherited from one of two healthy parents, both of whom are heterozygous for the same change, suggesting that it is also a polymorphism.

DNA sequencing of the protein-coding exons and splice junctions of the *WDR11* gene with 29 exons (NM\_018117.11, MIM 606417) was performed in 201 IHH or KS patients. Five different heterozygous missense variants that appear to be pathogenic mutations were identified in six unrelated probands (3%) (Table 1); none of these variants was present in  $\geq 420$  normal white controls or an additional 402 ethnically matched controls for the one proband who is of Turkish origin. Interestingly, the putative *WDR11* mutations were found in both IHH ( $n = 5$ ) and KS ( $n = 1$ ) patients, as shown in Table 1, and are described below. We identified one additional heterozygous change, c.2932A>C (p.K978Q), in a Turkish male IHH patient. Residue K978 is completely conserved in all 13 higher vertebrates (Figure S2) and this same nucleotide change was not identified in an initial set of 180 white and Turkish

controls. Although the K978Q alteration involves a base change in the first nucleotide of exon 24, exon-trapping experiments did not reveal any aberrant splicing in vitro (data not shown), suggesting that the variant does not affect splicing differences in vivo. Sequencing of additional Turkish control subjects identified the c.2932A>C variant in two Turkish controls (2/402). Thus, c.2932A>C (p.K978Q) is a sequence variant that is not sufficient to cause IHH or KS and has the characteristics of a rare polymorphism (Table S1). Whether it is associated with increased risk of IHH or KS will require investigation of additional IHH, KS, and ethnically matched control subjects.

The five putative heterozygous missense variants that were observed in IHH and KS patients but absent from all controls are c.1183C>T (p.R395W) in an IHH male, c.1303G>A (p.A435T) in a Turkish IHH male, c.1343G>A (p.R448Q) in an IHH female, c.2070T>A (p.H690Q) in a KS male, and c.3450T>G (p.F1150L) in one male and one female IHH patient. Three of the missense changes altering amino acid residues R395, H690, and F1150 are completely conserved in all higher vertebrates (Figure 3B). Two unrelated patients had an identical heterozygous missense alteration, F1150L, so we performed haplotype analysis with seven informative SNP markers—rs7077126, rs10886800, rs41287988, rs1045154, rs12259815, rs10871, and rs1045170—from the *WDR11* 3'-UTR region. Genotyping of these two patients (C17 and C100) indicated that they share a haplotype around *WDR11* and are therefore likely to be descended from a recent common ancestor (data not shown). All identified SNPs, including ten novel SNPs in *WDR11*, are listed in Table S1.

### RT-qPCR and Immunoblot Analysis of *WDR11* in the t(10;12) Patient

RT-qPCR and immunoblot analysis with antibody that we generated demonstrated reductions of approximately 20% for transcript and 10% for *WDR11* in the lymphoblast cell line from our translocation patient compared to two gender-matched controls (Figure S3). Because this line has one normal allele and one translocated allele, if the entire effect were due to reduction of the mRNA from the translocated chromosome, as would be suspected, the 20% overall reduction in transcript would correspond to an approximate 40% reduction in expression from the translocated chromosome. Although this reduction is modest, it does demonstrate proof-of-principle for a potential position effect of the breakpoint, which is about 547 kb 5' of *WDR11*. It is conceivable that, because of tissue-specific factors, the magnitude of the reduction might be significantly greater in some other tissues of the translocation subject, especially those tissues involved in the neuroendocrine control of reproduction, but it is not possible to measure this directly because the lymphoblasts represent the only tissue available from the subject. However, the small effect in lymphoblasts is consistent with an effect of the translocation on expression of this gene and

**Table 1. WDR11 Mutations in IHH and KS Patients**

Patient	Gender and Phenotype	Geographic Origin	Location (Exon/Intron)	Nucleotide change (NM_018117.11)	Amino Acid Change (NP_060587.8)	Class of Mutation	Confirmatory Method
C37	Male/KS; complete IHH; anosmia; testes, 1 ml bilaterally	United States	Translocation breakpoint is 547 kb away from the 5' end of <i>WDR11</i>	–	–	chromosomal structure mutation as a balanced translocation-t(10;12)(q26.12;q13.11)dn	FISH, array painting, and sequencing of junction fragment; de novo.
C71	Male/complete IHH; testes, 2 ml; FSH = 3 mIU/ml; LH = 1.2 mIU/ml	United States	Exon 8	c.1183C>T	R395W	missense	0/420 white controls; detrimental in protein modeling; invariant in 13 available <i>WDR11</i> orthologs.
T87	Male/IHH; complete IHH; bilat cryptorchidism; FSH = 0.3 mIU/ml; LH < 0.3 mIU/ml	Turkey	Exon 10	c.1303G>A	A435T	missense	0/426 white controls; 0/402 Turkish controls; within sixth WD domain; detrimental in protein modeling; abolished EMX1 binding.
C99	Female/IHH	United States	Exon 10	c.1343G>A	R448Q	missense	0/426 white controls; within sixth WD domain; detrimental in protein modeling; reduced EMX1 binding; predicted to destabilize the <i>WDR11</i> dimer and impair actin binding of the complex.
C127	Male/KS; anosmia	United States	Exon 16	c.2070T>A	H690Q	missense	0/420 white controls; within ninth WD domain; detrimental in protein modeling; abolished EMX1 binding; invariant in 13 available <i>WDR11</i> orthologs.
C17	Male/IHH; incomplete IHH; testes, 35, 25 ml; FSH = 6 mIU/ml	United States	Exon 28	c.3450T>G	F1150L	missense	0/420 white controls; invariant in 13 available <i>WDR11</i> orthologs.
C100	Female/IHH; complete IHH; breasts: Tanner 1; FSH < 1 mIU/ml; LH < 1 mIU/ml	United States	Exon 28	c.3450T>G	F1150L	missense	0/420 white controls; invariant in 13 available <i>WDR11</i> orthologs.

Complete IHH refers to the complete absence of sexual development, whereas incomplete IHH indicates partial pubertal development. Testis size is given (normal is 15–25 ml) when available. Breast development is given when available—Tanner 1 indicates no breast buds.

therefore with its role in generating the phenotype in this patient.

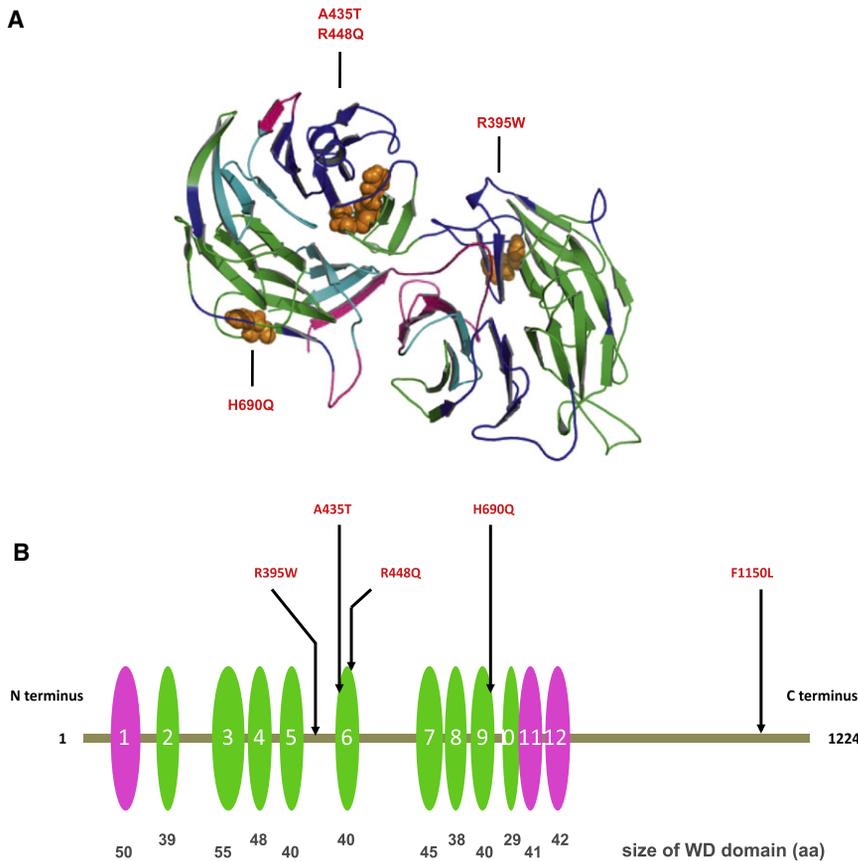
### Protein Structure Modeling of *WDR11*

To explore the potential structural features of *WDR11*, we generated a protein model by using multiple sequence alignment (ClustalW<sup>35</sup>) based on homology to the *C. elegans* Homolog of Yeast Actin Interacting Protein 1 (AIP1) (PDB code 1NR0),<sup>36</sup> shown in Figure S4. The close structural similarity of the model of *WDR11*, which features two  $\beta$  propeller structures in each protein chain<sup>36,37</sup> (Figure 4A), with the AIP1 dimer (PDB code 1PGU), indicates that *WDR11* will also be an actin-binding protein<sup>38</sup> (Figure S4). *WDR11* contains twelve WD domains, nine (second to tenth repeats) that are confirmed on the basis of direct comparison with the template structure of AIP1 (Figures 4A and 4B) and three additional repeats (first, 11<sup>th</sup>, and 12<sup>th</sup>) detected by sequence comparison outside

the region of the structural model (Figures 4A and 4B; Table S2). Like AIP1, *WDR11* is predicted to exhibit two  $\beta$  propellers; WD domains 2–6 are predicted to constitute the first, and WD domains 7–10 are predicted to constitute the second (Figure 4A). The structural model for *WDR11* overlaps well with the known AIP1 structure (Figure S4). This model predicts that *WDR11* has twelve WD domains and that nine of them (second through tenth) participate in the genesis of two consecutive  $\beta$  propellers.

### *WDR11* Interacts and Colocalizes with EMX1 In Vivo and In Vitro

*WDR11* is a 1,224 amino acid protein that contains multiple WD domains likely to mediate interaction with protein-binding partners (NCBI Conserved Domain Database cl02567). Consequently, we performed a yeast two-hybrid screen to identify potential cellular proteins whose interaction with *WDR11* might be disrupted by the



**Figure 4. WDR11 Structural Model Indicating the Mutation Sites**

(A) Model spanning amino acids 70–739 of WDR11. The model was obtained by alignment to 1NR0 via ClustalW and MOE (Molecular Operating Environment [MOE 2004.03], Chemical Computing Group, Montreal, Quebec, Canada H3B 3X3). WDR11 forms a double propeller structure, in which the WD domains indicated in (B) form the main structural constituent. The two propeller axes are tilted with respect to one another, so only the propeller structure on the left is clearly visible in this representation. Colors indicate side chains of the four mutations within the modeled sequence region, as follows: WD domains predicted on the basis of the model (green), on the basis of SMART (pink), or both (cyan). The sites of the mutations are indicated in orange.

(B) Positions of five missense mutations in WDR11; WD domains are depicted as ovals. The WD domains predicted on the basis of the model and by SMART are depicted in green and pink, respectively. The relative sizes and locations of WD domains are to scale. WDR11 contains twelve WD domains, nine (second to tenth repeats) that are confirmed on the basis of direct comparison with the template structure of AIP1 and three additional repeats (first, 11<sup>th</sup>, and 12<sup>th</sup>) detected by sequence comparison outside the region of the structural model. Note that three mutations directly affect WD domains 6 and 9.

apparent missense mutations in IHH and KS patients. We identified EMX1 (MIM 600034), a human ortholog of *Drosophila ems* (*empty spiracles*), as a WDR11-interacting protein. Although four other proteins, including Hey1 (hairy/enhancer-of-split related with YRPW motif 1, MIM 602953), Tagln2 (Transgelin 2, MIM 604634), Ndr4 (N-Myc downstream-regulated gene 4), and Nrnx3 (neurexin 3, MIM 600567) were also identified as potential WDR11 binding partners, we focused on EMX1 for further studies because of its functional relevance to the developing nervous system.<sup>39–43</sup>

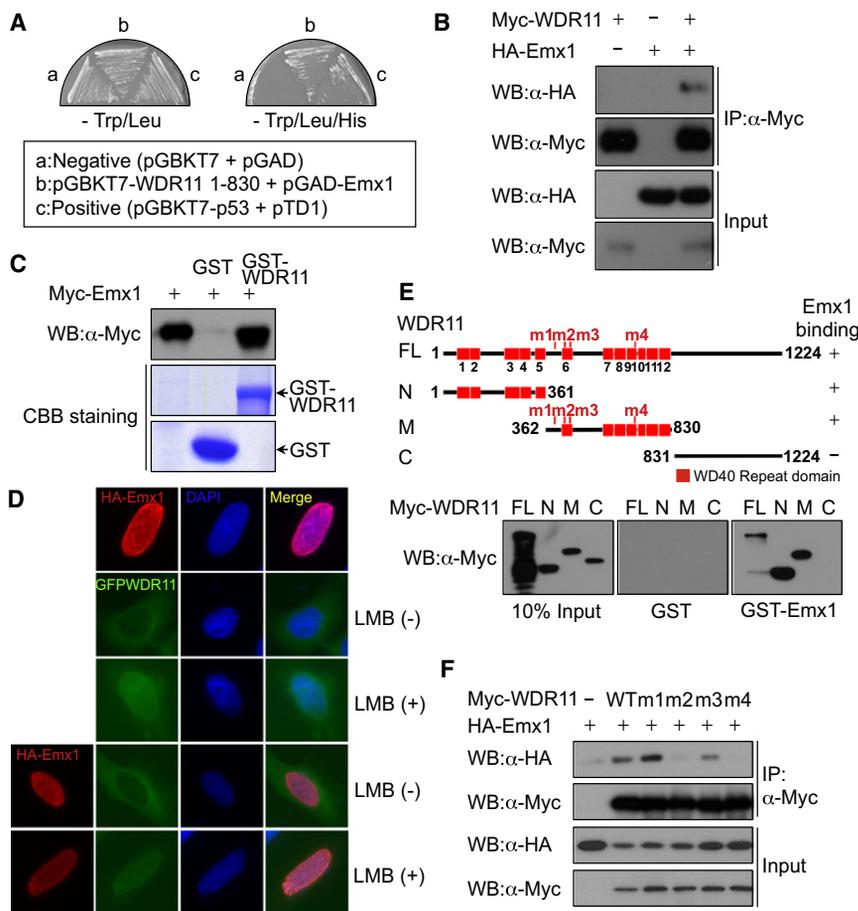
Specific interaction between WDR11 and EMX1 was revealed by cotransformation of yeast with an expression plasmid encoding WDR11 fused to the GAL4 DNA-binding domain and a plasmid encoding EMX1 fused to the GAL4 activation domain (Figure 5A). The interaction was confirmed in mammalian cells by coimmunoprecipitation assays, which demonstrated that HA-tagged EMX1 was recovered from immunoprecipitates of coexpressed Myc-WDR11 (Figure 5B), and further confirmed by GST pull-down analysis (Figure 5C).

To define the particular WDR11 domains necessary for interaction with EMX1, we generated WDR11 deletion mutants and performed GST pull-down assays to assess EMX1 binding. Both the N terminus and the central portion of WDR11 bound to EMX1, whereas the C terminus did not (Figure 5E). We next tested whether

any of the four missense mutations in the N terminus and the central region of WDR11 interfered with EMX1 binding in vitro. Coimmunoprecipitation of EMX1 with the translated proteins from each of the four WDR11 mutant constructs demonstrated that mutants 2 (A435T) and 4 (H690Q) did not associate with EMX1 in mammalian cells, whereas mutant 3 (R448Q) reduced, but did not eliminate, the EMX1 binding. Mutant 1 (R395W) had no effect upon EMX1 binding (Figure 5F).

#### Subcellular Localization of WDR11 and EMX1

Because EMX1 is a homeodomain transcription factor that participates in the development of olfactory neurons,<sup>40</sup> we examined the subcellular localization of both EMX1 and WDR11 in a physiologically relevant human cell system. Immunostaining with a polyclonal anti-WDR11 antibody (directed to N-terminal residues 32–48 of NP\_060587.8: QGLIAYGCHSLVVVIDS) localized WDR11 to the cytoplasm of FCNB4-hTERT cells, immortalized human embryonic olfactory GnRH neuroblasts isolated from olfactory epithelium of an 8- to 12-week-old human embryo (data not shown).<sup>44</sup> Similarly, fluorescence microscopy of GFP-WDR11 transfected into U2OS cells also yielded a cytoplasmic localization, whereas HA-EMX1 localized to the nucleus. However, when these cells were treated with leptomycin B, an inhibitor of nuclear export, both HA-EMX1 and GFP-WDR11 colocalized in the nucleus



**Figure 5. The WDR11 Interacts and Co-localizes with EMX1 In Vivo and In Vitro**

(A) EMX1 was identified in the yeast two-hybrid screen as a WDR11-interacting protein. The specific interactions between these two proteins were confirmed by streaking of transformed yeast cells onto synthetic drop-out plates lacking TL (Trp/Leu) or TLH (Trp/Leu/His). Yeast AH109 cells were transformed with empty vectors (pGBKT7 and pGAD), and plasmids encoding SV40 T antigen (pTD1) and p53 (pGBKT7-p53) were utilized as a negative and positive control, respectively.

(B) Myc-WDR11 expression plasmids were transfected alone or along with HA-EMX1 into HeLa cells, the cell lysates were immunoprecipitated with anti-Myc antibody, and coprecipitated HA-EMX1 was detected via immunoblotting with anti-HA antibody.

(C) In-vitro-translated Myc-EMX1 was subjected to GST pull-down analysis with GST (lane 2) or GST-WDR11 (lane 3). The bound proteins were detected via immunoblotting with anti-Myc antibody.

(D) HA-EMX1 and GFP-WDR11 expression plasmids were transfected into U2OS cells, and then cells were treated with leptomycin B (LMB), an inhibitor of nuclear export. Fluorescence microscopy analysis helped determine localizations of HA-EMX1 and GFP-WDR11. Nuclei were stained with DAPI.

(E) Wild-type Myc-WDR11 and its deletion mutants were synthesized in vitro and subjected to a GST pull-down assay with GST

(center panel) or GST-WDR11 (right panel). The N, M, and C denote the WDR11 N terminus (amino acids 1–361), middle portion (amino acids 362–830), and C terminus (amino acids 831–1224), respectively. The positions of missense mutations found within the N terminus and central region of *WDR11* in IHH patients are marked as m1 through m4 on the schematic diagrams of WDR11.

(F) The wild-type and missense mutant WDR11 expression plasmids were transfected into HEK293 cells along with HA-EMX1 expression plasmids. The cell lysates were immunoprecipitated with anti-Myc antibody, and the association of EMX1 with wild-type WDR11 or missense mutants was determined via immunoblot analysis with anti-HA antibody.

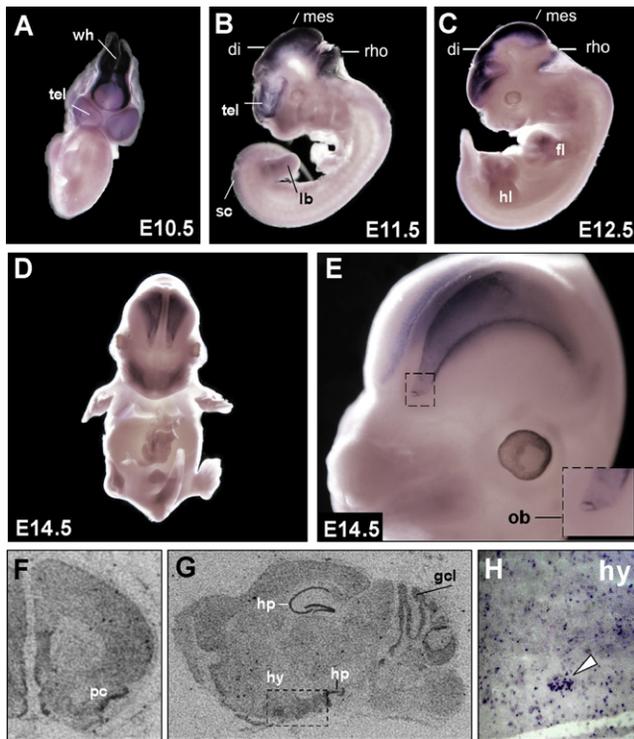
(Figure 5D), suggesting that WDR11 might be shuttling between the nucleus and cytoplasm. When each of the five WDR11 missense mutants was transfected into FCNB4-hTERT cells, none altered subcellular localization of the protein (data not shown).

### Mouse and Zebrafish Whole-Mount In Situ Hybridization

To investigate the developmental expression of *Wdr11*, we performed whole-mount in situ hybridization analysis in mouse embryos from days E10.5–E14.5. The pattern of expression observed was consistent with a role for WDR11 in IHH and KS. As early as E10.5, the entire developing central nervous system, except for the spinal cord, revealed *Wdr11* expression (Figures 6A–6C). The neuroepithelium, including the diencephalic region that gives rise to hypothalamic neurons where GnRH neurons reside, stained strongly for *Wdr11* at E11.5 and E12.5. Mouse neuroendocrine GnRH neurons migrate from the olfactory placode region alongside olfactory neurons to cross the cribriform plate and finally reach the hypothalamus,

a process that is active from E10.5–E14.5 and is usually completed by E18.5.<sup>45</sup> At E14.5 high levels of *Wdr11* expression were particularly noteworthy in the developing cortex and the olfactory bulb (Figures 6D and 6E). In the adult brain, intense *Wdr11* expression was restricted to the olfactory bulb, the olfaction-related piriform cortex (Figure 6F), the granule cell layer of the cerebellum, and neurons of the hippocampal formation (Figure 6G). Increased signal intensities within the hypothalamus were observed with <sup>35</sup>S-UTP labeled *Wdr11* antisense probes (Figure 6G). For higher-resolution images of this region, digoxigenin labeling of sagittal sections of adult brains demonstrated signals scattered throughout the hypothalamus, sometimes in clusters of neurons (Figure 6H). Notably, in mouse, *Emx1* expression is also detectable from day 9.5 of gestation in the presumptive cerebral cortex and olfactory bulbs.<sup>46</sup>

In the zebrafish, *wdr11* was expressed ubiquitously at 24 hpf; whereas *emx1* was highly expressed in the fore-brain, and a small group of cells was clustered around the same area of the diencephalic GnRH3 neurons at



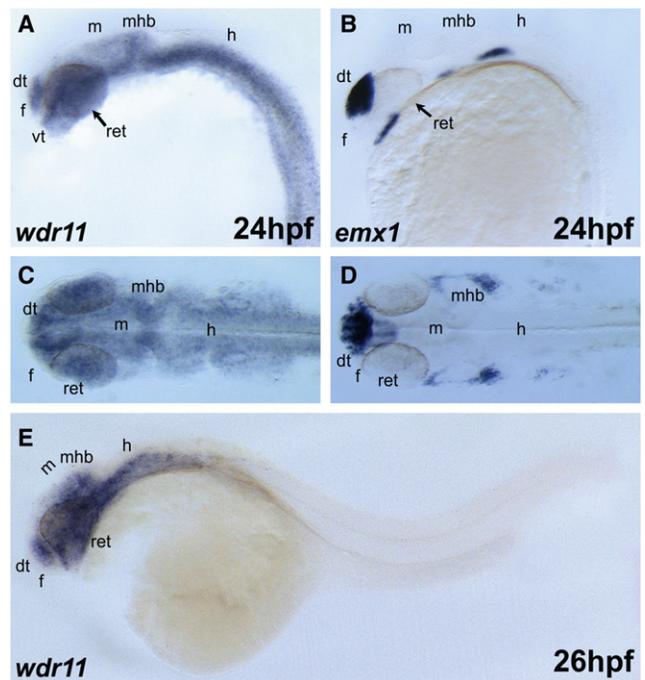
**Figure 6. *Wdr11* Expression during Murine Development**

(A–E) DIG-labeled whole-mount in situ hybridization with a *Wdr11* antisense probe at different embryonic stages. High expression levels are found in all structures of the developing brain as early as E10.5. Expression in the limbs is prominent at E12.5 and E13.5. Staining was also observed in both the hind and forelimb buds, but as limbs developed, it shifted toward the terminal phalanges. At E14.5 the olfactory bulb and the developing cortex show the highest expression levels. A magnification of the developing cortex and olfactory bulb is shown in (E). (F–H) Expression of *Wdr11* in the adult brain. [35S]-UTP-labeled in situ hybridizations show prominent *Wdr11* signals in the piriform cortex (F) as well as in the hippocampus and cerebellum (G). Note the higher signal intensity in the hypothalamic region within the dotted rectangle in (G). Single cells as well as clusters of neurons within the hypothalamic nuclei also showed *Wdr11* expression in DIG-labeled cryosections (H). Signals were absent with the sense control. Abbreviations are as follows: tel, telencephalic vesicle; wh, wall of hindbrain; di, diencephalon; mes, mesencephalon; rho, rhombencephalon; sc, spinal cord; lb, limb bud; hl, hind limb; fl, fore limb; ob, olfactory bulb; pc, piriform cortex; gcl, granule cell layer of the cerebellum; hp, hippocampus; and hy, hypothalamus.

30–36 hpf (Figure 7), again consistent with a potential role for a WDR11-EMX1 protein interaction.

## Discussion

Our mapping, cloning, and sequencing of the breakpoints of the de novo balanced chromosome translocation from a KS male with karyotype 46,XY,t(10;12)(q26.12;q13.11) dn and the consequent detection of multiple independent missense variants in *WDR11* in KS and IHH patients argues strongly for a causative role for *WDR11* in this disorder. Regrettably, DNA samples of the patients' parents are not



**Figure 7. Expression of *wdr11* in Developing Zebrafish Embryos** At 24 hpf, *wdr11* transcripts were broadly detected in forebrain, midbrain, and hindbrain.

(A, B, and E) Lateral view; anterior is to the left. (C and D) Dorsal view. The expression domain of *wdr11* in the brain partially overlapped with that of *emx1* (B and D), a dorsal telencephalon marker. Abbreviations are as follows: dt, dorsal telencephalon; f, forebrain; m, midbrain; h, hindbrain; mhb, midbrain-hindbrain boundary; vt, ventral telencephalon; and ret, retina.

available, and all efforts to track the parents have been unsuccessful. However, because all *WDR11* mutations were heterozygous, autosomal-dominant inheritance of the IHH phenotype is likely. Consistent with this interpretation, no *WDR11* mutation was found in any of 29 exons or splice junctions on the untranslocated second allele from our translocation patient, suggesting that a reduction in functional WDR11 as a result of dysregulation of the gene via a position effect of the translocation is the cause of KS in this subject (see Results and Figure S3). The absence of truncating nonsense and frameshift mutations indicates that these might produce a more severe phenotype or embryonic lethality.

*WDR11* was originally identified as a potential tumor suppressor by positional cloning of a somatically acquired t(10;19) chromosome translocation that generated an intragenic deletion in human glioblastoma cells.<sup>47</sup> This deletion disrupted *WDR11* and fused it to the *ZNF320* gene (MIM 606427). Interestingly, other genes associated with mammalian puberty are also known for their involvement in tumorigenesis.<sup>48</sup> For example, hypothalamic expression of certain tumor suppressor genes is increased at puberty in monkeys or decreased in mice with delayed puberty.<sup>49</sup> Similarly, *KISS1* (MIM 603286), a tumor metastasis suppressor gene<sup>50</sup> in melanomas<sup>51</sup> and breast carcinomas<sup>52</sup> encodes the peptide ligand of G-protein-coupled

receptor 54 (GPR54),<sup>53</sup> which plays a critical role in the initiation of puberty. *KISS1R* (encoding GPR54) mutations cause autosomal-recessive IHH in consanguineous families and mice.<sup>7,8</sup>

WDR11's function is unknown, but it contains twelve WD domains and is highly conserved throughout vertebrate evolution. The gene was previously annotated as *BRWD2* (Bromodomain and WD repeat domain containing 2) but has recently been more appropriately designated *WDR11* because it has WD domains but no Bromo domains.<sup>54</sup> Mutations of other WD proteins have been associated previously with mammalian reproduction. Both genders of *Repro5* mice, which have ENU-induced *Brwd1* (*Wdr9*) mutations, demonstrate infertility. In addition, patients with mutations of *BRWD3* (MIM 300553) have undescended testes and minimal facial or axillary hair.<sup>55</sup> However, this is the first report of a specific role for a WD protein in IHH or KS, and it implicates protein-protein interaction mediated by WD repeats of WDR11 as a critical requirement for normal puberty.

Proteins with repeated WD domains, each of which consists of four antiparallel  $\beta$  strands, form  $\beta$  propeller structures to support interactions with protein-binding partners and to organize and stabilize multiprotein complexes.<sup>56</sup> A  $\beta$  propeller is characterized by 4–8 blade-shaped  $\beta$  sheets arranged around a central axis; each sheet of four antiparallel  $\beta$  strands is twisted so that the first and fourth sheets are close to perpendicular. The last  $\beta$  strand of one WD repeat, and the first three  $\beta$  strands of the WD repeat form a blade of the  $\beta$  propeller (Figure 4A).

Three of the *WDR11* missense mutations leading to R395W, H690Q, and F1150L alter amino acid residues that are completely conserved in all 13 available mammalian and avian orthologs (human, chimpanzee, cow, horse, panda, pig, dog, rat, mouse, rabbit, opossum, chicken, and finch), and a fourth change, A435T, is shared in 11 out of 13 species, suggesting that these substitutions in six independent sporadic patients are very to be detrimental. Three of the missense alterations are located directly in the predicted propeller regions of WDR11: A435T and R448Q are in the sixth WD domain, and H690Q is in the ninth WD domain (Figure 4B). These three mutations are predicted by SPPIDER to alter protein-protein binding domains defined by protein modeling and therefore are likely to disrupt normal protein function (Figure S5).

The similarity in structural modeling between WDR11 and the known structure of AIP1 suggests that WDR11, like AIP1, may form a dimer stabilized by interaction with two zinc ions (Figure S6). Because the two protein structures are not identical, deviations arising from their alignment make the position of the Zn ion uncertain to a few angstroms (Figure S4). However, WDR11 has the required residues (Asp377, Glu384, His501, His508, and Glu510) for zinc binding in the vicinity of the putative zinc position (Figure S7). The R448Q mutation is less than 5 Å from the predicted zinc binding site. Arg residues

near Zn coordination sites do not directly interact with the Zn, but they stabilize their environment because Arg is highly positively charged. Replacing Arg with the much smaller Gln residue could influence the zinc-binding propensity of WDR11 and affect its dimer formation and interactions, including a potential actin interaction predicted by analogy with AIP1.<sup>57</sup>

In view of the importance of protein-protein interactions for the function of WD proteins, we sought binding partners for WDR11 and identified the transcription factor EMX1 as a novel interactor. EMX1 is a homeobox transcription factor involved in specifying cell fates in the developing central nervous system,<sup>39</sup> and it participates in the development of olfactory neurons.<sup>40</sup> This putative transcription factor has been shown to be one of the downstream target genes for Gli-Kruppel family member 3 (Gli3, MIM 165240) transcription factor,<sup>41,42</sup> which is a part of the Sonic hedgehog-Patched-Gli (Shh-Ptch-Gli) signaling pathway<sup>43</sup> important in endocrine signaling.

Analysis of the expression patterns of *WDR11* in human embryonic olfactory GnRH neuroblasts as well as in mouse and zebrafish development revealed overlapping patterns of expression with *EMX1* in regions critical for formation of the hypothalamus,<sup>45</sup> supporting the opportunity for the two proteins to interact in vivo and to act together during development. Deletion analysis revealed that WDR11 interacts with EMX1 via both its N terminus and its central region, where four (R395W, A435T, R448Q, and H690Q) of the five WDR11 missense alterations were found. R448Q reduced and both A435T and H690Q abolished binding to EMX1, physically decreasing the opportunity for productive interaction. Interestingly, the mutant R395W did not appear to affect EMX1 binding, but this represents alteration of an amino acid that is invariant in higher vertebrates, suggesting that this alteration permits binding but impairs the as-yet-undefined functional consequences of the WDR11-EMX1 interaction.

Taken together, our genetic and functional data provide strong evidence for missense sequence variants of *WDR11* as a cause of IHH and KS in a proportion of cases of this genetically heterogeneous condition. This adds to the growing list of genes known to be mutated in IHH and KS and will open new investigative routes for understanding the development of normal human puberty and reproduction. Importantly, the identification of EMX1 as a binding partner of WDR11 whose interaction can be disrupted by IHH- and KS-associated mutations is significant for two principal reasons. First, it places WDR11, whose biological function is not well understood, as a potential player in the Sonic hedgehog-Patched-Gli-Emx signaling pathway via its interaction with EMX1. Second, the demonstration of a developmental role for WDR11 in IHH and KS suggests a possible connection between Shh signaling and pubertal development. The potential for defects in this pathway in IHH and KS is worthy of further exploration.

## Supplemental Data

Supplemental Data include seven figures and four tables and can be found with this article online at <http://www.cell.com/AJHG/>.

## Acknowledgments

This paper is dedicated to the 80<sup>th</sup> birthday of Paul G. McDonough. We thank all families for their kind participation in this study. We are indebted to the following colleagues for this study: Stefanie Meien, Milton David Stuart, Karen Norris, Ihn-Sik Seong, Jae Ho Lee, Gabriela Alexandru, Wei Chen, Ines Müller, Fei Lan, Hans-Jürgen Kreienkamp, Enno Gößling, Duygu Duman, Lynn Chorich, Sandra Tho, Jill Mokry, Lisa Halvorson, and Temple Smith. We acknowledge support by the Landesstiftung Baden-Württemberg, the Deutsche Akademische Austauschdienst (German Academic Exchange Service) and the volunteers of [POEM@HOME](mailto:POEM@HOME). This study was supported in part by a Korea Science and Engineering Foundation (KOSEF) grant funded by the government of the Republic of Korea (The Ubiquitome Research Program, 2009-00983 to C.Y.C.) J.F.G. was supported by National Institutes of Health grant GM061354 for the Developmental Genome Anatomy Project. We also acknowledge support to L.C.L. from National Institutes of Health grants HD33004 and HD040287.

Received: May 20, 2010

Revised: August 10, 2010

Accepted: August 31, 2010

Published online: September 30, 2010

## Web Resources

The URLs for data presented herein are as follows:

NCBI Single Nucleotide Polymorphism Database, <http://www.ncbi.nlm.nih.gov/ezp-prod1.hul.harvard.edu/projects/SNP/>

NCBI BLAST, <http://blast.ncbi.nlm.nih.gov/>

Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/>

Primer design, [http://www.es.embnnet.org/cgi-bin/primer3\\_www.cgi](http://www.es.embnnet.org/cgi-bin/primer3_www.cgi)

UCSC Genome Browser, <http://genome.ucsc.edu/>

## Accession Numbers

Human *WDR11* mRNA, NM\_018117.11; human WDR11 protein, NP\_060587.8; mouse *WDR11* mRNA, NM\_172255.3; mouse WDR11 protein, NP\_758459.2; rat *WDR11* mRNA, XM\_219377.5; rat WDR11 protein, XP\_219377.5; chimpanzee WDR11 protein, XP\_508077.2; cow WDR11 protein, NP\_001094592.1; horse WDR11 protein, XP\_001496034.1; panda WDR11 protein, EFB13673.1; pig WDR11 protein, XP\_001924391.1; dog WDR11 protein, XP\_535039.2; rabbit WDR11 protein, XP\_002718739.1; opossum WDR11 protein, XP\_001376637.1; chicken WDR11 protein, XP\_421795.2; zebra finch WDR11 protein, XP\_002188976.1; zebrafish *wdr11* mRNA, XM\_682139.3; zebrafish *wdr11* protein, XP\_687231.2; human *PPAPDC1A* mRNA, NM\_001030059.1; human *SEC23IP* mRNA, NM\_007190.2.

## References

- Kim, H.G., Bhagavath, B., and Layman, L.C. (2008). Clinical manifestations of impaired GnRH neuron development and function. *Neurosignals* 16, 165–182.
- Franco, B., Guioli, S., Pragliola, A., Incerti, B., Bardoni, B., Tonlorenzi, R., Carozzo, R., Maestrini, E., Pieretti, M., Taillon-Miller, P., et al. (1991). A gene deleted in Kallmann's syndrome shares homology with neural cell adhesion and axonal path-finding molecules. *Nature* 353, 529–536.
- Legouis, R., Hardelin, J.P., Levilliers, J., Claverie, J.M., Compain, S., Wunderle, V., Millasseau, P., Le Paslier, D., Cohen, D., Caterina, D., et al. (1991). The candidate gene for the X-linked Kallmann syndrome encodes a protein related to adhesion molecules. *Cell* 67, 423–435.
- Dodé, C., Levilliers, J., Dupont, J.M., De Paepe, A., Le Dû, N., Soussi-Yanicostas, N., Coimbra, R.S., Delmaghani, S., Compain-Nouaille, S., Baverel, F., et al. (2003). Loss-of-function mutations in *FGFR1* cause autosomal dominant Kallmann syndrome. *Nat. Genet.* 33, 463–465.
- Kim, H.G., Herrick, S.R., Lemyre, E., Kishikawa, S., Salisz, J.A., Seminara, S., MacDonald, M.E., Bruns, G.A., Morton, C.C., Quade, B.J., and Gusella, J.F. (2005). Hypogonadotropic hypogonadism and cleft lip and palate caused by a balanced translocation producing haploinsufficiency for *FGFR1*. *J. Med. Genet.* 42, 666–672.
- Pitteloud, N., Acierno, J.S., Jr., Meysing, A., Eliseenkova, A.V., Ma, J., Ibrahim, O.A., Metzger, D.L., Hayes, F.J., Dwyer, A.A., Hughes, V.A., et al. (2006). Mutations in fibroblast growth factor receptor 1 cause both Kallmann syndrome and normosmic idiopathic hypogonadotropic hypogonadism. *Proc. Natl. Acad. Sci. USA* 103, 6281–6286.
- de Roux, N., Genin, E., Carel, J.C., Matsuda, F., Chaussain, J.L., and Milgrom, E. (2003). Hypogonadotropic hypogonadism due to loss of function of the *KISS1*-derived peptide receptor *GPR54*. *Proc. Natl. Acad. Sci. USA* 100, 10972–10976.
- Seminara, S.B., Messager, S., Chatzidaki, E.E., Thresher, R.R., Acierno, J.S., Jr., Shagoury, J.K., Bo-Abbas, Y., Kuohung, W., Schwinof, K.M., Hendrick, A.G., et al. (2003). The *GPR54* gene as a regulator of puberty. *N. Engl. J. Med.* 349, 1614–1627.
- Topaloglu, A.K., Reimann, F., Guclu, M., Yalin, A.S., Kotan, L.D., Porter, K.M., Serin, A., Mungan, N.O., Cook, J.R., Ozbek, M.N., et al. (2009). *TAC3* and *TACR3* mutations in familial hypogonadotropic hypogonadism reveal a key role for Neurokinin B in the central control of reproduction. *Nat. Genet.* 41, 354–358.
- de Roux, N., Young, J., Misrahi, M., Genet, R., Chanson, P., Schaison, G., and Milgrom, E. (1997). A family with hypogonadotropic hypogonadism and mutations in the gonadotropin-releasing hormone receptor. *N. Engl. J. Med.* 337, 1597–1602.
- Layman, L.C., Cohen, D.P., Jin, M., Xie, J., Li, Z., Reindollar, R.H., Bolbolan, S., Bick, D.P., Sherins, R.R., Duck, L.W., et al. (1998). Mutations in gonadotropin-releasing hormone receptor gene cause hypogonadotropic hypogonadism. *Nat. Genet.* 18, 14–15.
- Miura, K., Acierno, J.S., Jr., and Seminara, S.B. (2004). Characterization of the human nasal embryonic LHRH factor gene, *NELF*, and a mutation screening among 65 patients with idiopathic hypogonadotropic hypogonadism (IHH). *J. Hum. Genet.* 49, 265–268.
- Kim, H.G., Kurth, I., Lan, F., Melicani, I., Wenzel, W., Eom, S.H., Kang, G.B., Rosenberger, G., Tekin, M., Ozata, M., et al. (2008). Mutations in *CHD7*, encoding a chromatin-remodeling protein, cause idiopathic hypogonadotropic hypogonadism and Kallmann syndrome. *Am. J. Hum. Genet.* 83, 511–519.

14. Falardeau, J., Chung, W.C., Beenken, A., Raivio, T., Plummer, L., Sidis, Y., Jacobson-Dickman, E.E., Eliseenkova, A.V., Ma, J., Dwyer, A., et al. (2008). Decreased FGF8 signaling causes deficiency of gonadotropin-releasing hormone in humans and mice. *J. Clin. Invest.* *118*, 2822–2831.
15. Bouligand, J., Ghervan, C., Tello, J.A., Brailly-Tabard, S., Salenave, S., Chanson, P., Lombès, M., Millar, R.P., Guiochon-Mantel, A., and Young, J. (2009). Isolated familial hypogonadotropic hypogonadism and a GNRH1 mutation. *N. Engl. J. Med.* *360*, 2742–2748.
16. Chan, Y.M., de Guillebon, A., Lang-Muritano, M., Plummer, L., Cerrato, F., Tsiaras, S., Gaspert, A., Lavoie, H.B., Wu, C.H., Crowley, W.F., Jr., et al. (2009). GNRH1 mutations in patients with idiopathic hypogonadotropic hypogonadism. *Proc. Natl. Acad. Sci. USA* *106*, 11703–11708.
17. Dodé, C., Teixeira, L., Levilliers, J., Fouveaut, C., Bouchard, P., Kottler, M.L., Lespinasse, J., Lienhardt-Roussie, A., Mathieu, M., Moerman, A., et al. (2006). Kallmann syndrome: Mutations in the genes encoding prokineticin-2 and prokineticin receptor-2. *PLoS Genet.* *2*, e175.
18. Bhagavath, B., Podolsky, R.H., Ozata, M., Bolu, E., Bick, D.P., Kulharya, A., Sherins, R.J., and Layman, L.C. (2006). Clinical and molecular characterization of a large sample of patients with hypogonadotropic hypogonadism. *Fertil. Steril.* *85*, 706–713.
19. Suzuki, Y., Sasagawa, I., Nakada, T., and Onmura, Y. (1998). Bilateral cryptorchidism associated with terminal deletion of 10q. *Urol. Int.* *61*, 186–187.
20. Mutoh, A., Sasagawa, I., Tateno, T., Sawamura, T., and Nakada, T. (1999). Long arm deletion of chromosome 10 in a boy with monorchidism. *Scand. J. Urol. Nephrol.* *33*, 77–78.
21. Leonard, N.J., Harley, F.L., and Lin, C.C. (1999). Terminal deletion of chromosome 10q at band 26.1: Follow-up in an adolescent male with high-output renal failure from congenital obstructive uropathy. *Am. J. Med. Genet.* *86*, 115–117.
22. Baccetti, B., Bruni, E., Collodel, G., Gambera, L., Moretti, E., Marzella, R., and Piomboni, P. (2003). 10, 15 reciprocal translocation in an infertile man: Ultrastructural and fluorescence in-situ hybridization sperm study: case report. *Hum. Reprod.* *18*, 2302–2308.
23. Ogata, T., Muroya, K., Sasagawa, I., Kosho, T., Wakui, K., Sakazume, S., Ito, K., Matsuo, N., Ohashi, H., and Nagai, T. (2000). Genetic evidence for a novel gene(s) involved in urogenital development on 10q26. *Kidney Int.* *58*, 2281–2290.
24. Lukusa, T., and Fryns, J.P. (2000). Pure distal monosomy 10q26 in a patient displaying clinical features of Prader-Willi syndrome during infancy and distinct behavioural phenotype in adolescence. *Genet. Couns.* *11*, 119–126.
25. Bofinger, M.K., Opitz, J.M., Soukup, S.W., Ekblom, L.S., Phillips, S., Daniel, A., and Greene, E.W. (1991). A familial MCA/MR syndrome due to translocation t(10;16)(q26;p13.1): report of six cases. *Am. J. Med. Genet.* *38*, 1–8.
26. Schinzel, A., Lorda-Sanchez, I., Binkert, F., Carter, N.P., Bebb, C.E., Ferguson-Smith, M.A., Eiholzer, U., Zachmann, M., and Robinson, W.P. (1995). Kallmann syndrome in a boy with a t(1;10) translocation detected by reverse chromosome painting. *J. Med. Genet.* *32*, 957–961.
27. Higgins, A.W., Alkuraya, F.S., Bosco, A.F., Brown, K.K., Bruns, G.A., Donovan, D.J., Eisenman, R., Fan, Y., Farra, C.G., Ferguson, H.L., et al. (2008). Characterization of apparently balanced chromosomal rearrangements from the developmental genome anatomy project. *Am. J. Hum. Genet.* *82*, 712–722.
28. Crowley, W.F., Jr., Filicori, M., Spratt, D.I., and Santoro, N.F. (1985). The physiology of gonadotropin-releasing hormone (GnRH) secretion in men and women. *Recent Prog. Horm. Res.* *41*, 473–531.
29. Hermans-Borgmeyer, I., Hampe, W., Schinke, B., Methner, A., Nykjaer, A., Süsens, U., Fenger, U., Herbarth, B., and Schaller, H.C. (1998). Unique expression pattern of a novel mosaic receptor in the developing cerebral cortex. *Mech. Dev.* *70*, 65–76.
30. Kim, H.T., Kim, E.H., Yoo, K.W., Lee, M.S., Choi, J.H., Park, H.C., Yeo, S.Y., Lee, D.S., and Kim, C.H. (2008). Isolation and expression analysis of Alzheimer's disease-related gene *xb51* in zebrafish. *Dev. Dyn.* *237*, 3921–3926.
31. Morita, T., Nitta, H., Kiyama, Y., Mori, H., and Mishina, M. (1995). Differential expression of two zebrafish *emx* homeo-protein mRNAs in the developing brain. *Neurosci. Lett.* *198*, 131–134.
32. Bhagavath, B., Ozata, M., Ozdemir, I.C., Bolu, E., Bick, D.P., Sherins, R.J., and Layman, L.C. (2005). The prevalence of gonadotropin-releasing hormone receptor mutations in a large cohort of patients with hypogonadotropic hypogonadism. *Fertil. Steril.* *84*, 951–957.
33. Kalscheuer, V.M., FitzPatrick, D., Tommerup, N., Bugge, M., Niebuhr, E., Neumann, L.M., Tzschach, A., Shoichet, S.A., Menzel, C., Erdogan, F., et al. (2007). Mutations in autism susceptibility candidate 2 (*AUTS2*) in patients with mental retardation. *Hum. Genet.* *121*, 501–509.
34. Tzschach, A., Bisgaard, A.M., Kirchhoff, M., Graul-Neumann, L.M., Neitzel, H., Page, S., Ahmed, A., Müller, I., Erdogan, F., Ropers, H.H., et al. (2010). Chromosome aberrations involving 10q22: Report of three overlapping interstitial deletions and a balanced translocation disrupting *C10orf11*. *Eur. J. Hum. Genet.* *18*, 291–295.
35. Li, K.B. (2003). ClustalW-MPI: ClustalW analysis using distributed and parallel computing. *Bioinformatics* *19*, 1585–1586.
36. Mohri, K., Vorobiev, S., Fedorov, A.A., Almo, S.C., and Ono, S. (2004). Identification of functional residues on *Caenorhabditis elegans* actin-interacting protein 1 (*UNC-78*) for disassembly of actin depolymerizing factor/cofilin-bound actin filaments. *J. Biol. Chem.* *279*, 31697–31707.
37. Voegtli, W.C., Madrona, A.Y., and Wilson, D.K. (2003). The structure of *Aip1p*, a WD repeat protein that regulates Cofilin-mediated actin depolymerization. *J. Biol. Chem.* *278*, 34373–34379.
38. Franke, W.W. (2004). Actin's many actions start at the genes. *Nat. Cell Biol.* *6*, 1013–1014.
39. Bishop, K.M., Garel, S., Nakagawa, Y., Rubenstein, J.L., and O'Leary, D.D. (2003). *Emx1* and *Emx2* cooperate to regulate cortical size, lamination, neuronal differentiation, development of cortical efferents, and thalamocortical pathfinding. *J. Comp. Neurol.* *457*, 345–360.
40. Lichtneckert, R., Nobs, L., and Reichert, H. (2008). Empty spiracles is required for the development of olfactory projection neuron circuitry in *Drosophila*. *Development* *135*, 2415–2424.
41. Theil, T., Alvarez-Bolado, G., Walter, A., and Rütger, U. (1999). *Gli3* is required for *Emx* gene expression during dorsal telencephalon development. *Development* *126*, 3561–3571.

42. Tole, S., Ragsdale, C.W., and Grove, E.A. (2000). Dorsoroventral patterning of the telencephalon is disrupted in the mouse mutant extra-toes(J). *Dev. Biol.* *217*, 254–265.
43. Villavicencio, E.H., Walterhouse, D.O., and Iannaccone, P.M. (2000). The sonic hedgehog-patched-gli pathway in human development and disease. *Am. J. Hum. Genet.* *67*, 1047–1054.
44. Hu, Y., Guimond, S.E., Travers, P., Cadman, S., Hohenester, E., Turnbull, J.E., Kim, S.H., and Bouloux, P.M. (2009). Novel mechanisms of fibroblast growth factor receptor 1 regulation by extracellular matrix protein anosmin-1. *J. Biol. Chem.* *284*, 29905–29920.
45. Schwarting, G.A., Wierman, M.E., and Tobet, S.A. (2007). Gonadotropin-releasing hormone neuronal migration. *Semin. Reprod. Med.* *25*, 305–312.
46. Simeone, A., Gulisano, M., Acampora, D., Stornaiuolo, A., Rambaldi, M., and Boncinelli, E. (1992). Two vertebrate homeobox genes related to the *Drosophila* empty spiracles gene are expressed in the embryonic cerebral cortex. *EMBO J.* *11*, 2541–2550.
47. Chernova, O.B., Hunyadi, A., Malaj, E., Pan, H., Crooks, C., Roe, B., and Cowell, J.K. (2001). A novel member of the WD-repeat gene family, WDR11, maps to the 10q26 region and is disrupted by a chromosome translocation in human glioblastoma cells. *Oncogene* *20*, 5378–5392.
48. Ojeda, S.R., Dubay, C., Lomniczi, A., Kaidar, G., Matagne, V., Sandau, U.S., and Dissen, G.A. (2010). Gene networks and the neuroendocrine regulation of puberty. *Mol. Cell. Endocrinol.* *324*, 3–11.
49. Parent, A.S., Matagne, V., Westphal, M., Heger, S., Ojeda, S., and Jung, H. (2008). Gene expression profiling of hypothalamic hamartomas: A search for genes associated with central precocious puberty. *Horm. Res.* *69*, 114–123.
50. Lee, J.H., Miele, M.E., Hicks, D.J., Phillips, K.K., Trent, J.M., Weissman, B.E., and Welch, D.R. (1996). KiSS-1, a novel human malignant melanoma metastasis-suppressor gene. *J. Natl. Cancer Inst.* *88*, 1731–1737.
51. Lee, J.H., and Welch, D.R. (1997). Identification of highly expressed genes in metastasis-suppressed chromosome 6/human malignant melanoma hybrid cells using subtractive hybridization and differential display. *Int. J. Cancer* *71*, 1035–1044.
52. Lee, J.H., and Welch, D.R. (1997). Suppression of metastasis in human breast carcinoma MDA-MB-435 cells after transfection with the metastasis suppressor gene, KiSS-1. *Cancer Res.* *57*, 2384–2387.
53. Ohtaki, T., Shintani, Y., Honda, S., Matsumoto, H., Hori, A., Kanehashi, K., Terao, Y., Kumano, S., Takatsu, Y., Masuda, Y., et al. (2001). Metastasis suppressor gene KiSS-1 encodes peptide ligand of a G-protein-coupled receptor. *Nature* *411*, 613–617.
54. Philipps, D.L., Wigglesworth, K., Hartford, S.A., Sun, F., Pattabiraman, S., Schimenti, K., Handel, M., Eppig, J.J., and Schimenti, J.C. (2008). The dual bromodomain and WD repeat-containing mouse protein BRWD1 is required for normal spermiogenesis and the oocyte-embryo transition. *Dev. Biol.* *317*, 72–82.
55. Field, M., Tarpey, P.S., Smith, R., Edkins, S., O'Meara, S., Stevens, C., Tofts, C., Teague, J., Butler, A., Dicks, E., et al. (2007). Mutations in the BRWD3 gene cause X-linked mental retardation associated with macrocephaly. *Am. J. Hum. Genet.* *81*, 367–374.
56. Higa, L.A., and Zhang, H. (2007). Stealing the spotlight: CUL4-DDB1 ubiquitin ligase docks WD40-repeat proteins to destroy. *Cell Div.* *2*, 5.
57. Kudryashov, D.S., Sawaya, M.R., Adisetiyo, H., Norcross, T., Hegyi, G., Reisler, E., and Yeates, T.O. (2005). The crystal structure of a cross-linked actin dimer suggests a detailed molecular interface in F-actin. *Proc. Natl. Acad. Sci. USA* *102*, 13105–13110.

## Supplemental Data

### WDR11, a WD Protein that Interacts

### with Transcription Factor EMX1, Is Mutated in Idiopathic

### Hypogonadotropic Hypogonadism and Kallmann Syndrome

Hyung-Goo Kim, Jang-Won Ahn, Ingo Kurth, Reinhard Ullmann, Hyun-Taek Kim, Anita Kulharya, Kyung-Soo Ha, Yasuhide Itokawa, Irene Meliciani, Wolfgang Wenzel, Deresa Lee, Georg Rosenberger, Metin Ozata, David P. Bick, Richard J. Sherins, Takahiro Nagase, Mustafa Tekin, Soo-Hyun Kim, Cheol-Hee Kim, Hans-Hilger Ropers, James F. Gusella, Vera Kalscheuer, Cheol Yong Choi, Lawrence C. Layman

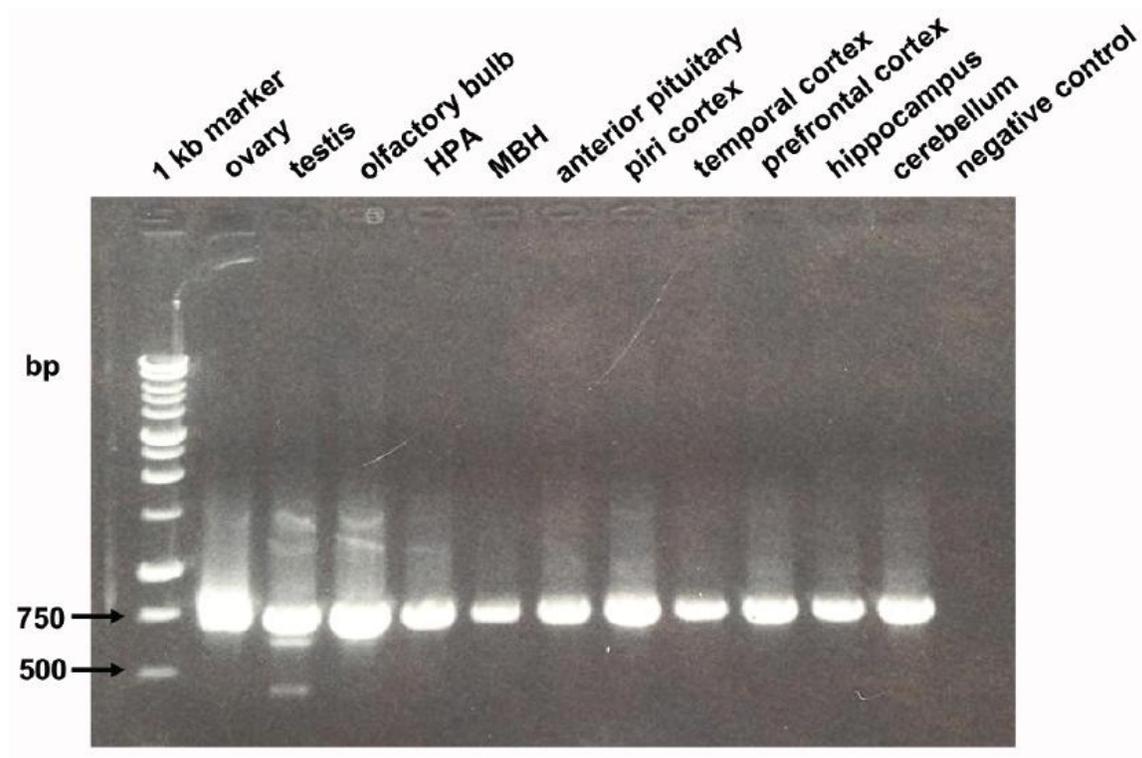
#### TABLE OF CONTENTS

1. Figure S1: Semi-quantitative RT-PCR analysis in rat tissues
2. Figure S2: Evolutionary conservation of the residue Lys978
3. Figure S3: RT-qPCR and western blot analysis of t(10;12)
4. Figure S4: Overlay of 1PGU and WDR11
5. Figure S5: Protein-protein interaction region of WDR11
6. Figure S6: Two zinc ion binding sites of 1PGU
7. Figure S7: Glu, His and Asp residues in the model of WDR11 in the vicinity of the proposed zinc position based upon 1PGU
8. Table S1: SNPs of *WDR11* (NM\_018117.11) found in IHH/KS patients and normal controls
9. Table S2: WD40 domains in WDR11
10. Table S3: BAC clones used for FISH analysis of t(10;12)
11. Table S4: Primers used in this study

## SUPPLEMENTARY FIGURES AND TABLES

**Figure S1: Semi-quantitative RT-PCR analysis in rat tissues**

Total RNA isolated from 11 rat IHH/KS-relevant tissues was used for RT-PCR using rat-specific *Wdr11* primers. 736 bp amplicons were detected in all tissues after 35 cycles; in particular, expression levels were higher in the ovary, testis, olfactory bulb and piri cortex. HPA: hypothalamus preoptic area; MBH: medial basal hypothalamus.



## Figure S2: Evolutionary conservation of the residue Lys978

ClustalW multiple alignment of partial protein sequence of WDR11 orthologs. The position of residue K978 altered by one heterozygous nucleotide change of *WDR11* is marked by arrow and red letters in the corresponding segment of the multiple alignment. The amino acid residues that differ from the sequence of the human WDR11 protein are indicated blue. Lys978 residue is evolutionarily fully conserved in all thirteen available WDR11 orthologs.

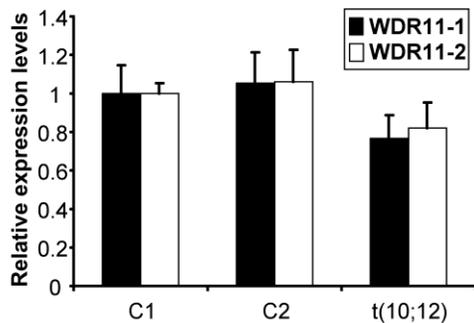
	K978
human	ENAYFQ <b>K</b> FQLERV
chimpanzee	ENAYFQ <b>K</b> FQLERV
cow	ENAYFQ <b>K</b> FQLERV
horse	ENAYFQ <b>K</b> FQLERV
panda	ENAYFQ <b>K</b> FQLERV
pig	ENAYFQ <b>K</b> FQLERV
dog	ENAYFQ <b>K</b> FQLERV
rat	ENAYFQ <b>K</b> FQLERV
mouse	ENAYFQ <b>K</b> FQLERV
rabbit	EN <b>T</b> YFQ <b>K</b> FQLERV
opossum	EN <b>S</b> YFQ <b>K</b> FQLERV
chicken	EN <b>H</b> YFQ <b>K</b> FQLERV
finch	EN <b>S</b> YFQ <b>K</b> FQLERV

**Figure S3: RT-qPCR and western blot analysis of t(10;12)**

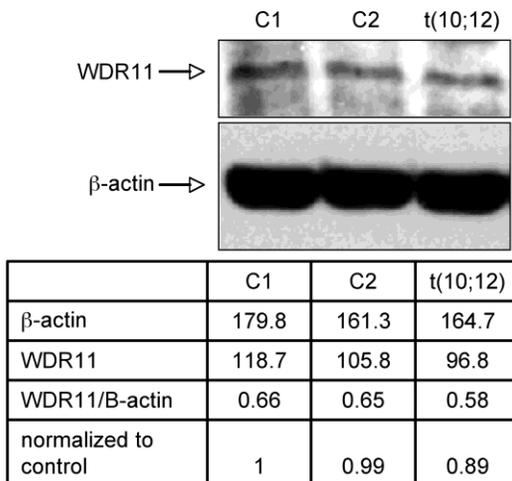
(A) *WDR11* mRNA levels in controls and the t(10;12) patient were measured by real time quantitative RT-PCR. Two different sets of primers (*WDR11*-1 and -2) were used to validate *WDR11* mRNA expression, which was normalized against *GAPDH*. The level of *WDR11* mRNA in the t(10;12) patient was reduced by 20% compared to two male controls C1 and C2, suggesting the positional effect of the translocation breakpoint at 10q26.12. The data show the mean of three independent experiments  $\pm$  SD.

(B) The anti-*WDR11* antibody recognizes ~120kDa *WDR11* protein in both male controls (C1 and C2) and t(10;12) lymphoblastoid cell extracts, showing about 10% reduced expression in t(10;12) (arrow).  $\beta$ -actin expression served as the internal loading control.

A

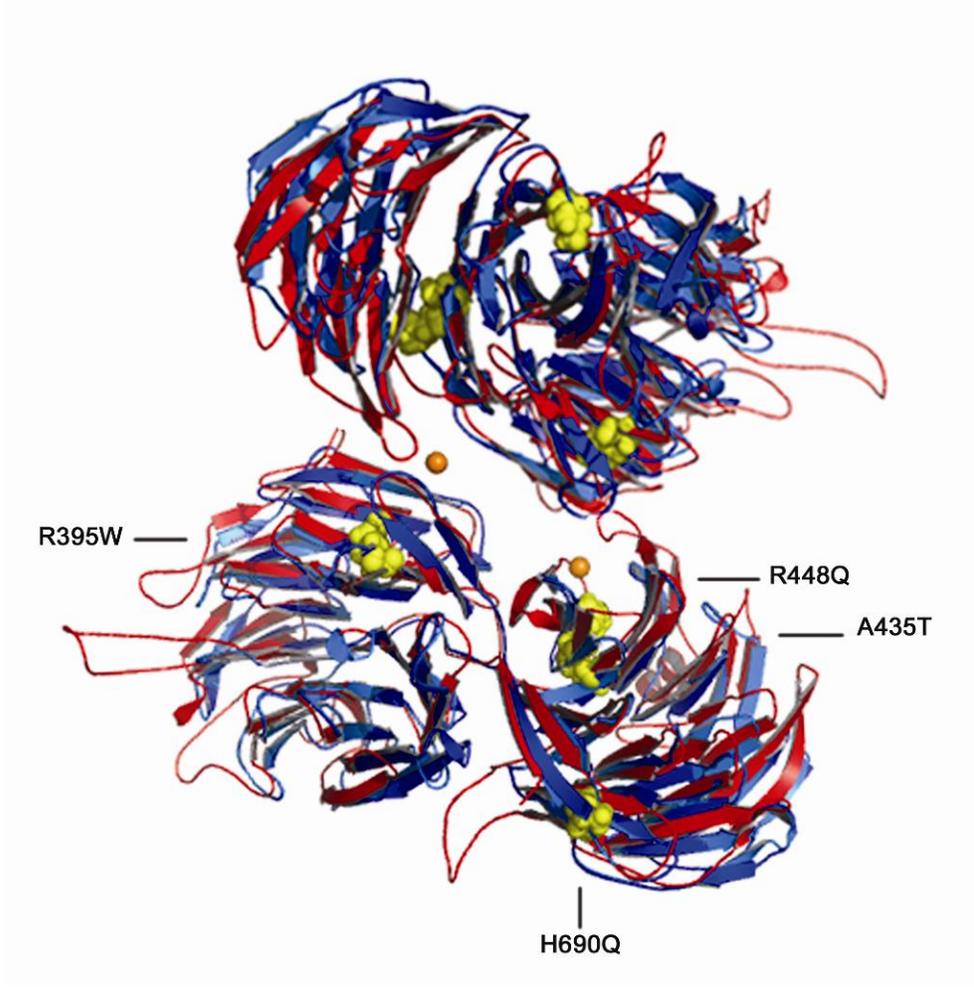


B



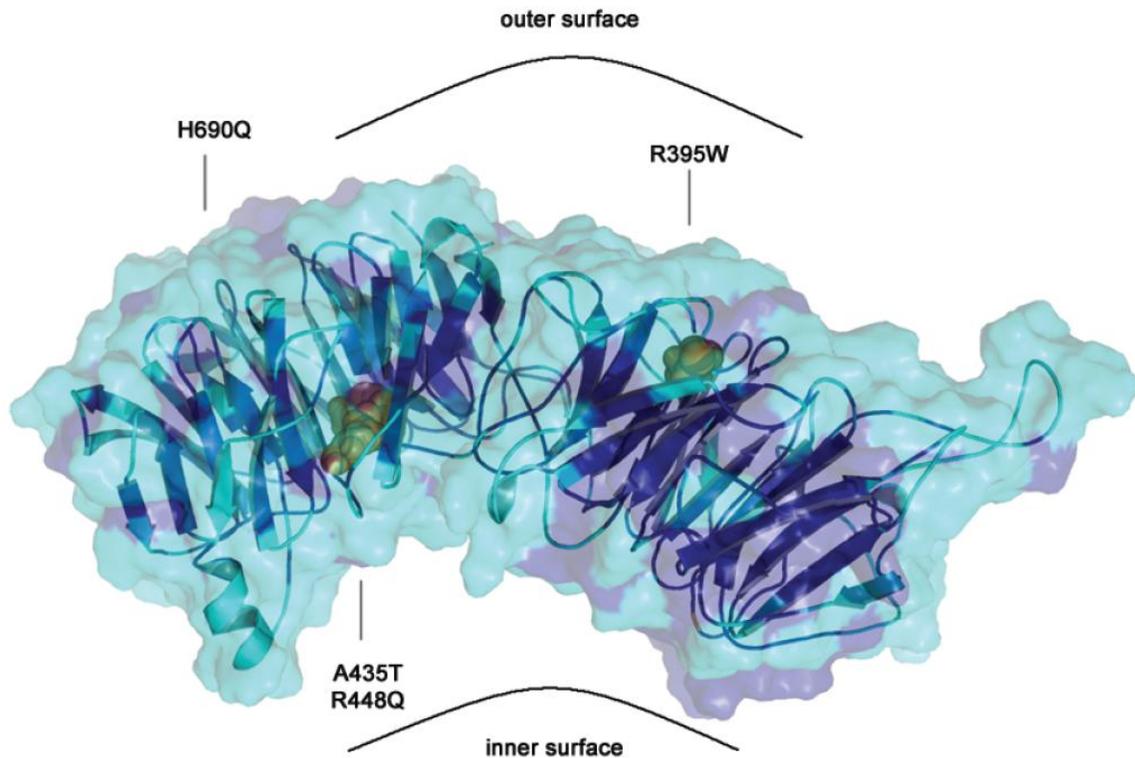
#### Figure S4: Overlay of 1PGU and WDR11

When we overlapped two replicas of the model of WDR11 (red) with the structure of 1PGU (blue) we found very good structural agreement, which led to the hypothesis that WDR11 may also dimerize by a similar mechanism involving two zinc coordination sites and might also be an actin binding protein. Because the two protein structures are not identical, deviations arising from their alignment make the position of the zinc ions (orange spheres) uncertain to a few angstroms. The sites of mutations in WDR11 are indicated with yellow spheres. To check our hypothesis we investigated whether WDR11 has a set of zinc-coordinating amino acids in the vicinity a putative zinc position, that we obtain, when we structurally align 1PGU with a dimer of WDR11. Because the two protein structures are not identical, deviations arising from their alignment make the position of the Zn ion uncertain to a few Angstrom. However, Fig. s7 clearly shows that also WDR11 has the required residues (ASP, GLU or HIS) for zinc binding in the vicinity of the putative zinc position. These residues are Asp 377, Glu 384, His 501, His 508, Glu 510. The two units, each of which has two beta propeller elements, are stabilized by two zinc ions as illustrated in Fig. S6<sup>1</sup>. 1PGU contains two zinc binding sites in which a typical pattern of residues (green sticks) coordinates with the zinc ions (orange). Since these residues belong to different chains the zinc coordination acts as a tether stabilizing the dimer.



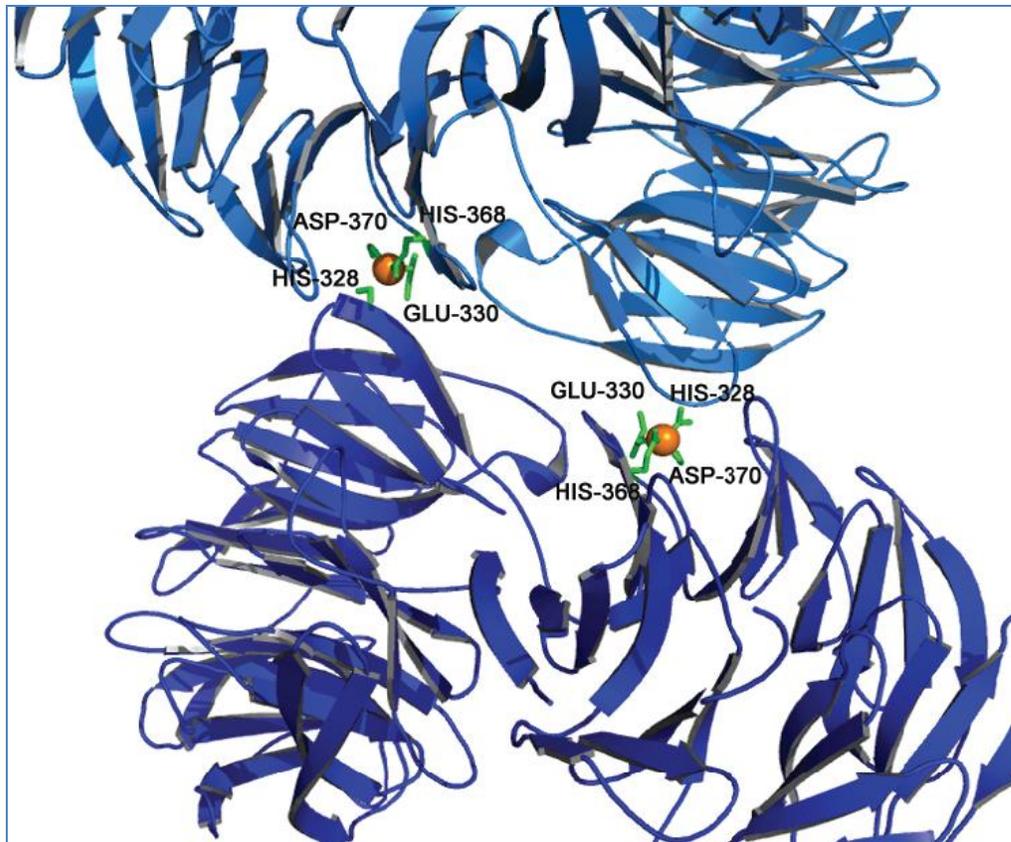
### Figure S5: Protein-protein interaction region of WDR11

We used SPPIDER to detect protein-protein interaction regions (in bright cyan) in the model for WDR11 (non-interacting regions in dark blue), here shown in a view obtained by rotating panel A 90 degrees out of the plane on a horizontal axis. The Figure shows two prominent external interaction regions (labeled inner & outer surface), where WDR11 may interact with multiple other proteins, such as EMX1 or actin, as well as an intramolecular interaction region at the interface of the two propeller units. Four mutations (those that are visible are illustrated with orange sidechains and labeled) lie within protein-protein interaction regions of WDR11 and may thus directly influence binding to EMX1.



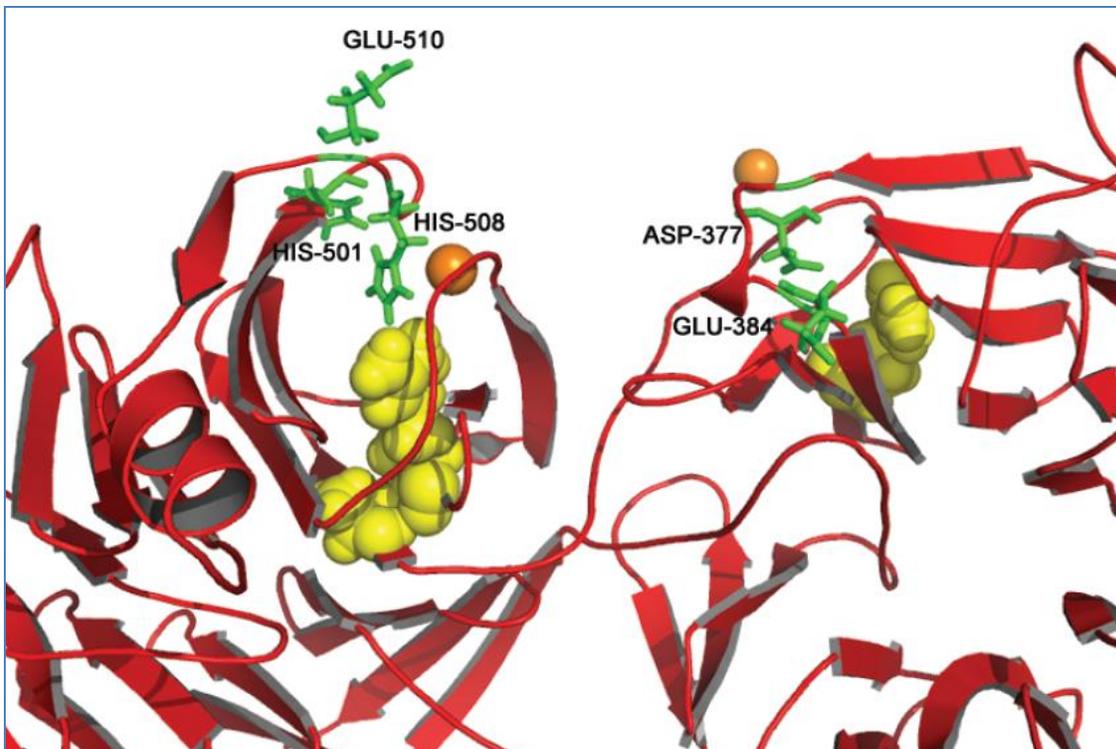
### Figure S6: Two zinc ion binding sites of 1PGU

We have identified other beta-propeller proteins that are highly homologous to WDR11, in particular, actin-interacting protein 1 (pdb code 1PGU), which is comprised of a dimer of two double propeller units that are present in 1NR0 and WDR11. The two units, each of which has two  $\beta$ -propeller elements, are stabilized by two zinc ions (shown in orange) as illustrated here <sup>1</sup>. Zinc ions in proteins are coordinated by a highly conserved set of amino acids, most commonly His, Glu, Asp and Cys <sup>2</sup>. The experimental structure of 1PGU shows the interaction between zinc and amino acids Asp, His and Glu. Both Zn ions are stabilized by residues coming from different molecules in the dimer and they therefore act as tethers that hold the dimer together.



**Figure S7: Glu, His and Asp residues in the model of WDR11 in the vicinity of the proposed zinc position based upon 1PGU**

A single unit of the proposed dimer is shown. The presence of the required set of amino acids demonstrates that WDR11 might also bind zinc and form a dimer that is stabilized by the same mechanism as 1PGU. WDR11 has the required residues (Asp 377, Glu 384, His 501, His 508, Glu 510) for zinc binding in the vicinity of the putative zinc position.



**Table S1: SNPs of *WDR11* (NM\_018117.11) found in IHH/KS patients and normal controls**

Location	Nucleotide change	Amino-acid change	Frequency of second allele in patients	dbSNP or frequency of second allele in normal controls
exon 1	c.-6G>A	–	3/402	rs17100985 5/96 (1) <sup>a</sup>
exon 1	c.51G>T	Gly17Gly	3/402	rs35692153 4/96 (1) <sup>a</sup>
intron 2	IVS2-7G>C	–	91/402 (8) <sup>a</sup>	rs2241846 17/96 (5) <sup>a</sup>
exon 6	c.834G>A	Thr278Thr	164/402 (31) <sup>a</sup>	rs10886789 32/96 (7) <sup>a</sup>
exon 6	c.789C>G	Leu263Leu	1/402	0/96
exon 8	c.1066G>A	Val356Ile	0/402	rs34304988 3/936
exon 8	c.1113C>T	Ala371Ala	1/402	rs41287986 2/936
exon 10	c.1425G>A	Pro475Pro	0/402	6/1466
exon 11	c.1542C>T	His514His	5/402	rs12355108 2/96
intron 12	IVS12+25G>A	–	6/402 (1) <sup>a</sup>	rs7923412 2/96
intron 13	IVS13+11A>C	–	1/402	
exon 15	c.1899A>T	Ala633Ala	169/402 (35) <sup>a</sup>	rs7899928 40/96 (10) <sup>a</sup>
exon 18	c.2304A>G	Ala768Ala	76/402 (5) <sup>a</sup>	rs2289337 19/96 (2) <sup>a</sup>
intron 18	IVS18-9C>T		1/402	1/548
intron 21	IVS21+48A>C	–	1/402	0/96
intron 21	IVS21-5C>G	–	1/402	1/456
intron 23	IVS23-48C>T	–	2/402	1/456
intron 23	IVS23-30G>C	–	2/402	7/648 (2) <sup>a</sup>
exon 24	c.2932A>C	Lys978Gln	1/402	2/1174
exon 24	c.2958A>G	Leu986Leu	163/402 (29) <sup>a</sup>	rs1652727 216/648 (38) <sup>a</sup>
intron 26	IVS26+11C>T	–	7/402	3/456
intron 26	IVS26+39C>T	–	259/402 (72) <sup>a</sup>	rs1866516 279/456 (87) <sup>a</sup>
intron 26	IVS26-12C>T	–	1/402	no SNP data

exon 27	c.3318C>T	Ala1106Ala	1/402	1/96 rs34937000
exon 27	c.3363C>G	Val1121Val	73/402 (2) <sup>a</sup>	rs3740307 12/96 (3) <sup>a</sup>
exon 27	c.3393C>T	Leu1131Leu	1/402	rs12268298 4/96 (1) <sup>a</sup>

<sup>a</sup>Total number of homozygotes among the individuals screened

**Table S2: WD domains in WDR11**

WD domains in WDR11 predicted based upon the structural model (WD domains from 2 to 10) and confidently predicted from the server SMART (WD domains 1, 11, and 12)

Number	Name	Begin	End	Size (AA)	E-value	Blast Confirmation
1	WD	50	99	50	0.04	Begin/end 50-99
2	WD	126	164	39	—	—
3	WD	220	274	55	—	—
4	WD	277	324	48	—	—
5	WD	335	374	40	—	—
6	WD	428	467	40	—	—
7	WD	561	605	45	—	—
8	WD	609	646	38	—	—
9	WD	652	691	40	—	—
10	WD	708	736	29	—	—
11	WD	738	778	41	10.30	738-778
12	WD	781	822	42	13.80	781-822

**Table S3: BAC clones used for FISH analysis of t(10;12)**

FISH result of BACs from Chromosome 10q26 (coordinates are from NCBI Build 36/hg18 assembly of UCSC Genome Browser)

No.	BAC ID	Start (bp)	End (bp)	Proximal	Spanning	Distal
1	RP11-592D19	118,702,524	118,893,472	X		
2	RP11-714M16	121,088,268	121,281,904	X		
3	RP11-717L13	121,765,054	121,963,545	X		
4	RP11-254K03	121,923,506	122,098,857		X	
5	RP11-572P18	122,039,380	122,255,369			X
6	RP11-313F09	122,302,049	122,499,951			X
7	RP11-499E06	122,574,466	122,745,931			X
8	RP11-152I12	122,709,923	122,891,863			X
9	RP11-604B15	122,847,905	123,036,763			X
10	RP11-251P02	122,860,000	123,009,780			X
11	RP11-753P11	123,034,946	123,207,777			X
12	RP11-300A10	123,166,882	123,350,993			X
13	RP11-7N04	123,407,405	123,552,165			X
14	RP11-105F10	123,764,473	123,939,046			X
15	RP11-619E04	124,103,531	124,311,827			X
16	RP13-63P20	124,872,781	125,056,306			X
17	RP11-47H10	125,699,910	125,916,390			X
18	RP11-8D22	130,325,185	130,493,639			X
19	RP11-91E2	135,150,439	135,301,208			X

FISH result of BACs from Chromosome 12q12-12q13 (coordinates are from NCBI Build 36/hg18 assembly of UCSC Genome Browser)

No.	BAC ID	Start (bp)	End (bp)	Proximal	Spanning	Distal
1	RP11-88L2	42,459,839	42,629,204	X		
2	RP11-462E15	42,875,295	43,061,579	X		
3	RP11-424H10	43,693,902	43,853,320	X		
4	RP11-19E18	44,441,691	44,596,850	X		
5	RP11-659M15	44,897,959	45,111,797	X		
6	RP11-755G20	45,222,971	45,399,371	X		
7	RP11-618L22	45,523,763	45,703,378	X		
8	RP11-72L16	45,581,486	45,734,180	X		
9	RP11-23J18	45,755,428	45,925,226	X		
10	RP11-479I2	45,789,177	45,975,868	X		
11	RP11-677P12	45,813,066	45,983,115	X		
12	RP11-71G17	45,875,143	46,046,421	X		
13	RP11-464D5	45,880,805	46,073,952		X	
14	RP11-543D1	46,000,452	46,161,075			X
15	RP11-648F12	46,021,097	46,184,193			X
16	RP11-379M8	46,074,480	46,267,797			X
17	RP11-650J13	46,467,143	46,641,379			X
18	RP11-204C20	46,894,554	47,075,616			X
19	RP11-141C14	50,423,519	50,612,911			X
20	RP11-762I7	54,296,170	54,469,906			X

**Table S4: Primers used in this study**

Semi-quantitative RT-PCR in rat tissues (rat <i>WDR11</i> ; XM_219377.5)	
Forward primer	5'-GACCTTGAAGTGAATCAAACAGTTGGCGTGATTGCAATTGAGCGAAC-3'
Reverse primer	5'-CACACGAGTGGACGCTTAGCTCCTTGTGTAGCAGACC-3'
Amplicon size	736 bp
Annealing temperature	68 °C

<i>WDR11</i> mutant constructs			
WDR11 mutants	Type	Outside primers	Mutant primers
m1	R395W	F01: 5'- GCTATCCACCCGCCAAATTA-3'	1183_mR1: 5'- TACTGTTCCATGAATTTGATTACAAACTG-3'
		R01: 5'- CTGCTCACCATCTGAACCTC-3'	1183_mF1: 5'- CGAAATTCATGGAACAGTAGTTCTGGTGTG-3'
m2	A435T	F01: 5'- GCTATCCACCCGCCAAATTA-3'	1303_mR1: 5'- CAGCAATTGTACTTTGCCCAATCATGTTAT-3'
		R01: 5'- CTGCTCACCATCTGAACCTC-3'	1303_mF1: 5'- GGGCAAAGTACAATTGCTGGGGAAGAACAT-3'
m3	R448Q	F01: 5'- GCTATCCACCCGCCAAATTA-3'	1343_mR1: 5'- TGCACTTCCTGCAGAATTGAACCTCTGGGA-3'
		R01: 5'- CTGCTCACCATCTGAACCTC-3'	1343_mF1: 5'- CAATTCTGCAGGAAGTGCACCTCAAGTTCC-3'
m4	H690Q	F01: 5'- GCTATCCACCCGCCAAATTA-3'	2070_mR1: 5'- AACAGTGAGTTGATACACTTGGCCATCAAT-3'
		R01: 5'- CTGCTCACCATCTGAACCTC-3'	2070_mF1: 5'- AGTGTATCAACTCACTGTTGAAGGAAACTC-3'
m5	F1150L	F06: 5'- ATCGCTTGGAAGGTGATAC-3'	3450_mR1: 5'- TGCTCTATCCAAGTATCTCATGCTGTGAAG-3'
		R02: 5'- CTTCAATGGGTTCTTCCTTG-3'	3450_mF1: 5'- GAGATACTTGGATAGAGCAGCCTTATTTGT-3'

Y2H and Binding Assay

Construct	Top primer	Bottom primer
pENTR3C-WDR11 1-361 [N]	5'- GATGGATCCA TGTTGCCCTA CACAGTG -3'	5' - GTAGTCGACC ATACTGAAGG GACGGAC - 3'
pENTR3C-WDR11 362-830 [M]	5'- GATGGATCCG TGTGCTGTCC TGTC AAT -3'	5' - GATGTCGACA AAGCACGCAG ACTTCAT - 3'
pENTR3C-WDR11 831-1224 [C]	5'- GATGGATCCA GAATGGATGA ACAAGAG -3'	5' - GATGTCGACT CACTCTTCAA TGGGTTC -3'
pENTR3C-WDR11 1-1224	5'- GATGGATCCA TGTTGCCCTA CACAGTG -3'	5' - GATGTCGACT CACTCTTCAA TGGGTTC -3'
pENTR3C-Emx1 1-290	5'- GAT GAA TTC ATG TGC CTG GCT GGG TGC -3'	5' - GAT CTC GAG CTA GTC ATT GGA GGT GAC -3'
pGBKT7-WDR11 1-830	5'- GATGGATCCA TGTTGCCCTA CACAGTG -3'	5' - GATGTCGACA AAGCACGCAG ACTTCAT - 3'

Mouse *in situ* hybridization analysis (NM\_172255.3)

WDR11 probe	Forward primer	Reverse primer
641-1035 of NM_172255.3	5'- GAATTCGGGAAGAAAGTGTACATCTCCAG-3'	5'-TCTAGATCATTGATGTAGTACAGATGC-3'
2892-3492 of NM_172255.3	5'- GAATTCTAACCAAGGAAGGTGCTCCTAA-3'	5'-TCTAGATATCTCATGCTGTGAAGCGTCT-3'

Mutation Screening of WDR11 (NM\_018117.11)

Exon number	Outer primer pair	Nested primer pair
1	F01: 5'-AGAGTGC GGAACCTAG-3'	F01: 5'-TAGGAAACTGAAGGCAAC-3'
	R01: 5'-GGCTATTATCACTTTTGTAACAA-3'	R01: 5'-TCATTGCGTGAGGGC-3'
2	F02: 5'-CAGATCTTTTATTTAAGTGGGATA-3'	F02: 5'-ATAAATACTGGCCTTTGGG-3'
	R02: 5'-ATGATCTAACAACTAGCATTTAATA-3'	R02: 5'-CCTGTATTGACATTAAGTCATATA-3'
3	F03: 5'-CCTTCAA AATGAAAACCAGTTT-3'	F03: 5'-TCTGTTTATTCTTGCTAAATGTTTA-3'
	R03: 5'-ATTTAGATCAGCAGAATAGG-3'	R03: 5'-CAATCAAGAGCTAACTTGA-3'
4	F04: 5'-ATTAGTTTTTCTCTGGAATATTGA-3'	F04: 5'-ACATTTGGGGCTGGTG-3'
	R04: 5'-GACTAGACATGCTTCTC-3'	R04: 5'-CTTCGCTCAGTCATTTAAC-3'
5	F05: 5'-TAATGGACCACCTGTTTCT-3'	F05: 5'-TTAGAGAACATTCCCATTATGTT-3'
	R05: 5'-AAAGGCAATGTTAAATGCTGAA-3'	R05: 5'-TTAGGGCAGGGGACAA-3'
6	F06: 5'-GATGCCAACCCATGTT-3'	F06: 5'-ACTTCAGGGAATAGTTGTAC-3'
	R06: 5'-AAAAACAATCTAAAATAGTGTGCT-3'	R06: 5'-CTAGTCTTTGATAAATTGAAAAG-3'

7	F07: 5'-ATCCTGAGATATTTATTGAACTATT-3'	F07: 5'-AGTCACATGGGTAAAGACAA-3'
	R07: 5'-TCAAGAGATTAATCTTGGGC-3'	R07: 5'-GCTTATACTTTTTATGCCTCTTAT-3'
8	F08: 5'-TAATTGTAAAAAGTGAAGCCATG-3'	F08: 5'-TGACATTGCATTCATGAAGG-3'
	R08: 5'-ATAGGTGAAATAGCTCTAGC-3'	R08: 5'-GCAGATTTACATAATCCCTAATA-3'
9	F09: 5'-GAGGGCTGGTTTTCTATAA-3'	F09: 5'-TGCTAGAGCTATTTACC-3'
	R09: 5'-GGCACTTTTCTACCTAAGTA-3'	R09: 5'-CTCTGATTTATTTAGGAAAGAATG-3'
10	F10: 5'-AGTCATTGTGTCCTTAATAAC-3'	F10: 5'-CAGCCACTAAAGTACTACAT-3'
	R10: 5'-CAGTTGGAATAAAAGGACAG-3'	R10: 5'-ATGAGTATTTACAGTTTCACTTAG-3'
11	F11: 5'-ATTTTTTTTATAAGGAAGTGGACC-3'	F01: 5'-CAGAGAAATTCATTCTTCTTTGT-3'
	R11: 5'-TCTACTGCTTATGTTCTTTAC-3'	R01: 5'-ATCCTTGATTAACATAGCCATT-3'
12	F12: 5'-GTATGGTAATACAGTCTCAA-3'	F12: 5'-CAAATAAAAGGTCATGTTAAGATAA-3'
	R12: 5'-TAAACAATGATGAGGCCTTT-3'	R12: 5'-TTCTAGGGCAATTATTAGAGAATA-3'
13	F13: 5'-CCGAAAAACAGAAACAGATT-3'	F13: 5'-AAAAAACAGAAACAGATTAGAAGC-3'
	R13: 5'-GTTACTTCATCTTTTTAGAAGTAG-3'	R13: 5'-GAATTAAGTGTAGCAACTTAAT-3'
14	F14: 5'-GAGCTTTACCTTCTTATGAC-3'	F14: 5'-TTAATCATGGTGCTAAATGG-3'
	R14: 5'-CCAAGAAAATAAGAATTATCTCCA-3'	R14: 5'-CATTTATTTAGATCATCTTCTCTG-3'
15	F15: 5'-AGCCTTTCTCTAAGTAGGTA-3'	F15: 5'-TAAAATCTAAAGAACAAATGTGGA-3'
	R15: 5'-ATTATCTTTTTAGTAATCTGGGTAT-3'	R15: 5'-ATGTACTGAAAGATGAAACTAAGT-3'
16	F16: 5'-GCTGTTAAGGAGTGATTCAT-3'	F16: 5'-CATAATTAGCTGGGCAAATAT-3'
	R16: 5'-AGTATGGAACAAGTTTAATCG-3'	R16: 5'-TCGATAATATCTCCTTAAATAAAGT-3'
17	F17: 5'-GATAGGCAATGTAGCTCTT-3'	F17: 5'-TTTGTTAGTGTCTAAAGTCTGTTT-3'
	R17: 5'-GCAGGAGAATGGCATGA-3'	R17: 5'-CGGGAGTCAGAGCTT-3'
18	F18: 5'-GACCTTAAAGGGCAGG-3'	F18: 5'-TGTCAGATGTGGCCC-3'
	R18: 5'-GGTCTCAGAGACAGTG-3'	R18: 5'-AGTTGTTACCTTTAGAACAACAAAT-3'
19	F19: 5'-TGGAAGATAGATTGCCAG-3'	F19: 5'-GCTGTTTCAGTTCTAAACTTG-3'
	R19: 5'-ACTCAAGCTTCAAGTTAATATATTT-3'	R19: 5'-AAACTCAGATTCTAAGAAAATAAGA-3'
20	F20: 5'-AGACTCTCTCTGCTCTT-3'	F20: 5'-TAAGCGGTGACTACATTGAA-3'
	R20: 5'-TTCCACAAAGGCCCGA-3'	R20: 5'-GCAATTCTGTAGTGTCTCT-3'
21	F21: 5'-CTCAATAAGATAAAAGCAGATATTA-3'	F21: 5'-GGTCCTGAGTAGATTG-3'
	R21: 5'-CTCTGTCACTGTAACACTAC-3'	R21: 5'-CTGAATAAGTACTTTTTCTGCAA-3'
22	F22: 5'-AGAACACTCTCAGCCC-3'	F22: 5'-CTTCAACTTTTATAATTTTCAGGGT-3'
	R22: 5'-CTAGATCTCACATTTAAATGCA-3'	R22: 5'-CAATTTGATTTTTCCCTATTTGTT-3'
23	F23: 5'-CTTTCACCCTGTCAGG-3'	F23: 5'-CCAGTATCCCAGTGG-3'
	R23: 5'-CAAATGTACCGAGGGC-3'	R23: 5'-GGCTTATTTCTATTACTGAAC-3'
24	F24: 5'-GATAATATCTATCAACATTTTTGCT-3'	F24: 5'-TTCAAATTCTTATTAATGTTTGGG-3'
	R24: 5'-TTTTTCTCATGAGGTTTATTGGTA-3'	R24: 5'-GTATTATTTGAAAGAAAGCCTAAA-3'
25	F25: 5'-TAGTGTAGCACCTCTG-3'	F25: 5'-CTGTTAATGTTTGTATTTGTAAAC-3'
	R25: 5'-ACAGTCTGGAGAATAAGTAG-3'	R25: 5'-ATAGTTCTCTCACAGGTTTTAT-3'
26	F26: 5'-TACCCTGGTTACTTGAG-3'	F26: 5'-GAGTAATAACAGTAAACTCTTAAT-3'
	R26: 5'-GATCCATTTTTAAATAAGTAATGCA-3'	R26: 5'-ACTCATAAAGAATAGGAAAAGG-3'
27	F27: 5'-CAGAGTATGCAGCAGCT-3'	F27: 5'-GGCAGTGTGCTAGTCA-3'
	R27: 5'-AGTTCCATTCTCCTTTGG-3'	R27: 5'-TTGACAAACCTGAAATGTATACTT-3'
28	F28: 5'-CTGTAGATTTTTGTGTATTTAAACT-3'	F28: 5'-TGAGTATTTCCCTCCTT-3'
	R28: 5'-TGCAGCCACCTTCTTAT-3'	R28: 5'-TGATTCTTTGTCTCTAATTTCT-3'
29	F29: 5'-AAAAGTCTCCTGTGTTTCATC-3'	F29: 5'-ACACCAGGTCCCTTC-3'
	R29: 5'-GGCACACATAGTCTTAGAA-3'	R29: 5'-TCAAACAGCTCACAGG-3'

Confirmation of <i>WDR11</i> missense mutations			
Mutation	Location	Forward primer	Reverse primer
R395W	Exon 8	5'- CCATGCTTGACATTGCATTCATGAAGGA GTGATGC-3'	5'- CTTAAAATCTGCATTACGTATTAGGAAT TCAAAGTATTAAGTCCAG-3'
A435T	Exon 10	5'- GATGCAATTTAAGTCCTATAGCCAGGTA CAGAGTC-3'	5'- GCATTCCAATTCAATATAAGACAGTTGG AATAAAAGGACAG-3'
R448Q	Exon 10	5'- GTGTCCTTAATAACTACAGCCACTAAAG TACTACATTAAC-3'	5'- GAGTATTTACAGTTTCACTTAGTCTCTAT CTCCTTTTAAAGC-3'
H690Q	Exon 16	5'- CATGTTTACCCATAATTAGCTGGGCAA TATTAGTTTCAGAC-3'	5'- GCTGGATAACTTTGGGAAAGTTACTTAT TCTCCCACAG-3'
F1150L	Exon 28	5'- GCTTGTGTTTGGAGATGTGAGTATTCTT CCCTC-3'	5'-CTTCCTGCCTCCCTCCATCCTAGTC-3'

Long Range PCR for der(12) junction fragment	
Forward primer	Primer sequence
12q13.2-10kb-4208f	5'-CTCATATGTTTTCAAGAGGATCTTCTCAATATAACAATGTGTAGCTTAT-3'
12q13.2-10kb-4562f	5'-CTGCAACGATGAACAGATCAGAATAGGCGACATGC-3'
12q13.2-10kb-5338f	5'-GTTATATTGAAGCGTGAAATATTGATTACAATTTTGAAGCATAAAATATATCG-3'
12q13.2-10kb-5961f	5'-ATATTCTGTGCTTCAAAATAGTTGTAATCAATATTTGACACTTCAATATAAC-3'
12q13.2-10kb-7351f	5'-GTAATGACTTACTCACACCCTGCCCACTCTAAC-3'
Reverse primer	Primer sequence
10q26-10kb-3497r	5'-CAAAAGCTGGAGTGGCATTACTACTAAGTAACAGGTAGAACAAAAAC-3'
10q26-10kb-5710r	5'-CCCTGGTTTAAATGTTTCTGCATTGCCTAAACAATGAGATTTAAAC-3'
10q26-10kb-5750r	5'-CTAGAGTTGTCTTCTGGATAGCCAGAACTGATCACC-3'
Long Range PCR for der(10) junction fragment	
Forward primer	Primer sequence
10q26-10kb-1851f	5'-CTGTAGGTCAAGCCTTTAATACCCTCCTATTAGAGGG-3'
10q26-10kb-2367f	5'-GCCAAAAAATGTTAGCCATTATTACAGAGTATTCAGGAGTTCAG-3'
10q26-10kb-2801f	5'-ATCTACTTTATAAGGTCATGATGAACACTAGATTTAGCATGTGGTAAG-3'
10q26-10kb-2851f	5'-CCATGAAAATTTGCAGTATTATTATAGCAGTTGGAGGCCATG-3'
10q26-10kb-2914f	5'-GTATTGTGAAGAGGTCAAAGCCAGGCTGCAGAG-3'

Reverse primer	Primer sequence
12q13.2-10kb-7630r	5'-CTGGACATTGGGTATCAAATTGCTGATACCCTGATCTG-3'
12q13.2-10kb-8180r	5'-CGATCTCAGCTCACTGCAACCTCTGTCTCC-3'
12q13.2-10kb-8836r	5'-GCTATGTAAGCAGTTATCAGCAATTTGACCCCAGTC-3'
12q13.2-10kb-9664r	5'-CTGTATACTTTCTGTCTTCATTTTATATCATCATTCCATGTTTTGAAAAAAGTT-3'
12q13.2-10kb-9890r	5'-GCTCAATTATAAAAGGAAAGTTGATTCTAGATAGGAGCAGGCCATTG-3'

Zebrafish <i>wdr11</i> gene (XM_682139.3) isolation by RT-PCR	
Forward primer	Reverse primer
5'-CGGGATCCGCCAAAACATTCCACCTTTTCAT-3'	5'-CACCATGGAGTCCATATCCTGAATGGGTTTG-3'

RT-qPCR of <i>WDR11</i> in balanced translocation patient	
Primer	Primer sequence
WDR11-RT-3117f	5'-GCAACCCACTGGATATATGCTATGACGTGCTCTG-3'
WDR11-RT-3213r	5'-GACCGTTTCACTTCCTGTAGATTAACCCTTTCTAGCTG-3'
WDR11-RT-3221f	5'-GATCATACAAGGAAATGTACAGACCAGCTACTGCTCTTG-3'
WDR11-RT-3370r	5'-GCCTGACGAGGTGACAGTAGTGACTAACAGG-3'

## References

1. Voegtli, W.C., Madrona, A.Y. & Wilson, D.K. The structure of Aip1p, a WD repeat protein that regulates Cofilin-mediated actin depolymerization. *J Biol Chem* **278**, 34373-9 (2003).
2. Auld, D.S. Zinc coordination sphere in biochemical zinc sites. *Biometals* **14**, 271-313 (2001).

## 12.4 Paper IV

**Hyung-Goo Kim\***, Hyun Taek Kim\*\*, Natalia Leach\*\*, Fei Lan\*\*, Reinhard Ullmann, Asli Silahdaroglu, Ingo Kurth, Anja Nowka, Ihn Sik Seong, Yiping Shen, Michael E. Talkowski, Douglas Ruderfer, Ji-Hyun Lee, Caron Glotzbach, Kyungsoo Ha, Susanne Kjærgaard, Alex V. Levin, Bernd F. Romeike, Tjitske Kleefstra, Oliver Bartsch, Sarah H. Elsea, Ethylin W. Jabs, Marcy E. MacDonald, David J. Harris, Bradley J. Quade, Hans-Hilger Ropers, Lisa G. Shaffer, Kerstin Kutsche, Lawrence C. Layman, Niels Tommerup, Vera M. Kalscheuer, Yang Shi, Cynthia C. Morton, Cheol-Hee Kim, James F. Gusella

\*Corresponding author

\*\* These authors contributed equally to this work

***Translocations disrupting PHF21A in the Potocki-Shaffer syndrome region are associated with intellectual disability and craniofacial anomalies***

In press in July 2012 issue of the American Journal of Human Genetics

# Translocations Disrupting *PHF21A* in the Potocki-Shaffer-Syndrome Region Are Associated with Intellectual Disability and Craniofacial Anomalies

Hyung-Goo Kim,<sup>1,2,3,\*</sup> Hyun-Taek Kim,<sup>4,27</sup> Natalia T. Leach,<sup>5,27</sup> Fei Lan,<sup>6,7,27</sup> Reinhard Ullmann,<sup>3</sup> Asli Silaharoglu,<sup>8</sup> Ingo Kurth,<sup>9,10</sup> Anja Nowka,<sup>10</sup> Ihn Sik Seong,<sup>1,11</sup> Yiping Shen,<sup>1,6,12</sup> Michael E. Talkowski,<sup>1,11,13</sup> Douglas Ruderfer,<sup>1,13</sup> Ji-Hyun Lee,<sup>1</sup> Caron Glotzbach,<sup>14</sup> Kyungsoo Ha,<sup>15</sup> Susanne Kjaergaard,<sup>16</sup> Alex V. Levin,<sup>17</sup> Bernd F. Romeike,<sup>18</sup> Tjitske Kleefstra,<sup>19</sup> Oliver Bartsch,<sup>20</sup> Sarah H. Elsea,<sup>21</sup> Ethylin Wang Jabs,<sup>22,23</sup> Marcy E. MacDonald,<sup>1,11,13</sup> David J. Harris,<sup>24</sup> Bradley J. Quade,<sup>25</sup> Hans-Hilger Ropers,<sup>3</sup> Lisa G. Shaffer,<sup>14</sup> Kerstin Kutsche,<sup>10</sup> Lawrence C. Layman,<sup>2</sup> Niels Tommerup,<sup>8</sup> Vera M. Kalscheuer,<sup>3</sup> Yang Shi,<sup>6</sup> Cynthia C. Morton,<sup>5,13,25</sup> Cheol-Hee Kim,<sup>4</sup> and James F. Gusella<sup>1,13,26</sup>

Potocki-Shaffer syndrome (PSS) is a contiguous gene disorder due to the interstitial deletion of band p11.2 of chromosome 11 and is characterized by multiple exostoses, parietal foramina, intellectual disability (ID), and craniofacial anomalies (CFAs). Despite the identification of individual genes responsible for multiple exostoses and parietal foramina in PSS, the identity of the gene(s) associated with the ID and CFA phenotypes has remained elusive. Through characterization of independent subjects with balanced translocations and supportive comparative deletion mapping of PSS subjects, we have uncovered evidence that the ID and CFA phenotypes are both caused by haploinsufficiency of a single gene, *PHF21A*, at 11p11.2. *PHF21A* encodes a plant homeodomain finger protein whose murine and zebrafish orthologs are both expressed in a manner consistent with a function in neurofacial and craniofacial development, and suppression of the latter led to both craniofacial abnormalities and neuronal apoptosis. Along with lysine-specific demethylase 1 (LSD1), *PHF21A*, also known as BHC80, is a component of the BRAF-histone deacetylase complex that represses target-gene transcription. In lymphoblastoid cell lines from two translocation subjects in whom *PHF21A* was directly disrupted by the respective breakpoints, we observed derepression of the neuronal gene *SCN3A* and reduced LSD1 occupancy at the *SCN3A* promoter, supporting a direct functional consequence of *PHF21A* haploinsufficiency on transcriptional regulation. Our finding that disruption of *PHF21A* by translocations in the PSS region is associated with ID adds to the growing list of ID-associated genes that emphasize the critical role of transcriptional regulation and chromatin remodeling in normal brain development and cognitive function.

## Introduction

In many regions of the genome, microdeletions<sup>1,2</sup> or balanced translocations<sup>3–6</sup> are associated with phenotypic abnormalities and are presumably caused by haploinsufficiency of the various genes involved. Potocki-Shaffer syndrome (PSS [MIM 601224]) is a rare contiguous gene-deletion syndrome caused by heterozygous interstitial microdeletions of chromosomal region 11p11-p12 and is

characterized by developmental defects that include intellectual disability (ID), craniofacial anomalies (CFAs), multiple exostoses (MIM 133701), and parietal foramina (MIM 609597).<sup>7,8</sup> Genes responsible for the latter two phenotypes in this chromosomal region have been identified: Deletion of *EXT2* (MIM 608210) causes multiple exostoses,<sup>9</sup> and deletion of *ALX4* (MIM 605420) causes parietal foramina.<sup>10,11</sup> However, the cause of the ID and abnormal craniofacial development has remained uncertain.

<sup>1</sup>Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA 02114, USA; <sup>2</sup>Department of Obstetrics & Gynecology, Institute of Molecular Medicine and Genetics, Georgia Health Sciences University, Augusta, GA 30912, USA; <sup>3</sup>Department of Human Molecular Genetics, Max Planck Institute for Molecular Genetics, Ihnestraße 63–73, 14195 Berlin, Germany; <sup>4</sup>Department of Biology, Chungnam National University, Daejeon 305-764, Korea; <sup>5</sup>Department of Obstetrics, Gynecology, and Reproductive Biology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA; <sup>6</sup>Department of Pathology, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA; <sup>7</sup>Constellation Pharmaceuticals, Department of Biology, 215 First Street, Cambridge, MA 02142, USA; <sup>8</sup>WilhelmJohannsen Centre for Functional Genome Research, Department of Cellular and Molecular Medicine, University of Copenhagen, DK-2200 Copenhagen, Denmark; <sup>9</sup>Jena University Hospital, Institute of Human Genetics, 07743 Jena, Germany; <sup>10</sup>Institut für Humangenetik, Universitätsklinikum Hamburg-Eppendorf, 20249 Hamburg, Germany; <sup>11</sup>Department of Neurology, Massachusetts General Hospital, Boston, MA 02114, USA; <sup>12</sup>Shanghai Children's Medical Center, Shanghai Jiaotong University School of Medicine, Shanghai, 200127 China; <sup>13</sup>Program in Medical and Population Genetics, Cambridge, MA 02114, USA; <sup>14</sup>Signature Genomic Laboratories, PerkinElmer, Spokane, WA 99207, USA; <sup>15</sup>Georgia Health Sciences University Cancer Center, Augusta, GA 30912, USA; <sup>16</sup>Department of Clinical Genetics, University Hospital Rigshospitalet, DK-2100 Copenhagen, Denmark; <sup>17</sup>Pediatric Ophthalmology and Ocular Genetics, Wills Eye Institute, Thomas Jefferson University, Philadelphia, PA 19107, USA; <sup>18</sup>Department of Neuropathology, Friedrich Schiller University, 07747 Jena, Germany; <sup>19</sup>Department of Human Genetics, Radboud University Nijmegen Medical Centre, 6500 Nijmegen, The Netherlands; <sup>20</sup>Institute of Human Genetics, University Medical Center of the Johannes Gutenberg-University Mainz, 55101 Mainz, Germany; <sup>21</sup>Departments of Pediatrics and Human & Molecular Genetics, Virginia Commonwealth University School of Medicine, Richmond, VA 23298, USA; <sup>22</sup>Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY 10029, USA; <sup>23</sup>Institute of Genetic Medicine, The Johns Hopkins University, Baltimore, MD 21287, USA; <sup>24</sup>Division of Genetics, Children's Hospital Boston and Harvard Medical School, Boston, MA 02115, USA; <sup>25</sup>Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA; <sup>26</sup>Department of Genetics, Harvard Medical School, Boston, MA 02114, USA

<sup>27</sup>These authors contributed equally to this work

\*Correspondence: [hkim@chgr.mgh.harvard.edu](mailto:hkim@chgr.mgh.harvard.edu)

<http://dx.doi.org/10.1016/j.ajhg.2012.05.005>. ©2012 by The American Society of Human Genetics. All rights reserved.

Through identification of two independent subjects with balanced translocations and support from a third published translocation subject,<sup>12,13</sup> we have uncovered evidence that haploinsufficiency of a single gene, *PHF21A* (also known as *BHC80* [MIM 608325]), at 11p11.2 is associated with ID and CFA phenotypes. This evidence was complemented by comparative deletion mapping of PSS subjects with diverse phenotypes that positioned *PHF21A* within the critical region associated with ID and CFAs, functional in vitro analysis of *PHF21A* in cells from translocation subjects, and generation and rescue of zebrafish phenotypes through the suppression of *phf21a* with morpholino oligonucleotides (MOs). *PHF21A* specifically binds unmethylated histone H3 lysine 4 (H3K4me0) and participates in the lysine-specific demethylase 1 (LSD1 [MIM 609132]) demethylase complex, implicating it as a regulatory protein in histone-methylation dynamics<sup>14</sup> and suggesting that disruption of this process might underlie the ID and CFA phenotypes in these translocation individuals.

## Subjects and Methods

### Human Subjects

DGAP012 and MCN1762 (Table 1) were ascertained through the Developmental Genome Anatomy Project (DGAP) and Mendelian Cytogenetic Network, respectively, and blood samples were obtained for initiation of a lymphoblastoid cell line.<sup>17</sup> PSS subjects PSS02, PSS08, PSS10, and PSS-Romeike were previously described.<sup>7,15,18</sup> Where possible, all subjects were tested for copy-number variants (CNVs) by array comparative genomic hybridization (CGH) with the use of Agilent (Santa Clara, CA) 244K arrays as described (Table 2).<sup>19</sup> DGAP012 displayed one 156 kb deletion CNV (chr13: 48,430,969–48,587,500; hg18) that is located at 13q14.2 and has not been reported previously and that involves *FNDC3A*, which is expressed in spermatids and Leydig cells and produces male sterility when homozygously inactivated.<sup>20</sup> All human studies were performed under informed-consent protocols approved by the Partners HealthCare System Human Research Committee.

#### DGAP012

At birth, this white male was small but normally proportioned and suffered from supraventricular tachycardia, hyperbilirubinemia, and hypoglycemia, all of which resolved. Chromosome analysis revealed an apparently balanced chromosomal translocation with the karyotype 46,XY,t(11;19)(p11.2;p13.3)dn, which was revised by molecular analysis here to 46,XY,t(11;19)(p11.2;p13.2)dn. The family history was unremarkable, and parental chromosomes were normal. At 1 year of age, the subject displayed bifrontal biparietal atrophy on a computed tomography (CT) scan, and at 15 months of age, he showed significant global developmental delay, digitalized thumbs, brachycephaly, microcephaly (a head circumference of 46 cm; tenth percentile), a small downturned mouth, mild midfacial hypoplasia, a flat midface, a narrow nasal bridge, a very small nose, large ears, bilateral epiblepharon (an extra skin fold was medially under each lower lid) without trichiasis, small hands and feet, and an absence of emotional expression. He also displayed hand flapping and had feeding problems prior to the age of 3 years.

A magnetic resonance image (MRI) at the age of 2 years revealed prominent cerebrospinal fluid, which might represent volume loss. At the age of 5 years, DGAP012 displayed hypotonia and dystonic movement and could not walk and showed significant nearsightedness and astigmatism (−3.00 +3.00 × 90 in both eyes), as well as some diffuse pigmentary mottling suggesting a possible retinal dystrophy.

#### MCN1762

This white female, aged 42 years, was delivered at term after an uneventful pregnancy but was lethargic and had feeding difficulties. At age 3.5 years, her gross motor development was normal, but her speech development was delayed, and she displayed hyperactivity, poor concentration, and mild myopia (−2.00 [right eye], −2.50 [left eye]). A verbal intelligence quotient (IQ) test (Terman-Merrill) indicated a delay of 1 year, and a nonverbal IQ test (Leiter) was normal. The following dysmorphic features were noted: brachycephaly, microcephaly, a long narrow nose, mild midfacial hypoplasia, a downturned mouth, thin lips, and prominent ear lobes and, in childhood only, downslanting palpebral fissures and epicanthal folds. GTG-banded karyotyping revealed an apparently balanced reciprocal translocation, 46,XX,t(1;11)(p13;p11)dn, which was revised by molecular analysis to 46,XX,t(1;11)(p21.1p11.2)dn. Her parents and two siblings are healthy, and parental karyotypes are normal. MCN1762 attended a school for special needs and is able to read. Her linear growth (to 173 cm) and pubertal development were normal, and she has developed truncal obesity (see Figure 4C). She now has mild ID but manages to live independently and have a sheltered part-time job.

#### GM03316

This female Venezuelan subject was 3 years old when her blood was submitted to the National Institute of General Medical Sciences Human Genetic Cell Repository at the Coriell Institute in January of 1978 (prior to identification of PSS) and showed an apparently balanced translocation: t(X;11)(q11.1;p11.2)dn. The clinical symptoms were listed as ID (quantitative intelligence [Gq] = 60) with a strikingly unusual dysmorphology syndrome including epicanthus, hypertelorism, oblique palpebral fissures, trigonocephaly, and micrognathia. At 5 years of age, her vocabulary was progressing well, and she had a good memory, but her Gq corresponded to that of a 3-year-old girl. Her principal problem was an inability to concentrate. She could feed herself when she wished and had gained control of her sphincters both day and night by age 3.5 years. She was shy and easily frightened and was clumsy with both hands and legs.

#### GC14361

This 2.25-year-old male from Bangladesh had a history of static encephalopathy and developmental delay, which were first noted when he was 6 months old. He displayed microcephaly, short stature, a small phallus, a unilateral absent testis, and dysmorphic features, including a short forehead, prominent biparietal foramina, a midline parietal cortical defect, a flat midface, a flat occiput, sensorineural hearing loss, epicanthal folds, protuberant ears, a bulbous nasal tip that continued below the columella, a depressed nasal root, a small mouth and small chin (micrognathia), hypotonia, a slight pectus excavatum, recurrent otitis media, and slender fingers.

### Breakpoint Mapping, Cloning, and Characterization

In brief, bacterial artificial chromosomes (BACs), followed by fosmid or cosmid probes chosen on the basis of the human genome map, were used for fluorescence in situ hybridization

**Table 1. Phenotypic Comparison of Individuals with Balanced Translocations and Their Breakpoints**

	Individual			Typical PSS phenotype <sup>d</sup>
	DGAP012 <sup>a</sup>	MCN1762 <sup>b</sup>	GILLE <sup>c</sup>	
Intellectual disability	+	+	+	+
Facial dysmorphism	+	+	+	+
Narrow nose	+	+	+	+
Downturned mouth	+	+ <sup>e</sup>	ND	+
Large ears	+	–	–	+
Brachycephaly	+	+	mild craniofacial asymmetry and thin corpus callosum	+
Microcephaly	+	+	hypoplasia of inferior cerebellar vermis	+
Myopia	+	+	ND	+
Strabismus	+	–	nystagmus	+
Heart defect	+	–	ND	+
Hypotonia	+	+	+	+
Small hands and feet	+	–	ND	+
Multiple exostoses	–	–	–	+
Parietal foramina	–	–	–	+
Tapering fingers	–	–	ND	+
Digitalized thumbs	+	–	ND	–
Retinal dystrophy	+	–	iris hypoplasia, superior atypical coloboma, and foveal hypoplasia	–
Revised karyotype	t(11;19)(p11.2;p13.2)dn	t(1;11)(p21.1;p11.2)dn	t(X;11)(p22.2;p11.2)dn <sup>f</sup>	
Location of breakpoint in <i>PHF21A</i>	intron 14	intron 5	breakpoint mapped within <i>PHF21A</i>	
Disruption of <i>PHF21A</i> (confirmation method)	truncated (FISH and breakpoint cloning)	truncated (FISH and breakpoint cloning)	disrupted (FISH); breakpoint-spanning BAC clone at 11p11.2 is RP11-618K13; fine mapped with 10 kb PCR product	
Breakpoint at 11p11.2 (Human Genome Browser hg18)	chr11: 45,921,645–45,921,646	chr11: 46,020,273–46,020,274	chr11: 45,907,446–46,099,561	

The following abbreviations are used: PSS, Potocki-Shaffer syndrome; ND, not determined; and FISH, fluorescence in situ hybridization.

<sup>a</sup>DGAP012 was assessed at age 15 months.

<sup>b</sup>MCN1762 was assessed at ages 3.5 years and 42 years.

<sup>c</sup>GILLE was assessed only at ages 4 and 8 months when her phenotype was thought to suggest Gillespie syndrome.<sup>12,13</sup>

<sup>d</sup>Typical PSS phenotypes from reported PSS subjects,<sup>15</sup> including patients 1, 2, and 3 in Wuyts et al.<sup>16</sup>

<sup>e</sup>Not visible in Figure 4 as a result of laughing expression.

<sup>f</sup>The initial karyotype was reported as t(X;11)(p22.32;p12),<sup>12</sup> which was revised as t(X;11)(p22.3;p12). However, on the basis of the locations of two genes, *ARHGAP6* and *PHF21A*, reported as disrupted,<sup>13</sup> this karyotype has been further revised to t(X;11)(p22.2;p11.2).

(FISH) of metaphase chromosomes as previously described.<sup>6</sup> For breakpoint fine mapping by array painting, metaphase chromosomes were flow sorted,<sup>21,22</sup> and DNA was extracted, amplified, labeled, and hybridized to a custom array (Agilent, Santa Clara, CA) with an average spacing of three unique probes per 1 kb as described previously.<sup>4</sup> The array design has been made public at the eArray website (name: Translok4; design number: 022990). Standard DNA blotting with PCR-generated probes was used for delineating the breakpoints, which were then isolated with suppression PCR and sequenced with an ABI 3730XL DNA Analyzer (Applied Biosystems, Foster City, CA). This suppression-PCR method exploits a specially designed

adaptor to walk in unknown genomic DNA regions from known adjacent sequences.<sup>23</sup> Specific steps for breakpoint mapping and cloning in each case are the following.

*DGAP012 with t(11;19)(p11.2;p13.2)*

On chromosome 11, BAC clones RP11-618K13, CTD-2254P23, and RP11-142O14 all span the breakpoint (Figure S1A and Table S1, available online) that lies in the 14 kb between BAC clones CTD-2580L7 and CTD-2353H9. Four PCR probes (D012-A, D012-B, D012-C, and D012-D; see Table S3 for primers) detected junction fragments with various enzymes (e.g., Figure S1C). This Southern analysis narrowed down the breakpoint to 729 bp (Figure S1A).

**Table 2. PSS Subjects with Deletions Tested for Inclusion of *PHF21A***

Individual ID	Phenotype	Karyotype	Deletion Size	Deletion of <i>PHF21A</i> (Confirmation Method)	Deletion by Array CGH (Human Genome Browser hg18)
PSS02	Potocki-Shaffer syndrome with ID and CFAs	del(11)(p11.2p12)	5.6 Mb	deleted (FISH and CGH)	chr11: 40,434,231–46,041,804
PSS08	Potocki-Shaffer syndrome with ID and CFAs	del(11)(p11.2p13)	15.2 Mb	deleted (FISH and CGH)	chr11: 31,009,420–46,204,605
PSS10	Potocki-Shaffer syndrome with ID and CFAs	del(11)(p11.2p12)	N/A	deleted (FISH)	DNA unavailable because primary cells did not survive
PSS Romeike <sup>18</sup>	Potocki-Shaffer syndrome with ID and CFAs	del(11)(p11.12p12)	11.9 Mb	deleted (CGH)	chr11: 38,824,655–50,638,770
GM03316	ID and CFAs	t(X;11)(q11.1;p11.2)dn del(11)(p11.2p12)	3.6 Mb	deleted (FISH and CGH)	chr11: 43,125,403–46,706,549
GC14361	Potocki-Shaffer syndrome with ID and CFAs	del(11)(p11.12p12)	13.8 Mb	deleted (FISH and CGH)	chr11: 33,430,130–47,276,580

Deletions of all subjects except PSS10 (for whom a microarray could not be performed) are also depicted as blue bars in Figure 3, and *PHF21A* FISH results are shown in the Figure 3 inset for all subjects except PSS Romeike. The following abbreviations are used: ID, intellectual disability, CFAs, craniofacial anomalies; FISH, fluorescence in situ hybridization; and CGH, comparative genomic hybridization.

We reported previously that FISH with BAC CTD-3193O13 spanned the chromosome 19 breakpoint<sup>24</sup> (Figure S1B and Table S1). Additional FISH with five cosmids and three restriction fragments narrowed down the breakpoint region to 11 kb between the telomeric end of a 30 kb restriction fragment and the telomeric end of RP11-585H16 (Figure S1B). Three PCR probes (D012-E, D012-F, and D012-G; primers in Table S3) detected junction fragments (e.g., Figure S1D), narrowing the breakpoint to 3,441 bp at 19p13.2 (Figure S1B).

For suppression PCR, adaptors 1 and 2 were ligated to ScaI-digested DGAP012 genomic DNA (Table S3).<sup>23</sup> The ligated DNA was diluted 1:10 with pure water and used as a template for suppression nested PCR (two-step PCR in which first boost PCR is followed by nested PCR) with the chromosome 19 primers (Table S3) designed from the 3,441 bp deduced breakpoint region (Figure S1B) and adaptor primers (Table S3). This suppression PCR generated products composed of sequences from both chromosomes 11 and 19 and thereby identified an unknown chromosome 11 sequence that confirmed the 729 bp breakpoint region at 11p11.2 by Southern analysis (Figure S1A).

The der(19) junction fragment was amplified by nested PCR with the following primer pairs: chr11 forward 1 + chr19 reverse 1 for boost PCR and chr11 forward 2 + chr19 reverse 2 for nested PCR. The der(11) junction fragment was amplified by nested PCR with the following primer pairs: chr19 forward 1 + chr11 reverse 1 for boost PCR and chr19 forward 2 + chr11 reverse 2 for nested PCR (Table S3).

#### *MCN1762 with t(1;11)(p21.1;p11.2)*

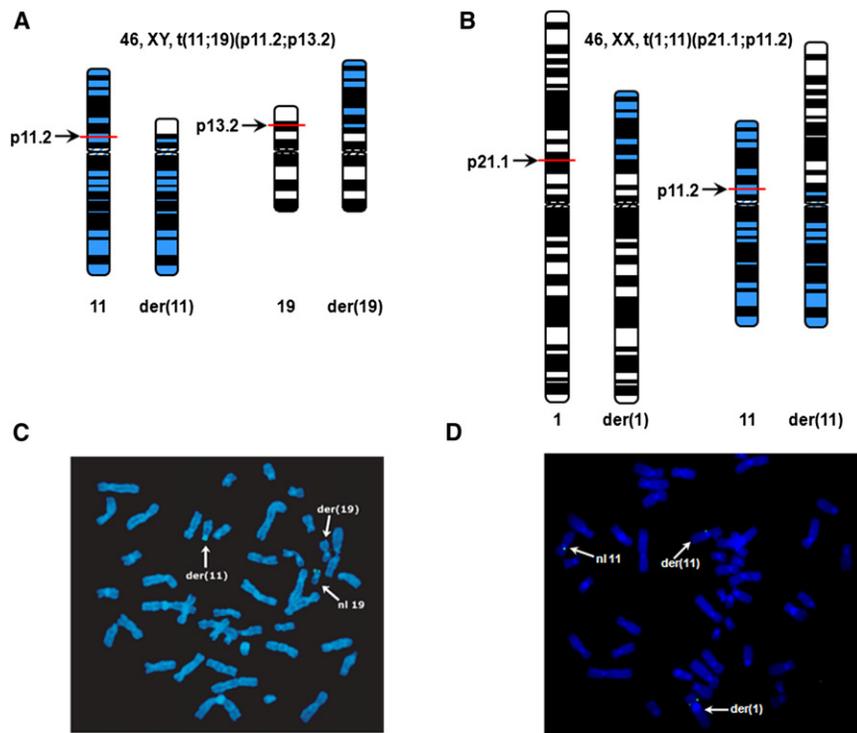
We first performed FISH with BAC clones RP11-992G23 and RP11-425L10 (which hybridize distally and proximally to *PHF21A*, respectively) in 11p11.2 to confirm localization of the translocation to the *PHF21A* region. Then, guided by array painting, we performed FISH with additional BACs and subsequently with fosmids and identified RP11-177H1 as a BAC spanning the breakpoint lying within the 37 kb segment covered by fosmid G248P88722D5 (Figure 1D and Table S2). From the array-painting results, breakpoint regions were refined to 11.4 kb in 1p21.1 and 93 bp in 11p11.2, resulting in the revision of the previously designated karyotype to 46,XY,t(1;11)(p21.1;p11.2)dn.

We then used PCR to amplify junction fragments by using primers in the refined regions of chromosomes 1 and 11 by array painting.

The junction fragment of ~0.8 kb from der(1) was amplified by normal PCR from one primer pair of chr11 forward and chr1 reverse and was confirmed by sequencing and subsequent BLAST searching. The ~650 bp junction fragment of der(11) was amplified by primer set chr1 forward and chr11 reverse (Table S3). Partial sequences from chromosomes 1 and 11 of this junction fragment were confirmed by sequencing and BLAST searching.

### Mutation Screening of *PHF21A* in Various Subjects with Partially Overlapping Phenotypes

We performed mutation analysis by direct sequencing of exonic PCR products to exclude an alteration in the nontranslocated allele of *PHF21A* in DGAP012 and MCN1762, as well as in 200 additional subjects with various phenotypes: 25 with features of Saethre-Chotzen syndrome (MIM 101400) and without mutations in *TWIST* (MIM 601622) or hotspot regions of *FGFR1* (MIM 136350), *FGFR2* (MIM 176943), and *FGFR3* (MIM 134934); 14 with undiagnosed syndromes and isolated features of Rubinstein-Taybi syndrome (MIM 180849) and without *CREBBP* (MIM 600140) mutations;<sup>25</sup> 16 with micropenis or hypogonadotropic hypogonadism and ID and no mutations in known genes involved in Kallmann syndrome or hypogonadotropic hypogonadism; five with metopic craniosynostosis of unknown etiology; 19 with typical or atypical Smith-Magenis syndrome (MIM 182290) and no mutations in *RAI1* (MIM 607642); 25 with Kleefstra syndrome (MIM 610203) and no mutations in *EHMT1* (MIM 607001);<sup>2</sup> and 96 with Cornelia de Lange syndrome (MIM 122470, 610759, and 300590) and no mutations in *NIPBL* (MIM 608667), *SMC3* (MIM 606062), or *SMC1* (MIM 300040). We screened the entire *PHF21A* coding region and intron-exon boundaries comprising 16 exons defined from RefSeq accession numbers NM\_001101802.1 and NM\_016621.3. We found no potentially pathogenic variants. We also screened exons 2–6 of *ELAVL1* (RefSeq NM\_001419.2) in 18 subjects with features of Saethre-Chotzen syndrome, 12 subjects with isolated features of



**Figure 1. Ideograms of Balanced Chromosome Translocations and FISH Mapping of the Chromosomal Breakpoints**

(A and B) Ideograms illustrating the  $t(11;19)(p11.2;p13.2)dn$  karyotype in DGAP012 and the  $t(1;11)(p21.1;p11.2)dn$  karyotype in MCN1762; this latter karyotype was revised from the original karyotypic assessment,  $t(1;11)(p13;p11)dn$ .

(C) FISH analysis with a 30 kb *Sna*BI restriction fragment (signal shown in green) from BAC CTD-3193O13, which has previously been reported as a breakpoint-crossing clone.<sup>24</sup> Probe hybridization is seen on the normal 19, der(19), and der(11) chromosomes, indicating that this 30 kb restriction fragment spans the chromosome 19 breakpoint region (Figure S1B).

(D) FISH analysis with 37 kb fosmid G248P88722D5 spanning the breakpoints on the normal 11, der(1), and der(11) chromosomes, indicating that the translocation breakpoint of the chromosome 11 is located within the sequence of this genomic clone.

Rubinstein-Taybi syndrome, 16 subjects with micropenis or hypogonadotropic hypogonadism and ID, and five subjects with metopic craniosynostosis; again, we detected no potentially pathogenic variants.

### RNA and ChIP Analysis in Lymphoblasts

For analysis of expression of neuronal genes in lymphoblastoid cell RNA, cells were harvested and RNAs were prepared with Trizol reagent (Sigma, St. Louis, MO). Lymphoblastoid cell cDNA samples were obtained by oligo dT and Superscript III (Invitrogen, San Diego, CA) with primers designed to cross multiple introns for the exclusion of genomic amplification. RT-PCR was carried out for 34 cycles by annealing at 60°C for 30 s and elongation at 72°C for 30 s with Platinum PCR SuperMix (Invitrogen).

LSD1 chromatin immunoprecipitation (ChIP) experiments were done as previously described.<sup>26,27</sup> Primer pairs for the *SCN3A* (MIM 182391) promoter target and control oligos designed in an intergenic region between *ACTG1* (MIM 102560) and *FSCN2* (MIM 607643) are listed in Table S5.

### Immunoblot Analysis

Whole cell protein extracts were prepared from harvested lymphoblastoid cells and were lysed on ice for 30 min in RIPA buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1% NP40, 0.5% sodium deoxycholate, and 0.1% SDS) containing a protease inhibitor mixture (Roche) and 1 mM phenylmethylsulfonyl fluoride. Total lysates were then cleared by centrifugation at 14,000 × g for 30 min, and supernatants were collected. Protein concentration was determined by a Bio-Rad (detergent compatible) protein assay. Fifty micrograms of protein extract was mixed with 4 × SDS sample buffer, boiled for 2 min, and subjected to 10% SDS-PAGE. After electrophoresis, proteins were transferred to Schleicher & Schuell nitrocellulose membranes (Whatman, UK) and incubated for 30 min in blocking solution containing 5%

nonfat powdered milk in TBS-T (50 mM Tris-HCl, 150 mM NaCl, pH 7.4, and 0.1% Tween 20). Blots were probed overnight at 4°C with antibodies against an N-terminal 93 residue polypeptide of mouse Phf21a<sup>28</sup> or  $\alpha$ -tubulin (Sigma), rinsed four times for 10 min in TBS-T, and incubated for 1 hr at room temperature with a horseradish peroxidase-conjugated anti-rabbit antibody. After extensive washing for 30 min, membranes were processed with an enhanced chemiluminescence (ECL) substrate kit (New England Biolabs, Beverly, MA) and exposed to autoradiographic film (Hyperfilm ECL, Amersham Bioscience, Piscataway, NJ).

### Immunostaining

Cells were grown on glass coverslips overnight. After being washed with PBS, cells were fixed in 4% paraformaldehyde and 0.5% Triton X-100 for 10 min. Coverslips were washed three times with PBS and incubated in block-permeabilization solution (15% goat serum [Sigma], 0.2% fish skin gelatin [Sigma], 0.03% NaN<sub>3</sub>, and 0.5% Triton X-100 in PBS) for 10 min. Cells were rinsed and incubated with blocking solution (15% goat serum, 0.2% fish skin gelatin, and 0.03% NaN<sub>3</sub> in PBS) overnight. After this, cells were incubated with a PHF21A antibody<sup>28</sup> in blocking solution for 2 hr at 37°C. Coverslips were washed three times with PBS and incubated for 1.5 hr with an Alexa Fluor 488-conjugated secondary antibody in blocking solution. Washing three times with PBS removed the excess antibody. DNA was stained with DAPI (Invitrogen). Images of fixed cells were collected with an LSM 510 META confocal microscope (Zeiss).

### Mouse In Situ Hybridization Analysis

The primers used for amplification of three independent probes of the murine *Phf21a* transcript (RefSeq NM\_138755.2) are listed in Table S6. Because multiple splice variants for *Phf21a* are listed in the Ensembl Browser (Ensembl Mouse, based on the National Center for Biotechnology Information [NCBI] m37

mouse assembly), probes were designed to detect the majority of the known isoforms. Probes were labeled with [ $\alpha$ -<sup>35</sup>S]-UTP for hybridization on 15  $\mu$ m cryosections. The day of plug was not counted for the specification of embryonic stages. No specific signals were detected with the respective sense probes except as indicated in the results section.

### Fish Stocks and Maintenance

Zebrafish were maintained at 28.5°C under a 14 hr light/10 hr dark cycle in 1/3 Ringer's solution. Transgenic fish *Tg[flkl:GFP]* (kindly provided by Dr. Tao P. Zhong) and *Tg[huc:EGFP]<sup>29</sup>* were used in overexpression or knockdown experiments. Embryos older than 24 hr postfertilization (hpf) were usually incubated in 0.003% 1-phenyl-2-thiourea (PTU, Sigma) for the inhibition of pigmentation. Embryos at appropriate stages were fixed with 4% paraformaldehyde in PBS.

### Zebrafish *phf21a* Constructs

Zebrafish *phf21a* was isolated from the 24 hpf zebrafish cDNA library by RT-PCR and was first cloned in a pGEM-T easy vector (Promega, Madison, USA) and then subcloned into the EcoRI site in the pCS2+ multipurpose expression vector. For the construction of the *phf21a-RFP* fusion reporter, specific enzyme-linked primers were designed for PCR amplification. PCR primers are listed in Table S7. PCR products were subcloned into the ClaI site in a pCS2+ RFP vector.

### Whole-Mount In Situ Hybridization and Alcian-Blue and Acridine-Orange Staining

Antisense digoxigenin-labeled RNA probes for *dlx2a*, *ngn1*, *huc*, and *phf21a* were produced with a DIG-RNA labeling kit (Roche, Germany) according to the manufacturer's instructions. Whole-mount in situ hybridization was performed with digoxigenin-labeled probes as previously described.<sup>30</sup> Cartilage staining was carried out with Alcian blue.<sup>31</sup> For the detection of apoptotic cells, embryos were placed in 10  $\mu$ g/ml acridine orange (Sigma) for 30 min and were washed in egg water.

### Microinjection of mRNA and Antisense MOs

Synthetic capped mRNAs for *PHF21A* and *phf21a* were transcribed in vitro with the linearized plasmid DNA as a template. mRNA was dissolved in 0.2% phenol red (as a tracking dye) and then microinjected into 1- to 2-cell-stage embryos. Antisense MOs for *phf21a* MO 5'-GCGTCATAAATGATATTACCTGTG-3' and standard control MO 5'-CCTCTTACCTCAGTTACAATTTATA-3' were synthesized by Gene Tools (Corvallis, OR, USA). Each morpholino was resuspended in 1  $\times$  Danieau buffer (58 mM NaCl, 0.7 mM KCl, 0.4 mM MgSO<sub>4</sub>, 0.6 mM Ca(NO<sub>3</sub>)<sub>2</sub>, 5.0 mM HEPES, and pH 7.6) and injected into 1- to 2-cell-stage embryos at the concentration of 5 ng/embryo.

## Results

### *PHF21A* Is Disrupted in Unrelated Subjects with Chromosomal Translocations, ID, and CFAs

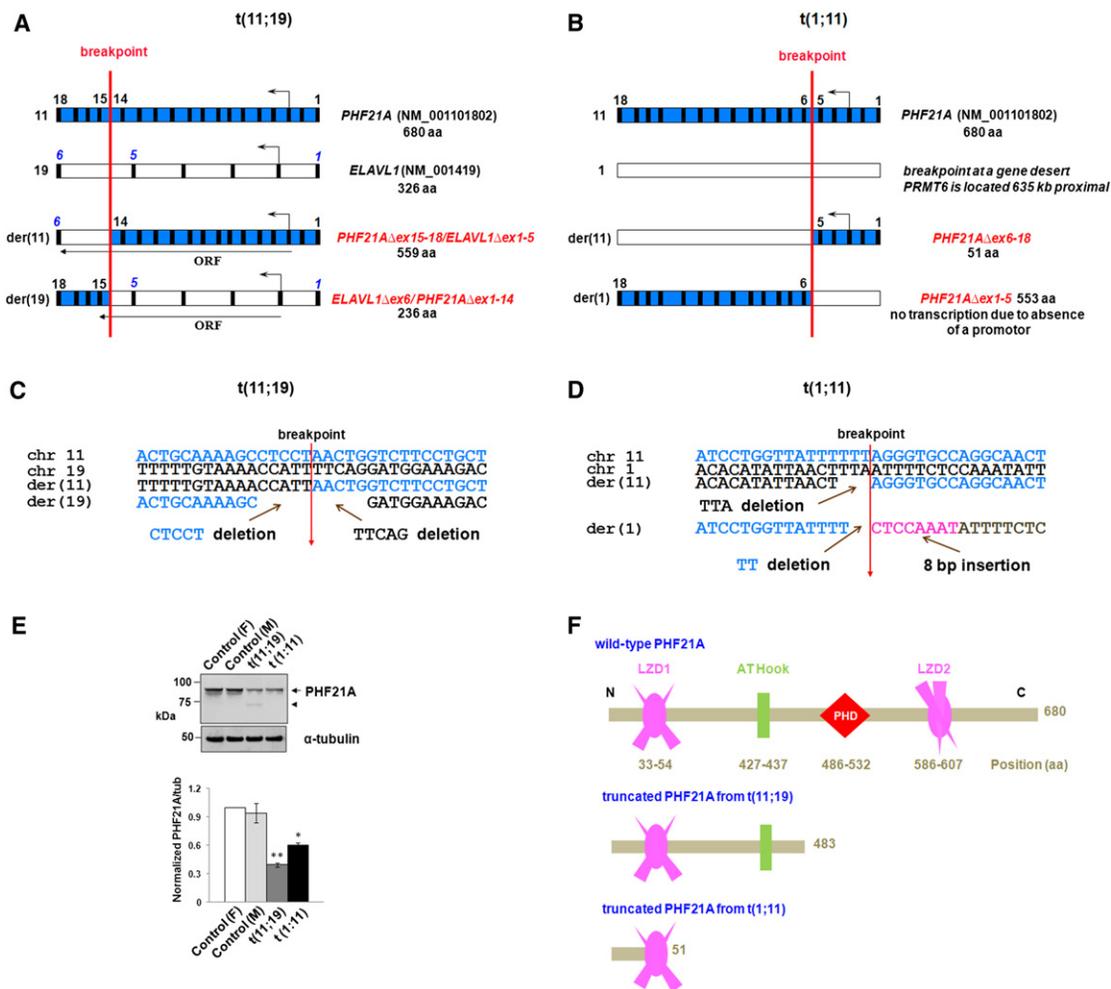
DGAP is a collaborative effort to identify genes of developmental importance through the study of individuals with apparently balanced chromosomal abnormalities and developmental defects.<sup>32</sup> Identification of multiple cases in whom the same gene is disrupted in independent

subjects with de novo translocations and similar phenotypes provides particularly strong evidence of the causative nature of the lesion. Through DGAP, we identified a subject (DGAP012) with an apparently balanced de novo translocation between chromosomes 11 and 19; this translocation resulted in a 46,XY,t(11;19)(p11.2;p13.2)dn karyotype (Figure 1A). A second subject, MCN1762 (MCN19730002-227), identified through the Mendelian Cytogenetic Network database, had an apparently balanced de novo translocation between chromosomes 1 and 11; this was initially reported as 46,XX,t(1;11)(p13;p11)dn (Figure 1B). They both display evidence of ID with CFAs, as well as other typical PSS features, except for multiple exostoses and parietal foramina, as summarized in Table 1, suggesting that the disruption in each case might affect the same gene in or near the PSS region in 11p11.2.

To map precisely the translocation in DGAP012, we first used FISH to bracket a candidate region and then to define a breakpoint-crossing BAC (Figure 1C and Figures S1A and S1B).<sup>24</sup> After DNA blotting (Figures S1C and S1D) to refine the breakpoints, we used suppression cloning<sup>23</sup> and targeted PCR for subsequent isolation and sequencing of junction fragments. Full details of the breakpoint-cloning steps are given in the Subjects and Methods section. The chromosome 19 breakpoint lies within a *SINE/Alu* sequence in intron 5 of *ELAVL1*, whereas the chromosome 11 breakpoint interrupts a *LINE/L2* repetitive element in intron 14 of *PHF21A* (Figure 2A). At the der(19) breakpoint, there was a 5 nt CTCCT deletion of chromosome 11 sequence and a 5 nt TTCAG deletion of chromosome 19 sequence, whereas at the der(11) breakpoint, there was no loss or gain of nucleotides as a result of the translocation (Figure 2C). The breakpoints were also confirmed independently with a multiplexed targeted-capture and sequencing approach as previously described.<sup>33</sup>

For the second subject, MCN1762, we used FISH (Figure 1D and Table S2) and array painting followed by PCR amplification to isolate the breakpoints, and this resulted in the revision of the previously designated karyotype to 46,XY,t(1;11)(p21.1;p11.2)dn (Figure 1B). The chromosome 1 breakpoint lies within a small 48 bp non-repetitive sequence surrounded by *LINE/L1* sequences in a region devoid of annotated genes; this region is 635 kb distal from the 5' end of *PRMT6* (MIM 608274). The chromosome 11 breakpoint is in intron 5 of *PHF21A* (Figure 2B). At the der(1) breakpoint, there was a 2 bp TT deletion of chromosome 1 sequence and an 8 bp CTCCAAAT insertion, whereas at the der(11) breakpoint, there was a 3 bp TTA deletion of chromosome 1 sequence (Figure 2D).

Immunofluorescence in cultured cells with an antibody specific to the amino-terminal segment of PHF21A revealed that the majority of the protein resides in the nuclei (Figure S3), consistent with a role for PHF21A in nuclear processes, such as chromatin association and transcriptional regulation. As expected from the sites of the translocations in both DGAP012 and MCN1762, disruption



## Figure 2. Mapping of the Breakpoints in Two Balanced Translocations

(A and B) Disruption of *PHF21A* by two translocations. A schematic diagram of the DGAP012 breakpoints (not to scale) shows, in blue, *PHF21A* on chromosome 11 (exons 1–18) and, in white, either (A) *ELAVL1* on chromosome 19 (exons 1–6) or (B) chromosome 1. All exons are depicted as vertical black boxes, and the start and direction of transcription are indicated by an arrow above the corresponding exons. The translocation occurred at the site of the red vertical line. In DGAP012, the translocation produced der(11) and der(19) chromosomes that encode potential fusion transcripts (A). Note that the stop codon of the potential *PHF21A $\Delta$ ex15-18/ELAVL1 $\Delta$ ex1-5* fusion gene is at the same location as wild-type *ELAVL1* because there is no frameshift, whereas *ELAVL1 $\Delta$ ex6/PHF21A $\Delta$ ex1-14* has a frameshift with a premature stop codon in the new exon 6 (equivalent to exon 15 of *PHF21A*). The translocation in MCN1762 does not predict any potential fusion product because the breakpoint on chromosome 1 is located in a gene desert.

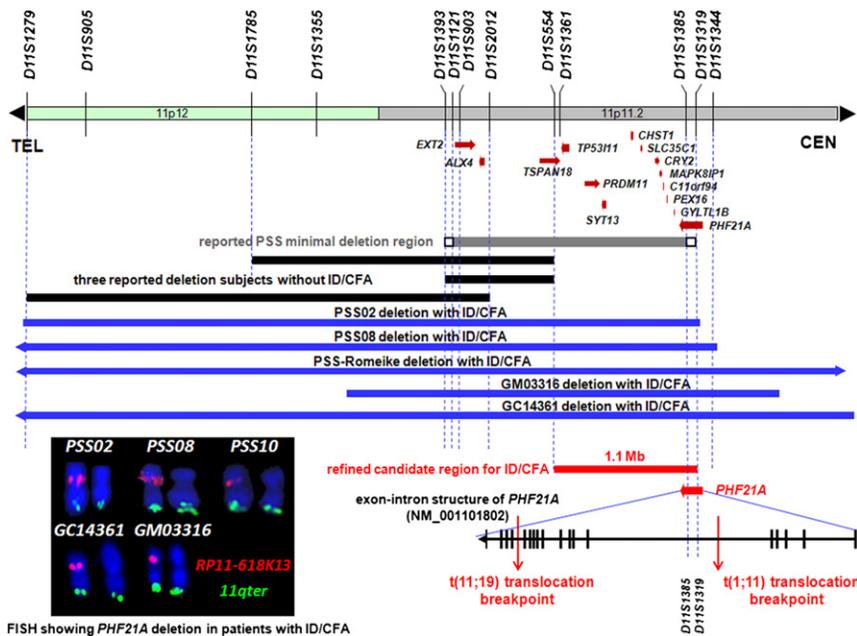
(C and D) Genomic DNA sequence from the normal chromosomes and at the breakpoints on derivative chromosomes. In DGAP012, at the der(19) breakpoint, there was a 5 nt CTCCT deletion of chromosome 11 sequence and a 5 nt TTCAG deletion of chromosome 19 sequence, whereas at the der(11) breakpoint, there was no loss or gain of sequence (C). In MCN1762, the junction sequences revealed a 3 bp TTA deletion of chromosome 1 sequence on der(11) and a 2 bp TT deletion of chromosome 11 sequence and an 8 bp CTCCAAAT insertion on der(1). Details of the mapping of both breakpoints in DGAP012 and MCN1762 are described in Figure S1 as well as in the Subjects and Methods section.

(E) Immunoblot analysis of *PHF21A* levels in DGAP012 and MCN1762 and in controls. The *PHF21A* antibody against an N-terminal 93 residue polypeptide<sup>28</sup> recognizes the ~92 kDa *PHF21A* in both female (“F”) and male (“M”) controls, as well as in DGAP012 and MCN1762 lymphoblastoid cell line extracts. It shows notably reduced protein levels due to disruption of *PHF21A* (arrow) in both DGAP012 and MCN1762. A 73 kDa protein (arrowhead) was noted in DGAP012 and is likely to be a product of the *PHF21A $\Delta$ ex15-18/ELAVL1 $\Delta$ ex1-5* fusion gene, which deletes the critical plant homeodomain (PHD) finger domain.  $\alpha$ -tubulin was used as an internal loading control. The bar graph shows the mean and  $\pm$  standard deviations from three independent experiments (\* $p < 0.001$ , \*\* $p < 0.0001$ ). The subcellular localization of *PHF21A* with the same antibody is described in Figure S3.

(F) *PHF21A* functional domains in wild-type and theoretical truncated proteins in two balanced translocation subjects. *PHF21A* contains two leucine zipper domains (LZD1 and LZD2), one AT-hook domain, and one PHD zinc finger domain. The amino acid positions of all domains are indicated as numbers below the domain structures. Note that if any protein were produced from the truncated *PHF21A* of DGAP012 or MCN1762, it would lack the PHD finger domain essential for binding H3K4me0.<sup>14</sup>

of *PHF21A* resulted in reduced protein levels of full-length *PHF21A* (as detected by immunoblot analysis) (Figure 2E). In both subjects, the *PHF21A* promoter could theoretically

drive expression of a truncated *PHF21A* either alone or, in the case of DGAP012, as part of a fusion protein (Figure S2). However, in neither case would such a protein product



**Figure 3. *PHF21A* Is Disrupted in DGAP012 and MCN1762 and Maps within the Refined Interval Associated with ID and CFAs**

The minimal deletion region associated with the full spectrum of PSS phenotypes, compiled from a consensus of ten PSS subjects, is shown as a gray line encompassing deleted markers and flanked by small white boxes leading to markers D11S1393 and D11S1319, which are at the end of each box and are definitely not deleted.<sup>15</sup> The PSS-associated genes *EXT2* and *ALX4*, together with *PHF21A*, all map to this deletion region. Black lines show regions of deletion in three subjects showing neither ID nor CFAs.<sup>34–36</sup> These deletion regions, subtracted from the above gray line, predict the ID and CFA candidate region (shown as a red line), a segment encompassing ~1.1 Mb and containing 12 annotated genes between D11S554 and D11S1319. The proximal end of the ID and CFA candidate region overlaps with the terminal 98 kb of *PHF21A*. The blue lines of PSS subjects with ID and CFAs show unanimous

deletion of *PHF21A*. At the bottom right, the breakpoints of DGAP012 and MCN1762 are shown. *PHF21A* is disrupted in both of these balanced translocation subjects with ID and CFAs but without multiple exostoses and parietal foramina. All together, the above genotype-phenotype analyses, combined with these two translocation cases, support a causative role for *PHF21A* in ID and CFAs. In the lower left corner is FISH mapping showing deletion of *PHF21A* in the five 11p11.2-deletion ID and CFA subjects shown above. FISH was carried out with BAC RP11-618K13 (red), which encompasses the 3' portion of *PHF21A*, for three PSS subjects (PSS02, PSS08, and PSS10) whose centromeric PSS deletion boundaries had not been fully delineated<sup>15</sup> and for two additional 11p11.2-deletion subjects (GC14361 and GM03316). Both chromosomes 11 from the same metaphase spread are shown, indicating the absence of *PHF21A* from one chromosome in each subject. Clone GS-770G7 from 11q25 (green) was used as a positive chromosome 11q control.

contain the plant homeodomain (PHD) required for binding H3K4me0<sup>14</sup> (Figure 2F).

Localization of the chromosome 11 breakpoint from each subject within the same gene with consequent reduction in the protein product strongly supports a role for disruption of *PHF21A* in generating the ID and CFA phenotypes. This possibility was further supported by a third translocation subject, a young girl who displayed both ID and CFAs in conjunction with phenotypes reminiscent of Gillespie syndrome (MIM 206700). In this female subject, the translocation directly disrupted both *PHF21A* and the X chromosome gene *ARHGAP6* (MIM 300118), encoding Rho-GTPase-activating protein 6.<sup>12,13</sup> The precise breakpoints in this case were not determined (Table 1). The phenotypes of all three translocation subjects are compared in Table 1.

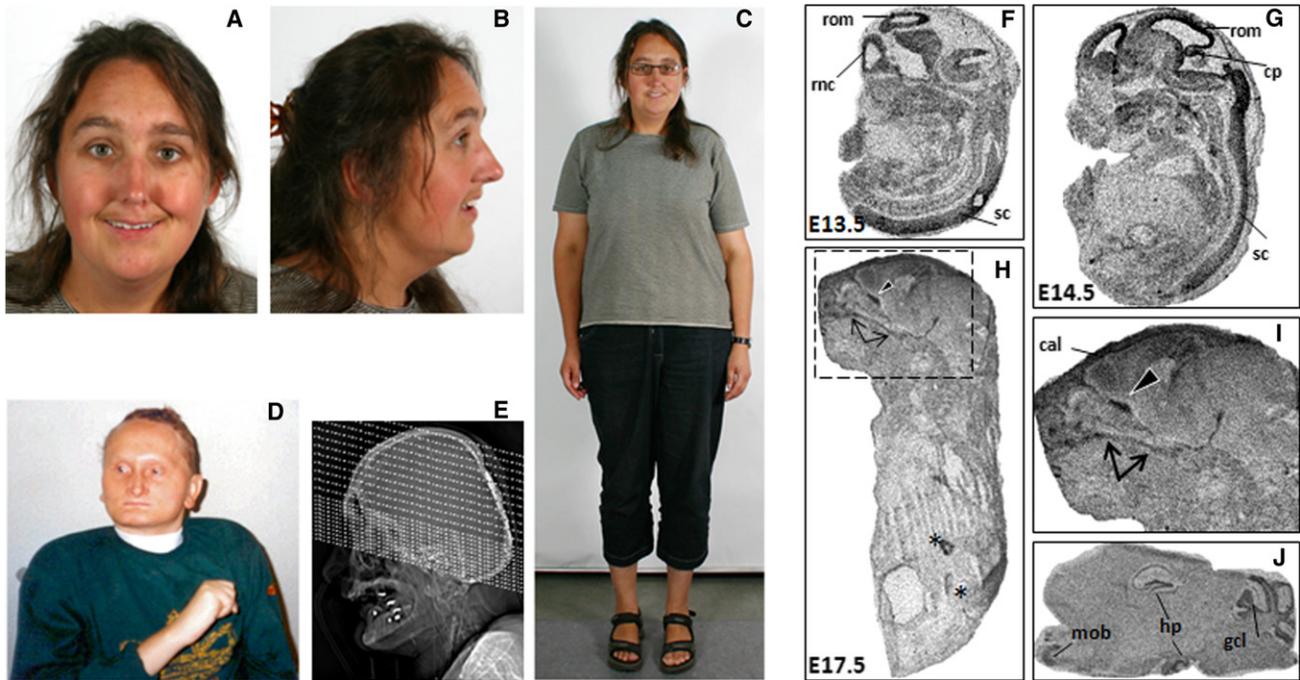
#### ***PHF21A* Maps to the Critical PSS Interval Associated with ID and CFAs**

Mapping of the translocation sites in *PHF21A* places them in proximity to the 11p11.2 genomic deletion region associated with PSS, in which phenotypic manifestations depend upon the precise location and extent of the deletion.<sup>15</sup> The full phenotypic spectrum of PSS is manifested when deletions are at least 2.1 Mb in size,<sup>15</sup> in which case they span the segment from D11S1393 to D11S1385/D11S1319 and contain 16 annotated genes

and (Figure 3). The three translocation subjects share phenotypes with PSS subjects, but the notable exceptions are multiple exostoses and parietal foramina (Table 1), which are known to be caused by two genes (*EXT2* and *ALX4*, respectively) distal to *PHF21A*. As an example, Figure 4 displays one of our translocation subjects and a PSS subject who both have CFAs that include microcephaly (Figures 4A, 4B, and 4D), brachycephaly (Figures 4A, 4B, and 4D), midfacial hypoplasia (Figure 4B), and a hypoplastic mandible (Figures 4D and 4E). Notably, neither of the two translocation subjects investigated here shows a second alteration on the nontranslocated chromosome by exon sequencing of *PHF21A*.

A critical reassessment of the in silico comparative deletion mapping of published deletion subjects, including those who display multiple exostoses and parietal foramina without ID or CFAs,<sup>34–36</sup> excludes the distal portion of the PSS interval. This leaves a ~1.1 Mb ID and CFA candidate interval that is between D11S554 and D11S1319 and encompasses 12 annotated genes (*TSPAN18*, *TP53I11*, *PRDM11*, *SYT13*, *CHST1*, *SLC35C1*, *CRY2*, *MAPK8IP1*, *C11orf94*, *PEX16*, *GYLTL1B*, and *PHF21A*) (Figure 3). *PHF21A* maps at the centromeric end of this candidate interval, and markers D11S1385 and D11S1319 are located within intron 5 (Figure 3).

To test whether *PHF21A* disruption occurs in PSS with ID and CFAs, we investigated five subjects clinically



**Figure 4. Human Craniofacial Anomalies and Murine *Phf21a* Expression in Craniofacial and Brain Regions**

(A–E) Human craniofacial anomalies with *PHF21A* truncation and *PHF21A* deletion.

(A–C) Female subject MCN1762 with balanced translocation  $t(1;11)$  exhibits ID and CFAs that include brachycephaly, microcephaly, a long narrow nose, mild midfacial hypoplasia, thin lips, and prominent ear lobes.

(D) The PSS male proband (PSS-Romeike) with a deletion including *PHF21A* is shown at the age of 31 years. He shows microcephaly, brachycephaly, a broad forehead, a long narrow nose, a hypoplastic mandible, very thin lips, hypotelorism, and dysplastic low set ears, in addition to ID and no speech, among other phenotypes.

(E) A topogram of a cerebral CT of individual PSS-Romeike at the age of 33 years, shortly before he died, shows a hypoplastic mandible and the parietal foramen at the back of the head.

(F–J) *Phf21a* expression during murine craniofacial development.

(F and G) Expression of *Phf21a* at embryonic days (E) 13.5 and 14.5. Highest transcript abundance is seen in the developing CNS with [ $^{35}$ S]-UTP labeled in situ probes. The following abbreviations are used: rom, roof of midbrain; rnc, roof of neopallial cortex; sc, spinal cord; and cp, intraventricular portion of cerebellar primordium.

(H and I) At E17.5, prominent signals were detected with *Phf21a* antisense probes in bones of the facial skeleton. The palatine bone is marked by arrows, and the orbitosphenoidal bone is marked by a black arrowhead. Signals in the intestine were also detected with the sense probes and are most likely not specific (these are marked by asterisks). The following abbreviation is used: cal, calvaria.

(J) Sagittal sections of the adult mouse brain showed expression of *Phf21a* within the hippocampal formation, the cerebellum, and the olfactory bulb. Three different *Phf21a* antisense probes gave consistent results. The following abbreviations are used: mob, main olfactory bulb; hp, hippocampus; and gcl, granule cell layer of the cerebellum.

diagnosed with PSS and one subject with a PSS-like phenotype in association with a 11p11.2 deletion whose extent had not been previously resolved (Table 2). For five subjects (PSS02, PSS08, PSS10, GC14361, and GM03316), we were able to perform FISH with BAC clone RP11-618K13, which contains the telomeric 65 kb of *PHF21A* and spans the breakpoint in DGAP012. This BAC revealed heterozygous deletions in all five subjects (Figure 3, inset). In addition, we carried out array-CGH analysis on PSS02, PSS08, GC14361, GM03316, and one additional PSS subject (PSS-Romeike; Figure 4) reported recently,<sup>18</sup> and this confirmed the loss of *PHF21A* in the first four subjects and also established it for the final subject (Figure 3 and Table 2). The coordinates of the deletions found by array CGH are listed in Table 2. Thus, *PHF21A* is hemizygous in all six subjects, five of whom were clinically diagnosed with PSS and all of whom display

both ID and CFA phenotypes comparable to those of the translocation subjects.

To test the frequency of *PHF21A* hemizyosity in apparently normal individuals, we obtained a CNV dataset for 13,991 independent control individuals (collected across multiple studies<sup>37–42</sup>) from the International Schizophrenia Consortium and from Cooper et al.<sup>43</sup> In this collection, the only structural variations (SVs) that crossed an exon of *PHF21A* were two identical 590 bp duplications that overlapped portions of exon 1 and intron 1 (the coding sequence of *PHF21A* begins in exon 3). We also examined the Database of Genomic Variants (DGV), which shows six SVs reported in this region; these include four small deletions (444–5,230 bp) within *PHF21A* introns (one in intron 1, two in intron 5, and one in intron 6) and one example of the aforementioned 590 bp 5' UTR duplication. Data from the 1,000 Genomes Project also

shows multiple CNVs at each of the exon 1 and intron 5 locations, as well as a single 545 bp CNV in the 3' UTR.<sup>44</sup> In addition to these presumably benign SVs, there is a single report in the DGV of a large 75 kb deletion encompassing six genes, including a portion of *PHF21A*. We are not in a position to validate that deletion or to confirm the absence of phenotype; however, there is precedent for even well-established disease-associated CNVs being nonpenetrant in some individuals.

### Murine *Phf21a* Is Expressed in the CNS and Cranial Bones

To determine whether the pattern of expression of *PHF21A* supports a role in craniofacial and neuronal development, we performed in situ hybridization experiments for the orthologous mouse gene. Predominant expression of *Phf21a* is detected in the developing CNS at early stages. At mouse embryonic days (E) 13.5 and 14.5, the roof of the neopallidal cortex and thus the developing cerebral cortex, as well as the roof of the midbrain and the spinal cord, showed the highest expression levels of *Phf21a*. The intraventricular portion of the cerebellar primordium also expressed *Phf21a* at E14.5 (Figure 4G). At early embryonic stages, facial bone and viscerocranial ossification initiates and, with ongoing ossification, high levels of *Phf21a* transcripts were found at E17.5 in the palatine bone (Figure 4H and magnification in Figure 4I, marked by arrows) and the orbitosphenoidal bone (Figure 4H, I black arrowhead), as well as in the calvaria. Signals observed in bone with *Phf21a* antisense probes were essentially restricted to cranial bones, suggesting a particular function for *Phf21a* in craniofacial development. In the adult mouse brain, the most abundant expression of *Phf21a* was observed in the neuronal layers of the hippocampus, the granule cell layer of the cerebellum, and the main olfactory bulb (Figure 4J). All together, these findings indicate that expression of *Phf21a* is consistent with an important role in the CNS and craniofacial skeletal development and in adult neuronal function. Interestingly, a single report of a mouse knockout for *Phf21a* described no gross morphological abnormality, although the potential for CFAs was not specifically evaluated. Neonatal mice died as a result of an inability to suckle properly; this inability was interpreted as a likely defect in neuronal control of milk-sucking behavior.<sup>45</sup>

### Suppression of Zebrafish *phf21a* Expression Causes CFAs and Neuronal Apoptosis

To directly test the developmental importance of *PHF21A*, we isolated the zebrafish *phf21a* ortholog, examined its expression pattern, and performed gain- and loss-of-function experiments in this model organism. The zebrafish *phf21a* is highly related to human and mouse *PHF21A* proteins: It exhibits an AT-hook domain, a PHD, and two coiled-coil domains (data not shown). Using RT-PCR, we confirmed that zebrafish *phf21a* showed maternal and zygotic transcripts during embryonic devel-

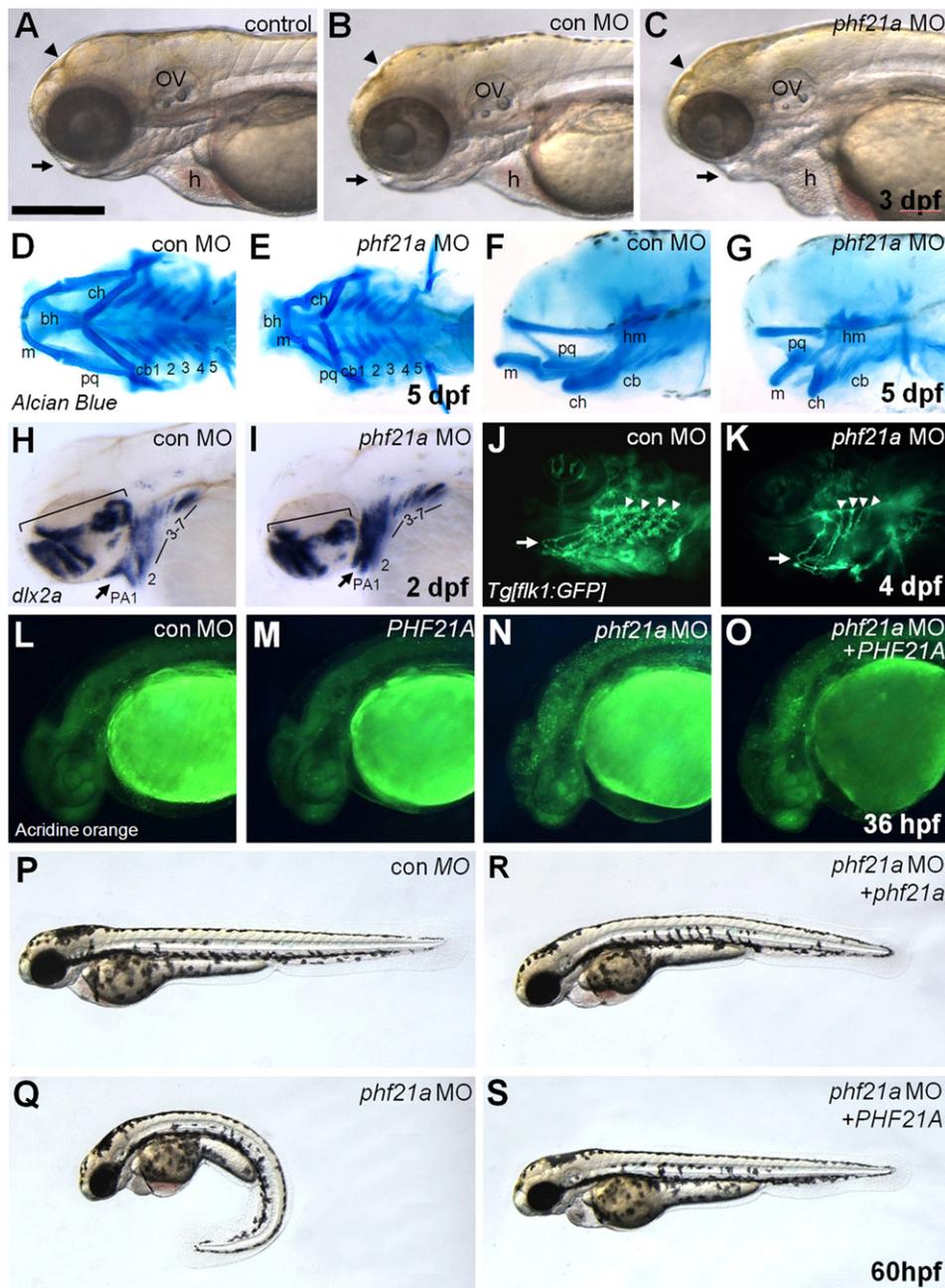
opment (Figure S4A). Whole-mount in situ hybridization analyses revealed *phf21a* transcripts ubiquitously distributed throughout the embryo during the stages of cleavage, blastula, gastrula, and early segmentation. Expression in the head region was increased from later somitogenesis and continued to 24 and 48 hpf (Figures S4B–S4J).

To investigate the function of *phf21a* in zebrafish development, we tested the effect of *phf21a* knockdown by antisense MO. Injection of the *phf21a* MO, but not of a standard control MO or no MO, caused a small-head phenotype and facial dysmorphism with a pronounced defect in growth of the lower jaw at 3 days postfertilization (dpf); these features are reminiscent of the microcephaly and dysmorphism seen in the translocation subjects (Figures 5A–5C and 5Q).

We examined the head structure of *phf21a* morphants in more detail by using Alcian blue to visualize the extent of cartilage development in larval fishes. At 5 dpf, Meckel's and palatoquadrate cartilages were severely distorted in their size and shape in *phf21a* morphants (Figures 5D–5G). Such defects were already manifest during early stages given that we also observed defects in *dlx2a*-positive pharyngeal-arch-cartilage progenitor cells in *phf21a* morphants at 2 dpf (Figures 5H and 5I). Defects in cranial-cartilage formation were also observed for the zebrafish *headless* mutation, which is known to be involved in the signaling pathway of vertebrate head formation and patterning.<sup>31</sup> To investigate further whether these defects also involve other arch-associated structures, we injected the *phf21a* MO into *flk1:GFP* transgenic zebrafish, in which the vascular endothelial cells were visualized by green fluorescent protein (GFP) fluorescence.<sup>46</sup> At 4 dpf, the aortic arches of the *phf21a*-MO-injected *flk1:GFP* transgenic embryos were found to be hypoplastic: They showed poor development of capillary networks associated with pharyngeal arches (Figures 5J and 5K). In vertebrates, Meckel's and palatoquadrate cartilages form the embryonic jaw apparatus.<sup>47,48</sup> Thus, it would be interesting to examine whether the *Phf21a/Bhc80*-deficient mice, which display a failure to suckle,<sup>45</sup> might also have a defect in jaw structure.

We also examined the effects of gain or loss of *phf21a* function on neuronal development but did not see any prominent change (Figure S5). However, injection of the *phf21a* MO, but not a standard control MO or *PHF21A* mRNA, caused apoptosis in the developing brain region at 36 hpf (Figures 5L–5N). Importantly, this apoptosis and the small-head phenotype can be rescued by introduction of wild-type human *PHF21A* mRNA (Figure 5O), suggesting that the ID phenotype in humans might be due to a requirement for *PHF21A* in the function of neuronal cell survival in the developing brain.

Overall, the *phf21a* MO caused craniofacial, morphological, and growth defects in the developing zebrafish embryo, as depicted by the notable ventral curvature of the body and small-head phenotype (Figure 5Q) relative to that of the control (Figure 5P). The body axis of the



### Figure 5. PHF21A Regulates Zebrafish Neuronal Cell Survival and Craniofacial Development

(A–C) A noninjected control embryo (A), an embryo injected with a control MO (B), and an embryo injected with a *phf21a* MO (C). Knockdown of *phf21a* causes a reduction in head size, a change of head and face shape (arrowhead), and a major reduction of the jaw (arrow) at 3 dpf; these features are reminiscent of the microcephaly and craniofacial dysmorphism seen in the translocation subjects. The scale bar in (A) represents 300  $\mu$ m. The following abbreviations are used: ov, otic vesicle; and h, heart.

(D–G) Cartilage staining of embryos injected with the control MO (D and F) or *phf21a* MO (E and G). Compared with that of the control embryo, Meckel's and palatoquadrate cartilage in the *phf21a*-MO-injected embryos are severely distorted in their size and shape. Five-day-old embryos are shown in ventral (D and E) or lateral (F and G) views. The following abbreviations are used: ch, ceratohyal; bh, basihyal cartilage; m, Meckel's cartilage; pq, palatoquadrate cartilage; and cb 1–5, ceratobranchial cartilage 1–5.

(H and I) *dlx2a* expression in the control (H) and the *phf21a* morphant (I) at 2 dpf. *phf21a* morphants fail to expand *dlx2a*-expressing pharyngeal-arch progenitor cells (arrow). The scale bar in (H) represents 300  $\mu$ m, and the scale bar in (I) represents 230  $\mu$ m. The following abbreviation is used: PA 1–7, pharyngeal arches 1–7.

(J and K) Formation and patterning of arch-associated blood vessels in the control (J) and the *phf21a* morphant (K). A lateral view of the *flk1:GFP* transgenic line at 4 dpf is shown. In *phf21a*-MO-injected embryos, aortic arches (arrowheads) form normally but fail to swing to a more anterior position (compare the position of arrows in J and K). Also, the vessels associated with gill filaments develop poorly in morphants.

(L–O) Effects on neuronal cell survival. Compared with the control MO (L) and *PHF21A*-mRNA-injected embryos (M), *phf21a*-MO-injected embryos (N) show a dramatically increased number of acridine-orange-positive apoptotic cells. The apoptotic phenotype of the *phf21a* morphant is rescued by coinjection of *PHF21A* mRNA (O).

zebrafish normally straightens from its early curvature in the pharyngula developmental stage (24–48 hr), during which *phf21a* is strongly expressed in the spinal cord (Figures S4G and S4I). Therefore, *phf21a* depletion in the spinal cord might impair the straightening process, resulting in persistent spinal and tail curvature. Two processes that have previously been implicated in such ventral curvature are dorsal midline development and cilia development, but the relationship of this phenotype to human PSS is unclear. Notably, injection of either wild-type *phf21a* or *PHF21A* mRNA rescued both the spinal curvature and small-head phenotypes (Figures 5R and 5S), demonstrating that they result from a lack of a conserved function of the protein.

### Disruption of *PHF21A* in the Translocation Subjects Derepresses *SCN3A*

Along with LSD1, *PHF21A* is known to be a component of the CoREST-related protein complex, BRAF-HDAC complex (BHC), which participates in neuron-specific gene repression presumably by regulating histone-demethylation activity.<sup>26,27</sup> *PHF21A* specifically recognizes unmethylated histone H3 lysine 4 residues and is required for LSD1-mediated transcriptional repression and LSD1 occupancy at target promoters.<sup>14</sup> To investigate whether LSD1-mediated transcriptional repression is functionally altered by the *PHF21A*-disrupting translocations, we first examined transcription levels of several reported LSD1 targets, including *SCN1A*, *SCN3A*, and *SYN1*, in lymphoblastoid cell lines derived from normal males, normal females, and translocation subjects DGAP012 and MCN1762. We found that *SYN1* and *SCN1A* have high and moderate levels, respectively, of expression, even in normal lymphoblastoid cells, indicating that they are not epigenetically repressed like they are in HeLa cells.<sup>26,27</sup> However, *SCN3A* mRNA was expressed at a lower level as measured by RT-PCR in lymphoblastoid cells from normal controls, suggesting that this LSD1 target gene is transcriptionally repressed. An increase of *SCN3A* transcript was readily detected in DGAP012 and MCN1762 cells (Figures 6A and 6B), indicating that a single functional allele of *PHF21A* might not be sufficient for effective repression of *SCN3A* transcription. This is consistent with the fact that correct dosage of *PHF21A* is important for its function. In support of this hypothesis, ChIP results showed a significant reduction of LSD1 binding to the *SCN3A* promoter in DGAP012 and MCN1762 cells compared to normal control cells (Figure 6C). These findings indicate that *SCN3A* might be repressed by the LSD1 corepressor complex through a mechanism similar to that reported in HeLa cells<sup>14,26</sup> and that *PHF21A* might be similarly

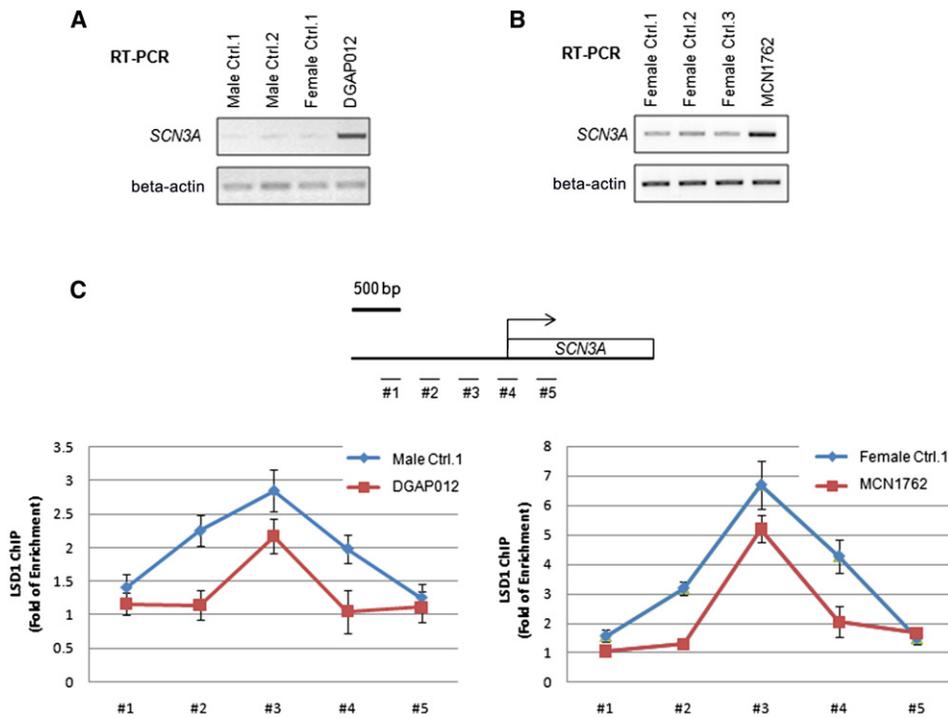
required for LSD1 promoter occupancy in lymphoblasts. This functional disruption of *PHF21A* in DGAP012 and MCN1762 lymphoblasts is probably mirrored by alterations in gene regulation in many tissues, including the CNS.

## Discussion

PSS is a contiguous gene syndrome involving ID and CFAs, along with other distinctive features, including eye abnormalities (severe myopia, nystagmus, and strabismus), skeletal anomalies (small hands and feet and tapering fingers), multiple exostoses, and enlarged parietal foramina.<sup>7,15,16,36,49,50</sup> The challenging aspect of positional cloning in such contiguous gene syndromes is assigning individual phenotypes to individual genes in the deleted region. Because the size and location of contiguous-gene-syndrome deletions can vary from individual to individual, comparison of overlapping regions for defining a minimal candidate region associated with a particular phenotype has often preceded candidate-gene analysis for identifying the associated gene.<sup>51,52</sup> These strategy and linkage analyses, respectively, led to two genes implicated in the pathogenesis of PSS: *ALX4*, associated with enlarged parietal foramina,<sup>53</sup> and *EXT2*, responsible for multiple exostoses.<sup>9</sup> However, these approaches have not identified the gene(s) responsible for the ID and CFA phenotypes. Attempts to identify the gene(s) underlying the latter PSS phenotypes have been hampered by the relatively large size of the minimal candidate interval (~2.1 Mb).<sup>15</sup> Both categories of clinical features are individually relatively common: ID affects ~2%–3% of humans,<sup>54</sup> and CFAs are present in ~1/3 of human congenital defects.<sup>55</sup> In many cases, these phenotypes manifest together, suggesting a common underlying etiology. Our reinterpretation of the PSS candidate region for ID and CFAs to ~1.1 Mb on the basis of published subjects with neither ID nor CFAs<sup>34–36</sup> and the identification of two independent translocation subjects with breakpoints at the proximal end of this region suggest that disruption of a single gene, *PHF21A*, is responsible for both ID and CFAs. This finding is supported by a third translocation case from the literature.<sup>12,13</sup>

Haploinsufficiency of *PHF21A* is the probable cause of ID and CFAs in all cases that we studied given that we detected no additional *PHF21A* mutation on the nontranslocated alleles in DGAP012 and MCN1762 and no clear differences between the critical ID and CFA features of these subjects and those of the PSS subjects with interstitial deletions (Tables 1 and 2). In MCN1762, *PHF21A* was the

(P–S) The *phf21a*-MO causes craniofacial, morphological, and growth defects in the developing zebrafish embryo. Reintroduction of wild-type *phf21a*/*PHF21A* rescues the zebrafish phenotype. A control-MO-injected embryo is shown in (P). Knockdown of *phf21a* causes ventral curvature of the body and a small-head phenotype (Q). These phenotypes of the *phf21a* MO are rescued by coinjection of either zebrafish *phf21a* mRNA (R) or human *PHF21A* mRNA (S). Expression of *phf21a* during zebrafish early development is described in Figure S4.



**Figure 6. *SCN3A* Transcriptional Derepression and Reduced LSD1 Promoter Occupancy in DGAP012 and MCN1762 Lymphoblastoid Cells**

*SCN3A* transcription is derepressed in affected individuals DGAP012 and MCN1762. Semiquantitative RT-PCR reactions were carried out in separate experiments for comparing the levels of *SCN3A* transcript (top box) in normal male and female controls with those of DGAP012 (A) and MCN1762 (B). Beta-actin (lower box) was used as a control in both experiments.

(C) LSD1 occupancy at the *SCN3A* promoter was compared in separate experiments between a normal male (#1 in left panel) and DGAP012 and between a normal female (#1 in right panel) and MCN1762 with the use of a ChIP assay. Five sets of PCR primers spanning the *SCN3A* promoter were used. These assays revealed reduced LSD1 occupancy at the *SCN3A* promoter in DGAP012 and MCN1762. The two panels cannot be directly compared because the two sets of assays were carried out at different times with different reagents (Abcam ab-17721 in left panel and Cell Signaling 2184 in right panel) and detection methods (radioactive PCR in left panel and SYBR Green real-time PCR in right panel). Error bars represent the standard error of the mean calculated on the basis of three independent experiments. An intragenic region upstream of the ACTG promoter not bound by LSD1 was used as a reference region in the ChIP assay for normalization.

only gene disrupted, whereas in DGAP012, *ELAVL1* was also truncated. *ELAVL1* encodes a protein that contains three RNA-binding domains and binds *cis*-acting AU-rich elements. It destabilizes mRNAs and thereby regulates gene expression.<sup>56,57</sup> Because both subjects show comparable phenotypes, it is unlikely that *ELAV1* haploinsufficiency or fusion proteins contribute strongly to the critical PSS-like phenotypes (Table 1), although they might be responsible for unusual features, such as possible retinal dystrophy and a digitized thumb, in DGAP012. The third translocation subject was only examined as a very young girl and displayed some phenotypes (bilateral superior coloboma, foveal hypoplasia, inferior cerebellar hypoplasia, and ID) thought to be suggestive of Gillespie syndrome (partial aniridia, cerebellar ataxia, and ID). Gillespie syndrome is known to be genetically heterogeneous and is caused in some cases by lesions in *PAX6* (MIM 607108). In this subject, it is likely that the disruption of *PHF21A* contributed to the ID and craniofacial dysmorphism in common with the other two translocation subjects (Table 1). The disparate phenotypes might then have been contributed by disruption of *ARHGAP6* or by an

independent lesion elsewhere. Notably, mice homozygous for a targeted null mutation of *ARHGAP6* do not exhibit any detected phenotypic or behavioral abnormalities.<sup>58</sup> A strict *PHF21A* dosage requirement for normal craniofacial and neurodevelopment is also supported by our findings in the zebrafish system where *phf21a* suppression produced abnormalities in the development of the head, face, and jaw, as well as increased neuronal apoptosis. Importantly, these deficiencies were rescued by the human *PHF21A*, indicating a conserved developmental function.

The discovery that translocations disrupting *PHF21A* are associated with abnormal craniofacial and intellectual development adds to the evidence that regulation of gene expression through chromatin modifications is crucial to both processes. To date, three genes, *NSD1* (MIM 606681), *KDM5C* (MIM 314690), and *PHF8* (MIM 300560), encoding proteins that have PHD domains and bind methylated histone tails have been found in syndromes with ID and distinctive craniofacial features. *NSD1*, which is deficient in Sotos syndrome, specifically binds methylated H3K4 and H3K9 via domains PHD1, PHD4, PHD5, and PHD6;<sup>59</sup> *SMCX/JARID1C* (encoded by

*KDM5C*) of nonsyndromic XLID binds histone H3K9me3 via its PHD1 domain;<sup>60</sup> and PHF8 of Siderius-Hamel syndrome binds histone H3K4me3 via its PHD.<sup>61</sup> At least the latter two act as demethylases by targeting H3K4me2 and H3K4me3 in the case of SMCX/JARID1C<sup>60</sup> and H3K9me1/2 and H4K20me1 in the case of PHF8.<sup>61,62</sup>

Unlike the above proteins, PHF21A is neither a methyltransferase nor a demethylase but instead specifically binds histone H3K4 when it is not methylated.<sup>14</sup> This suggests that both recognition of the unmodified state of histone tails and binding of proteins to methylated histone tails are critical for maintaining the appropriate balance and control of particular chromatin modifications for the support of normal intellectual and craniofacial development. Although the PSS-associated ID and CFAs appear to be due to haploinsufficiency of *PHF21A*, it has not been possible to ascertain and screen a large series of nontranslocation subjects with comparable phenotypes for *PHF21A* mutations. Therefore, we cannot state with certainty that missense, nonsense, splicing, or other mutations in *PHF21A* would lead to the same ID and CFA phenotypes. Indeed, it is conceivable that other types of genetic lesions in *PHF21A* could actually be associated with other developmental phenotypes; we were able to identify 200 individuals with ID and/or CFAs but without the full constellation of phenotypes exhibited by our translocation subjects, and we performed a mutation screen of *PHF21A*. We did not detect any truncating or missense mutations that have implicated particular PHD domains in NSD1 in binding to their methylated targets and that could thus aid in structure-function experiments. This is not surprising given the frequency of these two major phenotypes and the extent of genetic heterogeneity underlying each of them, but more extensive mutation analysis of subjects with various manifestations of ID, CFAs, and additional phenotypes seen in our translocation subjects might prove valuable to understanding the functional domains of PHF21A.

PHF21A (BHC80) is known to participate in the six-subunit BHC, which also comprises BRAF35 (MIM 605535), HDAC1 (MIM 601241), HDAC2 (MIM 605164), CoREST (MIM 607675), and LSD1 (BHC110); the latter is a histone demethylase that targets H3K4me2.<sup>26</sup> This complex interacts with the promoters of genes, such as synapsin and sodium-channel genes, to mediate repression of these neuron-specific genes through the *cis*-regulatory element known as repressor element 1 or neural restrictive silencer (RE1/NRS).<sup>63</sup> Specific binding of PHF21A to H3K4me0 is required for optimal LSD1 promoter occupancy *in vivo* and for LSD1-mediated gene repression.<sup>14</sup> Our data showing derepression of the neural gene *SCN3A* in lymphoblasts from the translocation subjects as a consequence of reduced levels of PHF21A are consistent with this role for the protein. Repression of neuronal-specific genes is of fundamental importance in the development of both neuronal and nonneuronal tissues,<sup>63</sup> so the failure of this particular function in the translocation subjects might have contributed to their ID and CFAs.

Another interesting XLID candidate, *ZMYM3* (MIM 300061), encodes a zinc finger protein that is predominantly expressed in the brain and that is a component of transcriptional corepressor complexes that also contain LSD1 (BHC110) and HDAC2.<sup>64,65</sup> The 5' UTR of *ZMYM3* is disrupted by a presumably balanced t(X;13) karyotype in a female with ID and preferential inactivation of the normal X chromosome.<sup>66</sup> In addition, Kleefstra syndrome, characterized by ID and CFAs comparable to PSS, has been associated with disruption of *EHMT1*, encoding euchromatin histone methyltransferase 1, which acts as a methyltransferase to modify H3K9 and has been reported as a component of the E2F6 transcription repressor complex and of a CtBP repressor complex that also contains LSD1.<sup>2,67-69</sup> The parallels between *PHF21A*, *ZMYM3*, and *EHMT1* suggest that other X-linked and autosomal loci underlying ID and/or CFAs might encode proteins that participate in complexes involving LSD1 or potentially other demethylases or methyltransferases. Our finding that decreased dosage of PHF21A, a histone-binding protein that interacts with and is required for the histone-demethylase activity of LSD1, leads to both ID and CFAs provides the proof of principle for investigation of other regulators of histone modification as genetic factors in ID and/or CFAs. Indeed, the recent finding of haploinsufficiency of *ARID1B* (MIM 614556), encoding an E3-ubiquitin-ligase component that functions with the chromatin-remodeling switch/sucrose nonfermentable complex,<sup>70-74</sup> suggests that genes involved with other aspects of chromatin modification might also contribute to ID and CFAs and that ultimately, human mutations affecting both regulatory and enzymatic components of histone-modification complexes might represent important tools for delineating the chromatin-regulation features that are critical for normal craniofacial and neurological development and cognitive function.

### Supplemental Data

Supplemental Data include five figures and seven tables and can be found with this article online at <http://www.cell.com/AJHG>.

### Acknowledgments

We are grateful to DGAP012 and MCN1762, as well as their family members, for their cooperation and participation in this study. We are also indebted to Amy Bosco, Heather L. Ferguson, and Chantal Kelly for obtaining informed consent and clinical information; to Joanne Sutherland, genetic counselor at the Hospital for Sick Children, Toronto, for her assistance in obtaining samples; to Shigeki Iwase and Tadashi Baba for the PHF21A antibody; to Mary Anne Anderson and Tammy Gillis in the Center for Human Genetic Research Tissue Culture Facility and Genomics Resource for technical assistance; to Ian Krantz, Stephanie Seminara, and Simeon A. Boyadjiev for providing samples of affected individuals; and to Ines Müller and Corinna Menzel for technical assistance. This work was supported by a grant from the Next-Generation BioGreen 21 Program (PJ00812701 to C.H.K.), Republic of Korea,

a grant from the Deutsche Forschungsgemeinschaft (KU 1240/5-1 to K.K.), the Danish National Research Foundation (N.T.), the Lundbeck Foundation (N.T. and A.S.), National Institutes of Health grants NCI118487 (to Y.S.) and RO1 GM071004 (to Y.S.), and United States Public Health Service grants GM061354 (Developmental Genome Anatomy Project to C.C.M. and J.F.G.) and HD065286 (to J.F.G.). Part of this work was financed by the European Union's Seventh Framework Program under grant agreement number 241995, project GENCODYS, and the German Federal Ministry of Education and Research through the German Mental Retardation Network (grant 01GS08161 to H.H.R.).

Received: December 27, 2011

Revised: March 18, 2012

Accepted: May 10, 2012

Published online: July 5, 2012

## Web Resources

The URLs for data presented herein are as follows:

Agilent Technologies, <https://earray.chem.agilent.com>

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

Mendelian Cytogenetics Network Online Database, <http://www.mcndb.org/index.jsp>

NCBI and GenBank, <http://www.ncbi.nlm.nih.gov>

NIGMS Human Genetic Cell Repository, <http://ccr.coriell.org/nigms/>

Online Mendelian Inheritance in Man (OMIM), <http://www.omim.org>

UCSC Genome Browser, <http://genome.ucsc.edu/>

## References

1. Koolen, D.A., Vissers, L.E., Pfundt, R., de Leeuw, N., Knight, S.J., Regan, R., Kooy, R.F., Reyniers, E., Romano, C., Fichera, M., et al. (2006). A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. *Nat. Genet.* **38**, 999–1001.
2. Kleefstra, T., Brunner, H.G., Amiel, J., Oudakker, A.R., Nillesen, W.M., Magee, A., Geneviève, D., Cormier-Daire, V., van Esch, H., Fryns, J.P., et al. (2006). Loss-of-function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome. *Am. J. Hum. Genet.* **79**, 370–377.
3. Kalscheuer, V.M., Tao, J., Donnelly, A., Hollway, G., Schwinger, E., Kübart, S., Menzel, C., Hoeltzenbein, M., Tommerup, N., Eyre, H., et al. (2003). Disruption of the serine/threonine kinase 9 gene causes severe X-linked infantile spasms and mental retardation. *Am. J. Hum. Genet.* **72**, 1401–1411.
4. Kalscheuer, V.M., FitzPatrick, D., Tommerup, N., Bugge, M., Niebuhr, E., Neumann, L.M., Tzschach, A., Shoichet, S.A., Menzel, C., Erdogan, F., et al. (2007). Mutations in autism susceptibility candidate 2 (AUTS2) in patients with mental retardation. *Hum. Genet.* **121**, 501–509.
5. Kim, H.G., Ahn, J.W., Kurth, I., Ullmann, R., Kim, H.T., Kulkharya, A., Ha, K.S., Itokawa, Y., Meliciani, I., Wenzel, W., et al. (2010). WDR11, a WD protein that interacts with transcription factor EMX1, is mutated in idiopathic hypogonadotropic hypogonadism and Kallmann syndrome. *Am. J. Hum. Genet.* **87**, 465–479.
6. Kim, H.G., Kishikawa, S., Higgins, A.W., Seong, I.S., Donovan, D.J., Shen, Y., Lally, E., Weiss, L.A., Najm, J., Kutsche, K., et al. (2008). Disruption of neurexin 1 associated with autism spectrum disorder. *Am. J. Hum. Genet.* **82**, 199–207.
7. Potocki, L., and Shaffer, L.G. (1996). Interstitial deletion of 11(p11.2p12): A newly described contiguous gene deletion syndrome involving the gene for hereditary multiple exostoses (EXT2). *Am. J. Med. Genet.* **62**, 319–325.
8. Shaffer, L.G., Hecht, J.T., Ledbetter, D.H., and Greenberg, F. (1993). Familial interstitial deletion 11(p11.12p12) associated with parietal foramina, brachymicrocephaly, and mental retardation. *Am. J. Med. Genet.* **45**, 581–583.
9. Stickens, D., Clines, G., Burbee, D., Ramos, P., Thomas, S., Hogue, D., Hecht, J.T., Lovett, M., and Evans, G.A. (1996). The EXT2 multiple exostoses gene defines a family of putative tumour suppressor genes. *Nat. Genet.* **14**, 25–32.
10. Mavrogiannis, L.A., Antonopoulou, I., Baxová, A., Kutílek, S., Kim, C.A., Sugayama, S.M., Salamanca, A., Wall, S.A., Morriss-Kay, G.M., and Wilkie, A.O. (2001). Haploinsufficiency of the human homeobox gene ALX4 causes skull ossification defects. *Nat. Genet.* **27**, 17–18.
11. Wu, Y.Q., Badano, J.L., McCaskill, C., Vogel, H., Potocki, L., and Shaffer, L.G. (2000). Haploinsufficiency of ALX4 as a potential cause of parietal foramina in the 11p11.2 contiguous gene-deletion syndrome. *Am. J. Hum. Genet.* **67**, 1327–1332.
12. Dollfus, H., Joanny-Flinois, O., Doco-Fenzy, M., Veyre, L., Joanny-Flinois, L., Khoury, M., Jonveaux, P., Abitbol, M., and Dufier, J.L. (1998). Gillespie syndrome phenotype with a t(X;11)(p22.32;p12) de novo translocation. *Am. J. Ophthalmol.* **125**, 397–399.
13. Fantes, J.A., Boland, E., Ramsay, J., Donnai, D., Splitt, M., Goodship, J.A., Stewart, H., Whiteford, M., Gautier, P., Harewood, L., et al. (2008). FISH mapping of de novo apparently balanced chromosome rearrangements identifies characteristics associated with phenotypic abnormality. *Am. J. Hum. Genet.* **82**, 916–926.
14. Lan, F., Collins, R.E., De Cegli, R., Alpatov, R., Horton, J.R., Shi, X., Gozani, O., Cheng, X., and Shi, Y. (2007). Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature* **448**, 718–722.
15. Wakui, K., Gregato, G., Ballif, B.C., Glotzbach, C.D., Bailey, K.A., Kuo, P.L., Sue, W.C., Sheffield, L.J., Irons, M., Gomez, E.G., et al. (2005). Construction of a natural panel of 11p11.2 deletions and further delineation of the critical region involved in Potocki-Shaffer syndrome. *Eur. J. Hum. Genet.* **13**, 528–540.
16. Wuyts, W., Waeber, G., Meinecke, P., Schüler, H., Goecke, T.O., Van Hul, W., and Bartsch, O. (2004). Proximal 11p deletion syndrome (P11pDS): Additional evaluation of the clinical and molecular aspects. *Eur. J. Hum. Genet.* **12**, 400–406.
17. Anderson, M.A., and Gusella, J.F. (1984). Use of cyclosporin A in establishing Epstein-Barr virus-transformed human lymphoblastoid cell lines. *In Vitro* **20**, 856–858.
18. Romeike, B.F., and Wuyts, W. (2007). Proximal chromosome 11p contiguous gene deletion syndrome phenotype: Case report and review of the literature. *Clin. Neuropathol.* **26**, 1–11.
19. Miller, D.T., Shen, Y., and Wu, B.L. (2008). Oligonucleotide microarrays for clinical diagnosis of copy number variation. *Curr. Protoc. Hum. Genet.* **58**, 8.12.1–8.12.17.

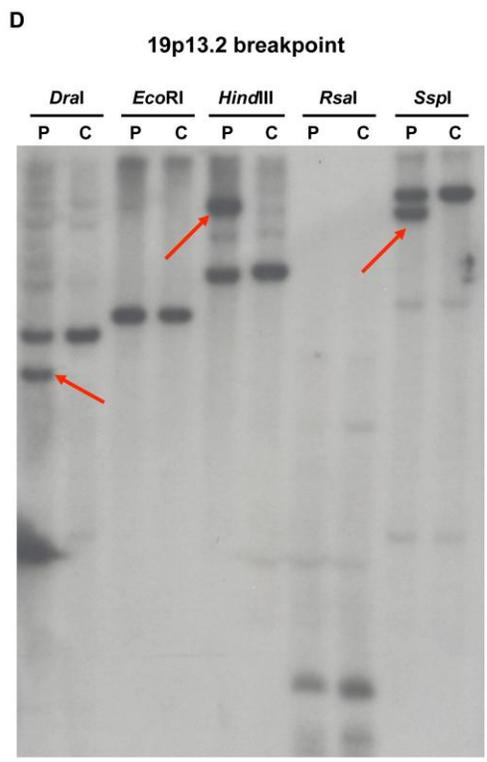
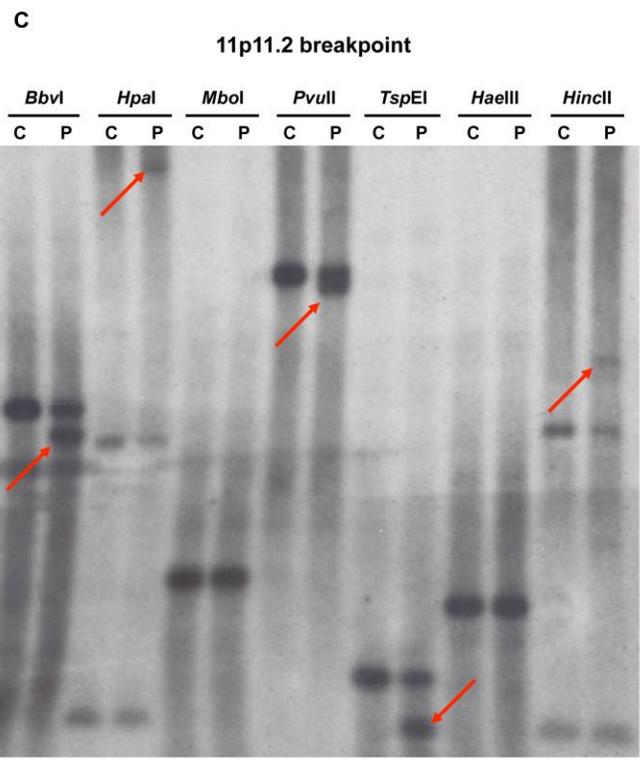
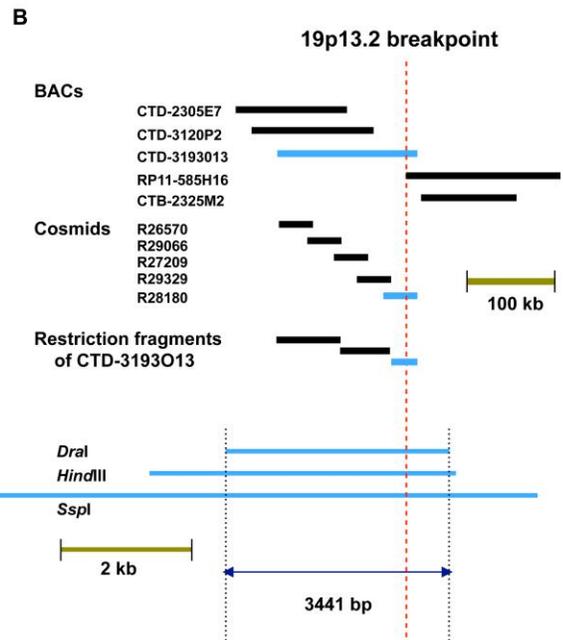
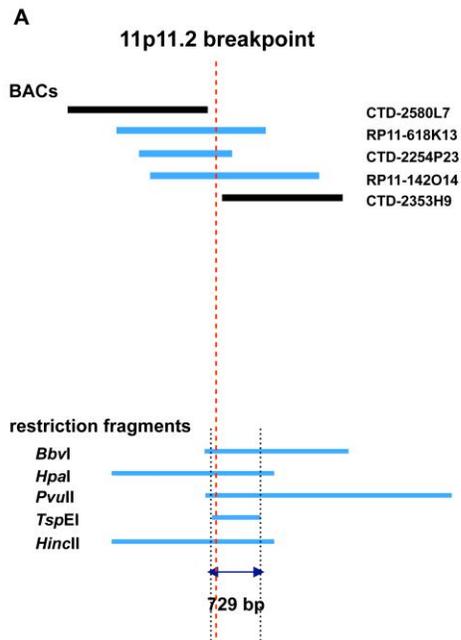
20. Obholz, K.L., Akopyan, A., Waymire, K.G., and MacGregor, G.R. (2006). FNDC3A is required for adhesion between spermatids and Sertoli cells. *Dev. Biol.* **298**, 498–513.
21. Arkesteijn, G., Jumelet, E., Hagenbeek, A., Smit, E., Slater, R., and Martens, A. (1999). Reverse chromosome painting for the identification of marker chromosomes and complex translocations in leukemia. *Cytometry* **35**, 117–124.
22. Chen, W., Kalscheuer, V., Tzschach, A., Menzel, C., Ullmann, R., Schulz, M.H., Erdogan, F., Li, N., Kijas, Z., Arkesteijn, G., et al. (2008). Mapping translocation breakpoints by next-generation sequencing. *Genome Res.* **18**, 1143–1149.
23. Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A., and Lukyanov, S.A. (1995). An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.* **23**, 1087–1088.
24. Cheung, V.G., Nowak, N., Jang, W., Kirsch, I.R., Zhao, S., Chen, X.N., Furey, T.S., Kim, U.J., Kuo, W.L., Olivier, M., et al; BAC Resource Consortium. (2001). Integration of cytogenetic landmarks into the draft sequence of the human genome. *Nature* **409**, 953–958.
25. Bartsch, O., Schmidt, S., Richter, M., Morlot, S., Seemanová, E., Wiebe, G., and Rasi, S. (2005). DNA sequencing of CREBBP demonstrates mutations in 56% of patients with Rubinstein-Taybi syndrome (RSTS) and in another patient with incomplete RSTS. *Hum. Genet.* **117**, 485–493.
26. Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953.
27. Shi, Y.J., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y. (2005). Regulation of LSD1 histone demethylase activity by its associated factors. *Mol. Cell* **19**, 857–864.
28. Iwase, S., Januma, A., Miyamoto, K., Shono, N., Honda, A., Yanagisawa, J., and Baba, T. (2004). Characterization of BHC80 in BRAF-HDAC complex, involved in neuron-specific gene repression. *Biochem. Biophys. Res. Commun.* **322**, 601–608.
29. Park, H.C., Kim, C.H., Bae, Y.K., Yeo, S.Y., Kim, S.H., Hong, S.K., Shin, J., Yoo, K.W., Hibi, M., Hirano, T., et al. (2000). Analysis of upstream elements in the HuC promoter leads to the establishment of transgenic zebrafish with fluorescent neurons. *Dev. Biol.* **227**, 279–293.
30. Kim, C.H., Ueshima, E., Muraoka, O., Tanaka, H., Yeo, S.Y., Huh, T.L., and Miki, N. (1996). Zebrafish elav/HuC homologue as a very early neuronal marker. *Neurosci. Lett.* **216**, 109–112.
31. Kim, C.H., Oda, T., Itoh, M., Jiang, D., Artinger, K.B., Chandrasekharappa, S.C., Driever, W., and Chitnis, A.B. (2000). Repressor activity of Headless/Tcf3 is essential for vertebrate head formation. *Nature* **407**, 913–916.
32. Higgins, A.W., Alkuraya, F.S., Bosco, A.F., Brown, K.K., Bruns, G.A., Donovan, D.J., Eisenman, R., Fan, Y., Farra, C.G., Ferguson, H.L., et al. (2008). Characterization of apparently balanced chromosomal rearrangements from the developmental genome anatomy project. *Am. J. Hum. Genet.* **82**, 712–722.
33. Talkowski, M.E., Ernst, C., Heilbut, A., Chiang, C., Hanscom, C., Lindgren, A., Kirby, A., Liu, S., Muddukrishna, B., Ohsumi, T.K., et al. (2011). Next-generation sequencing strategies enable routine detection of balanced chromosome rearrangements for clinical diagnostics and genetic research. *Am. J. Hum. Genet.* **88**, 469–481.
34. Hall, C.R., Wu, Y., Shaffer, L.G., and Hecht, J.T. (2001). Familial case of Potocki-Shaffer syndrome associated with microdeletion of EXT2 and ALX4. *Clin. Genet.* **60**, 356–359.
35. Mavrogiannis, L.A., Taylor, I.B., Davies, S.J., Ramos, F.J., Olivares, J.L., and Wilkie, A.O. (2006). Enlarged parietal foramina caused by mutations in the homeobox genes ALX4 and MSX2: From genotype to phenotype. *Eur. J. Hum. Genet.* **14**, 151–158.
36. Wuyts, W., Di Gennaro, G., Bianco, F., Wauters, J., Morocutti, C., Pierelli, F., Bossuyt, P., Van Hul, W., and Casali, C. (1999). Molecular and clinical examination of an Italian DEFECT11 family. *Eur. J. Hum. Genet.* **7**, 579–584.
37. Ferreira, M.A., O'Donovan, M.C., Meng, Y.A., Jones, I.R., Ruderfer, D.M., Jones, L., Fan, J., Kirov, G., Perlis, R.H., Green, E.K., et al; Wellcome Trust Case Control Consortium. (2008). Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat. Genet.* **40**, 1056–1058.
38. O'Donovan, M.C., Craddock, N., Norton, N., Williams, H., Peirce, T., Moskvina, V., Nikolov, I., Hamshere, M., Carroll, L., Georgieva, L., et al; Molecular Genetics of Schizophrenia Collaboration. (2008). Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat. Genet.* **40**, 1053–1055.
39. Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'Donovan, M.C., Sullivan, P.F., and Sklar, P.; International Schizophrenia Consortium. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752.
40. Smith, E.N., Bloss, C.S., Badner, J.A., Barrett, T., Belmonte, P.L., Berrettini, W., Byerley, W., Coryell, W., Craig, D., Edenberg, H.J., et al. (2009). Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol. Psychiatry* **14**, 755–763.
41. Sullivan, P.F., Lin, D., Tzeng, J.Y., van den Oord, E., Perkins, D., Stroup, T.S., Wagner, M., Lee, S., Wright, F.A., Zou, F., et al. (2008). Genomewide association for schizophrenia in the CATIE study: Results of stage 1. *Mol. Psychiatry* **13**, 570–584.
42. International Schizophrenia Consortium. (2008). Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455**, 237–241.
43. Cooper, G.M., Coe, B.P., Girirajan, S., Rosenfeld, J.A., Vu, T.H., Baker, C., Williams, C., Stalker, H., Hamid, R., Hannig, V., et al. (2011). A copy number variation morbidity map of developmental delay. *Nat. Genet.* **43**, 838–846.
44. 1000 Genomes Project Consortium. (2010). A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073.
45. Iwase, S., Shono, N., Honda, A., Nakanishi, T., Kashiwabara, S., Takahashi, S., and Baba, T. (2006). A component of BRAF-HDAC complex, BHC80, is required for neonatal survival in mice. *FEBS Lett.* **580**, 3129–3135.
46. Choi, J., Dong, L., Ahn, J., Dao, D., Hammerschmidt, M., and Chen, J.N. (2007). FoxH1 negatively modulates flk1 gene expression and vascular formation in zebrafish. *Dev. Biol.* **304**, 735–744.
47. Schilling, T.F., Piotrowski, T., Grandel, H., Brand, M., Heisenberg, C.P., Jiang, Y.J., Beuchle, D., Hammerschmidt, M., Kane, D.A., Mullins, M.C., et al. (1996). Jaw and branchial arch mutants in zebrafish I: Branchial arches. *Development* **123**, 329–344.

48. Wilkie, A.O., and Morriss-Kay, G.M. (2001). Genetics of craniofacial development and malformation. *Nat. Rev. Genet.* 2, 458–468.
49. Bartsch, O., Wuyts, W., Van Hul, W., Hecht, J.T., Meinecke, P., Hogue, D., Werner, W., Zabel, B., Hinkel, G.K., Powell, C.M., et al. (1996). Delineation of a contiguous gene syndrome with multiple exostoses, enlarged parietal foramina, craniofacial dysostosis, and mental retardation, caused by deletions in the short arm of chromosome 11. *Am. J. Hum. Genet.* 58, 734–742.
50. Yamamoto, T., Akaboshi, S., Ninomiya, H., and Nanba, E. (2001). DEFECT 11 syndrome associated with agenesis of the corpus callosum. *J. Med. Genet.* 38, E5.
51. Dodé, C., Levilliers, J., Dupont, J.M., De Paepe, A., Le Dû, N., Soussi-Yanicostas, N., Coimbra, R.S., Delmaghani, S., Compain-Nouaille, S., Baverel, F., et al. (2003). Loss-of-function mutations in FGFR1 cause autosomal dominant Kallmann syndrome. *Nat. Genet.* 33, 463–465.
52. Vissers, L.E., van Ravenswaaij, C.M., Admiraal, R., Hurst, J.A., de Vries, B.B., Janssen, I.M., van der Vliet, W.A., Huys, E.H., de Jong, P.J., Hamel, B.C., et al. (2004). Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat. Genet.* 36, 955–957.
53. Wuyts, W., Cleiren, E., Homfray, T., Rasore-Quartino, A., Vanhoenacker, F., and Van Hul, W. (2000). The ALX4 homeobox gene is mutated in patients with ossification defects of the skull (foramina parietalia permagna, OMIM 168500). *J. Med. Genet.* 37, 916–920.
54. Gécz, J. (2004). The molecular basis of intellectual disability: Novel genes with naturally occurring mutations causing altered gene expression in the brain. *Front. Biosci.* 9, 1–7.
55. Tassabehji, M., Hammond, P., Kamiloff-Smith, A., Thompson, P., Thorgeirsson, S.S., Durkin, M.E., Popescu, N.C., Hutton, T., Metcalfe, K., Rucka, A., et al. (2005). GTF2IRD1 in craniofacial development of humans and mice. *Science* 310, 1184–1187.
56. Ma, W.J., and Furneaux, H. (1997). Localization of the human HuR gene to chromosome 19p13.2. *Hum. Genet.* 99, 32–33.
57. Ma, W.J., Cheng, S., Campbell, C., Wright, A., and Furneaux, H. (1996). Cloning and characterization of HuR, a ubiquitously expressed Elav-like protein. *J. Biol. Chem.* 271, 8144–8151.
58. Prakash, S.K., Paylor, R., Jenna, S., Lamarche-Vane, N., Armstrong, D.L., Xu, B., Mancini, M.A., and Zoghbi, H.Y. (2000). Functional analysis of ARHGAP6, a novel GTPase-activating protein for RhoA. *Hum. Mol. Genet.* 9, 477–488.
59. Pasillas, M.P., Shah, M., and Kamps, M.P. (2011). NSD1 PHD domains bind methylated H3K4 and H3K9 using interactions disrupted by point mutations in human sotos syndrome. *Hum. Mutat.* 32, 292–298.
60. Iwase, S., Lan, F., Bayliss, P., de la Torre-Ubieta, L., Huarte, M., Qi, H.H., Whetstine, J.R., Bonni, A., Roberts, T.M., and Shi, Y. (2007). The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. *Cell* 128, 1077–1088.
61. Feng, W., Yonezawa, M., Ye, J., Jenuwein, T., and Grummt, I. (2010). PHF8 activates transcription of rRNA genes through H3K4me3 binding and H3K9me1/2 demethylation. *Nat. Struct. Mol. Biol.* 17, 445–450.
62. Qi, H.H., Sarkissian, M., Hu, G.Q., Wang, Z., Bhattacharjee, A., Gordon, D.B., Gonzales, M., Lan, F., Ongusaha, P.P., Huarte, M., et al. (2010). Histone H4K20/H3K9 demethylase PHF8 regulates zebrafish brain and craniofacial development. *Nature* 466, 503–507.
63. Hakimi, M.A., Bochar, D.A., Chenoweth, J., Lane, W.S., Mandel, G., and Shiekhattar, R. (2002). A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. *Proc. Natl. Acad. Sci. USA* 99, 7420–7425.
64. Hakimi, M.A., Dong, Y., Lane, W.S., Speicher, D.W., and Shiekhattar, R. (2003). A candidate X-linked mental retardation gene is a component of a new family of histone deacetylase-containing complexes. *J. Biol. Chem.* 278, 7234–7239.
65. Scheer, M.P., van der Maarel, S., Kübart, S., Schulz, A., Wirth, J., Schweiger, S., Ropers, H., and Nothwang, H.G. (2000). DXS6673E encodes a predominantly nuclear protein, and its mouse ortholog DXHS6673E is alternatively spliced in a developmental- and tissue-specific manner. *Genomics* 63, 123–132.
66. van der Maarel, S.M., Scholten, I.H., Huber, I., Philippe, C., Suijkerbuijk, R.F., Gilgenkrantz, S., Kere, J., Cremers, F.P., and Ropers, H.H. (1996). Cloning and characterization of DXS6673E, a candidate gene for X-linked mental retardation in Xq13.1. *Hum. Mol. Genet.* 5, 887–897.
67. Kleefstra, T., Smidt, M., Banning, M.J., Oudakker, A.R., Van Esch, H., de Brouwer, A.P., Nillesen, W., Sistermans, E.A., Hamel, B.C., de Bruijn, D., et al. (2005). Disruption of the gene Euchromatin Histone Methyl Transferase1 (Eu-HMTase1) is associated with the 9q34 subtelomeric deletion syndrome. *J. Med. Genet.* 42, 299–306.
68. Ogawa, H., Ishiguro, K., Gaubatz, S., Livingston, D.M., and Nakatani, Y. (2002). A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. *Science* 296, 1132–1136.
69. Shi, Y., Sawada, J., Sui, G., Affar, B., Whetstine, J.R., Lan, F., Ogawa, H., Luke, M.P., Nakatani, Y., and Shi, Y. (2003). Coordinated histone modifications mediated by a CtBP co-repressor complex. *Nature* 422, 735–738.
70. Halgren, C., Kjaergaard, S., Bak, M., Hansen, C., El-Schich, Z., Anderson, C., Henriksen, K., Hjalgrim, H., Kirchhoff, M., Bijlsma, E., et al. (2011). Corpus callosum abnormalities, intellectual disability, speech impairment, and autism in patients with haploinsufficiency of ARID1B. *Clin. Genet.*
71. Li, X.S., Trojer, P., Matsumura, T., Treisman, J.E., and Tanese, N. (2010). Mammalian SWI/SNF—a subunit BAF250/ARID1 is an E3 ubiquitin ligase that targets histone H2B. *Mol. Cell Biol.* 30, 1673–1688.
72. Hoyer, J., Ekici, A.B., Ende, S., Popp, B., Zweier, C., Wiesener, A., Wohlleber, E., Dufke, A., Rossier, E., Petsch, C., et al. (2012). Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. *Am. J. Hum. Genet.* 90, 565–572.
73. Santen, G.W., Aten, E., Sun, Y., Almomani, R., Gilissen, C., Nielsen, M., Kant, S.G., Snoeck, I.N., Peeters, E.A., Hillhorst-Hofstee, Y., et al. (2012). Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat. Genet.* 44, 379–380.
74. Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., Kaname, T., Naritomi, K., Kawame, H., Wakui, K., et al. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat. Genet.* 44, 376–378.

## **Supplemental Data**

### **Translocations Disrupting *PHF21A* in the Potocki-Shaffer-Syndrome Region Are Associated with Intellectual Disability and Craniofacial Anomalies**

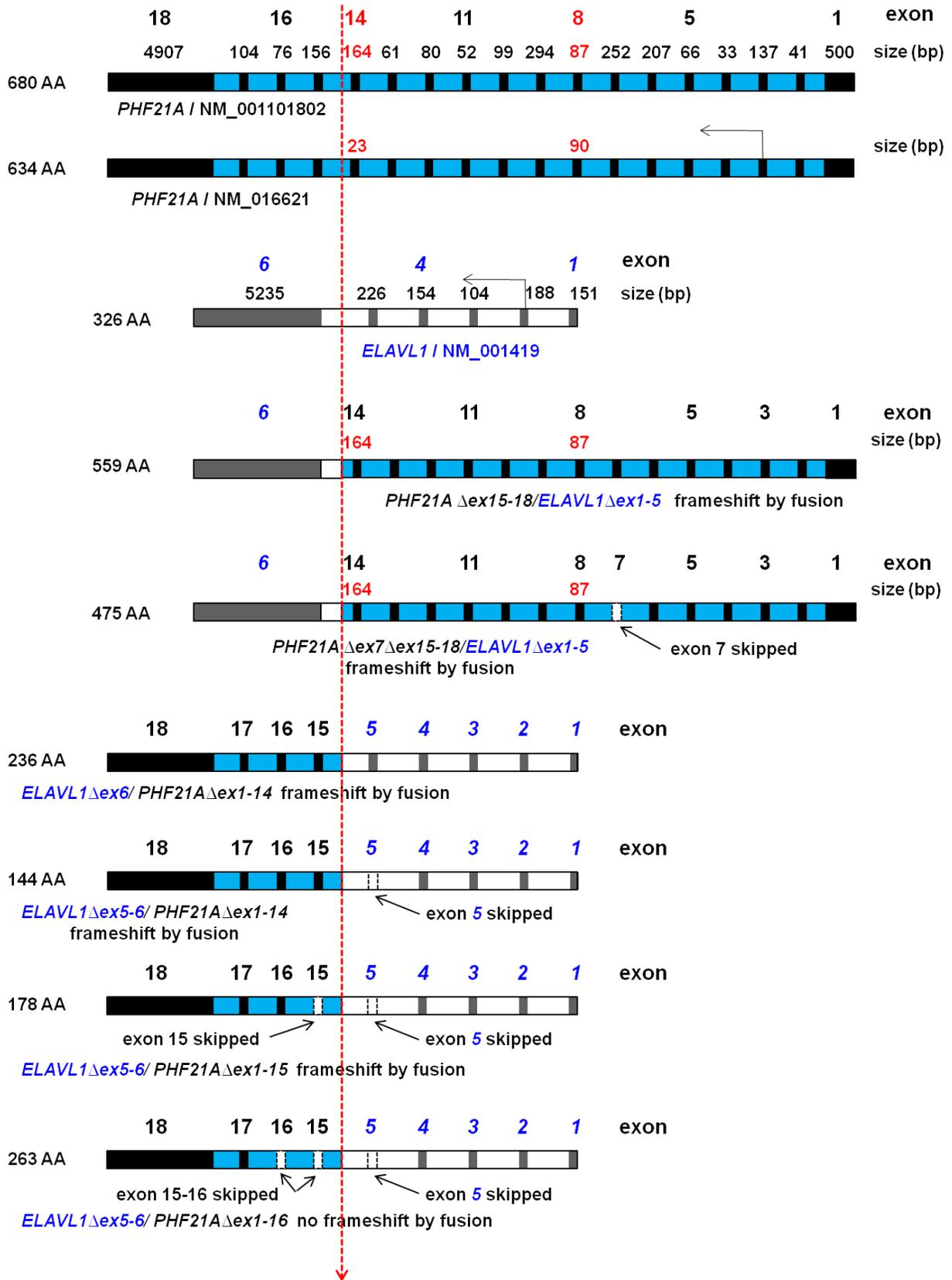
Hyung-Goo Kim, Hyun-Taek Kim, Natalia T. Leach, Fei Lan, Reinhard Ullmann, Asli Silahdaroglu, Ingo Kurth, Anja Nowka, Ihn Sik Seong, Yiping Shen, Michael E. Talkowski, Douglas Ruderfer, Ji-Hyun Lee, Caron Glotzbach, Kyungsoo Ha, Susanne Kjaergaard, Alex V. Levin, Bernd F. Romeike, Tjitske Kleefstra, Oliver Bartsch, Sarah H. Elsea, Ethylin Wang Jabs, Marcy E. MacDonald, David J. Harris, Bradley J. Quade, Hans-Hilger Ropers, Lisa G. Shaffer, Kerstin Kutsche, Lawrence C. Layman, Niels Tommerup, Vera M. Kalscheuer, Yang Shi, Cynthia C. Morton, Cheol-Hee Kim, and James F. Gusella



**Figure S1. Mapping 11p11.2 and 19p13.2 Breakpoints of DGAP012 with t(11;19)(p11.2;p13.2)**

(A and B) The respective breakpoints in chromosomes 11 and 19 are shown as dashed red vertical lines. BAC clones used in FISH experiments detect both der(11) and der(19) (blue lines) or only one of the derivative chromosomes (black lines) in addition to either chromosome 11 or chromosome 19. Therefore, the clones in blue span the corresponding breakpoint. In refining the 19p13.2 breakpoint by FISH, additional cosmid clones and three restriction fragments from a breakpoint spanning clone were used. At 11p11.2 the breakpoint was narrowed to ~14 kb between BAC clones CTD-2580L7 and CTD-2353H9, while the breakpoint at 19p13.2 was refined to a 30 kb *Sna*BI restriction fragment from BAC CTD-3193O13. Below the BACs are shown restriction fragments found in DNA-blotting experiments (see panels C and D) that confine the breakpoints to small segments of 11p11.2 and 19p13.2, respectively.

(C and D) Genomic DNA blots hybridized with probes from 11p11.2 and 19p13.2 breakpoint regions, respectively. Each lane contains genomic DNA digested with the designated restriction enzyme from either DGAP012 (P) or a normal control (C). Additional bands shown with red arrows in the P lanes indicate novel restriction fragments detected by probe D012-B in Panel C and D012-F in Panel D (Table S3). These additional bands were not seen in the control because they represent junction fragments composed of sequence of both chromosomes 11 and 19. Based on the junction fragments detected in DGAP012, the breakpoint of 11p11.2 and 19p13.2 was narrowed to 729 bp and 3441 bp, respectively, as shown in panels A and B.



**Figure S2. Two Splice Variants of *PHF21A* and Six Amplified Fusion Transcripts with *ELAVL1* from DGAP012 with t(11;19)(p11.2;p13.2)**

As both *PHF21A* and *ELAVL1* are transcribed in a centromeric to telomeric direction and the translocation predicts two fusion genes that might produce novel transcripts (Fig. 2A), we performed RT-PCR on lymphoblastoid cell line RNA, using *PHF21A*- and *ELAVL1*-derived forward and reverse primers. On the der(11), *PHF21A* $\Delta$ ex15-18/*ELAVL1* $\Delta$ ex1-5 would result from fusion of *PHF21A* exons 1-14 to exon 6 of *ELAVL1* and on the der(19), *ELAVL1* $\Delta$ ex6/*PHF21A* $\Delta$ ex1-14 would result from fusion of exons 1-5 of *ELAVL1* to exons 15-18 of *PHF21A*. The former predicts a frameshift, leading to a fusion protein of 559 amino acids, 483 from *PHF21A* and an aberrant 76 before a premature stop TGA in exon 6 of *ELAVL1*, while the latter predicts a fusion protein of 236 amino acids containing residues 1-218 from *ELAVL1*, followed by a frameshift to 18 novel residues prior to a novel stop TGA in exon 15 of *PHF21A*. To corroborate this *in silico* analysis, we performed RT-PCR, obtaining specific products that represent six different fusion transcripts due to alternative splicing from the two fusion genes; no cDNA could be amplified using a reverse transcribed control RNA with either of the primer pairs.

For the fusion gene on der(19), which would be transcribed from the *ELAVL1* promotor and could encompass exons 1-5 of *ELAVL1* and exons 15-18 of *PHF21A*, we amplified a mixture of fragments of 0.9-1.4 kb from DGAP012 cDNA. When these were cloned and sequenced, we identified four different transcripts, depending upon alternative splicing of exon 5 of *ELAVL1* and exons 15 and 16 of *PHF21A*: *ELAVL1* $\Delta$ ex6/*PHF21A* $\Delta$ ex1-14 (236 aa), *ELAVL1* $\Delta$ ex5-6/*PHF21A* $\Delta$ ex1-14 (144 aa), *ELAVL1* $\Delta$ ex5-6/*PHF21A* $\Delta$ ex1-15 (178 aa), and *ELAVL1* $\Delta$ ex5-6/*PHF21A* $\Delta$ ex1-16 (263 aa). For the fusion gene from der(11), which would be transcribed from the *PHF21A* promotor, the transcript could theoretically include *PHF21A* exons 1-14 and *ELAVL1* exon 6. As in the case of the der(19) fusion gene, multiple RT-PCR products were observed, ranging from 1.5-1.8 kb, and subcloning and sequencing revealed evidence of two transcripts due to novel alternative splicing of *PHF21A* exon 7 in the longer known transcript (NM\_001101802.1) of *PHF21A*: *PHF21A* $\Delta$ ex15-18/*ELAVL1* $\Delta$ ex1-5 (559 aa) and *PHF21A* $\Delta$ ex7 $\Delta$ ex15-18/*ELAVL1* $\Delta$ ex1-5 (475 aa).

The relevant exon representation in the fusion transcripts of *PHF21A* and *ELAVL1* is shown here with the location of the translocation breakpoint depicted by a dashed red arrow. *PHF21A* is shown in blue background with exons in black, whereas *ELAVL1* is shown in white backdrop with gray exons. The start and direction of transcription are indicated by an arrow above the corresponding exons. Note that normal variation in splice site utilization in *PHF21A* results in two different sizes for each of exons 8 and 14 depicted in red. In total, we amplified six different fusion transcripts due to variable exon skipping. All but one (*ELAVL1* $\Delta$ ex5-6/*PHF21A* $\Delta$ ex1-16) of these six entailed frameshifts that introduced premature stop codons. Thus, DGAP012 lymphoblasts have the potential to express full-length *PHF21A* from the nontranslocated allele, and a variety of fusions of *PHF21A* and *ELAVL1* from the chromosomes derived by translocation, though whether any of the protein products would be stable is unknown.

For RT-PCR, total RNA was isolated from subject and control lymphoblastoid cell lines with the RNeasy Mini Kit (Qiagen, Valencia, CA). Reverse transcription of total RNA (1  $\mu$ g) was undertaken by using either random hexanucleotide priming and Superscript II (Gibco BRL,

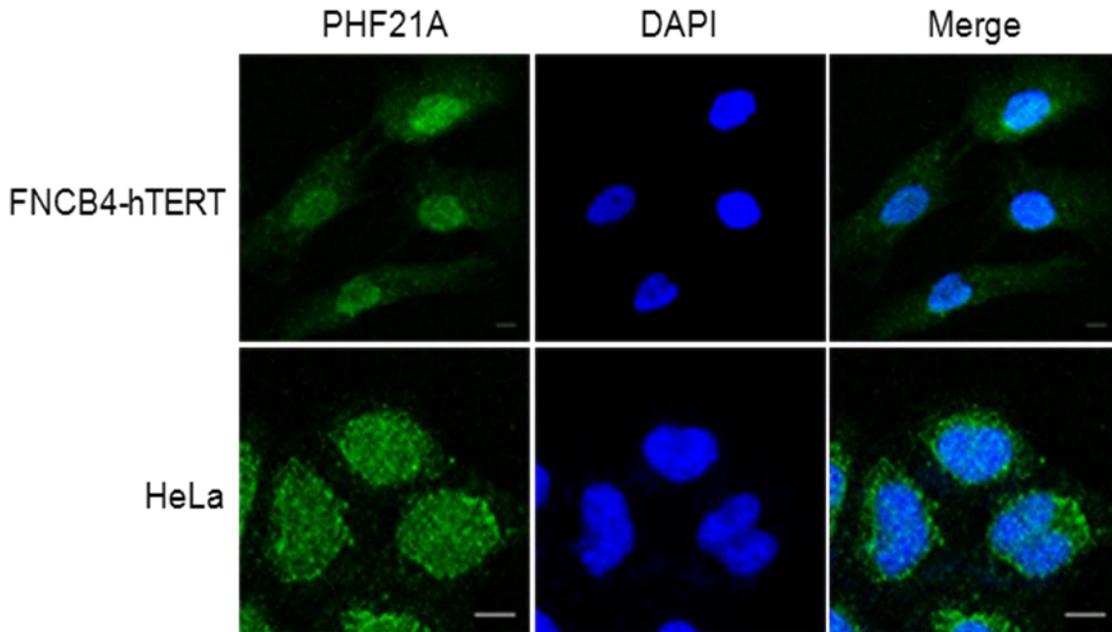
Gaithersburg, MD) or the SMART-PCR cDNA synthesis kit (Clontech, Palo Alto, CA) according to the protocols provided. In each experiment, DNA contamination was excluded by absence of a PCR product in the sample without reverse transcriptase, amplified under the same conditions as the reverse transcribed RNA sample, and by design of primers to cross multiple exons. Nested PCR was carried out using *Pfu* polymerase (Gibco BRL) with the following primer sets, annealing at 56 °C for 30 seconds with an extension for two minutes:

*PHF21A-ELAVLI:*

TCGCTGCTGCTGCTGAGAGG + AACACTTGTGAAAATTGGCGC;  
ATGGAGTTGCAGACTCTACAGGAG + TCATAGTTTGTTCATGGTCACAAAGC.

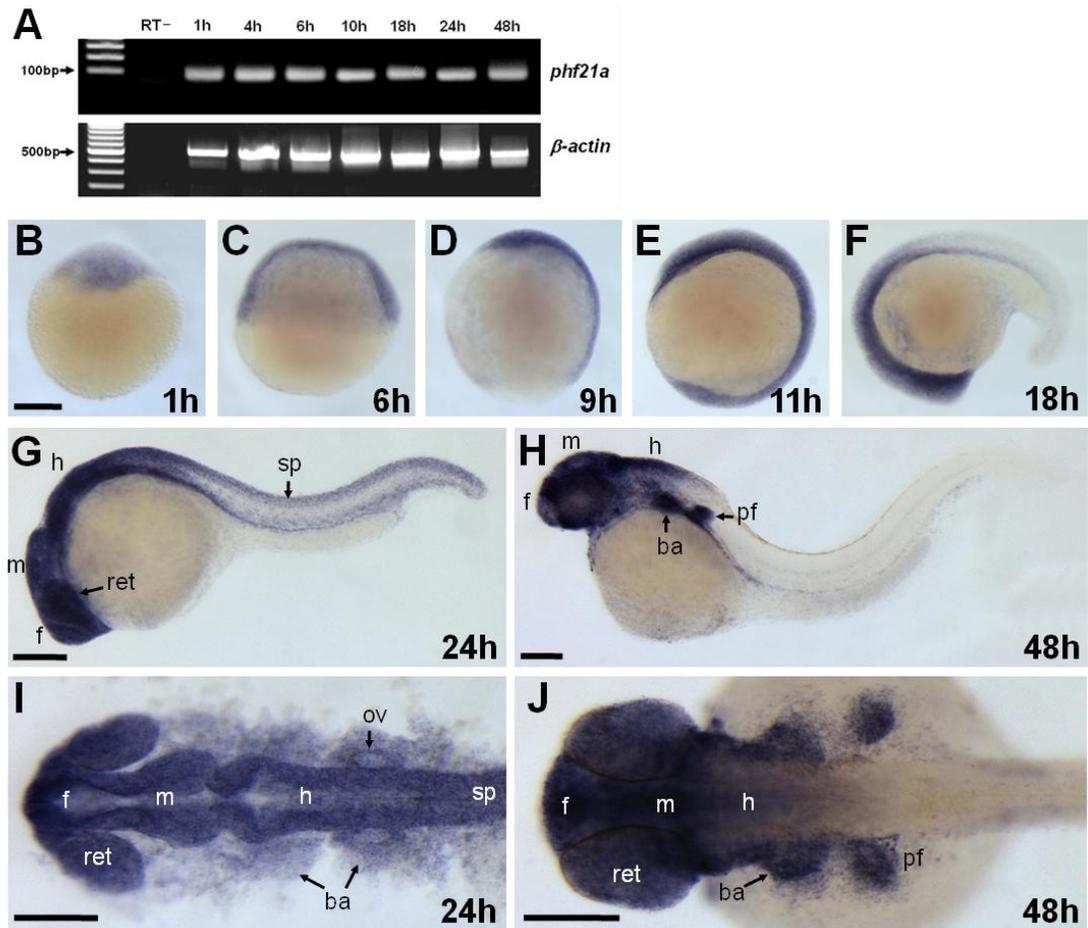
*ELAVLI-PHF21A:*

CGCATCCAGATTTTTGAAAAATACA + CCAAAGAATTCTGCACTTTCCAG;  
CGCATCCAGATTTTTGAAAAATACA + GAGTCTTCAGTGTTCTGTTCTCCTTG.



**Figure S3. Immunofluorescence Analysis of PHF21A Localization in FNCB4-hTERT and HeLa Cells**

Subcellular localization of PHF21A in FNCB4-hTERT and HeLa cells was measured by immunofluorescence analysis. FNCB4-hTERT and HeLa cells were cultured on glass coverslips overnight and stained with anti-PHF21A antibody. PHF21A mainly showed nuclear localization, but a small amount of PHF21A signal was also detected in the cytoplasm in both cell lines. Nuclei were visualized with DAPI staining. Scale bars represent 10  $\mu$ m.



**Figure S4. *phf21a* Expression during Zebrafish Early Development**

Temporal and spatial expression of *phf21a* analyzed by RT-PCR and whole-mount *in situ* hybridization.

(A) *phf21a* transcript is detected from the 1-cell stage (1h) to the long-pec stage (48h) in zebrafish. *ACTB* is used as a loading control.

(B–E) *phf21a* is ubiquitously expressed from the early cleavage stage to three-somite stage (11 h).

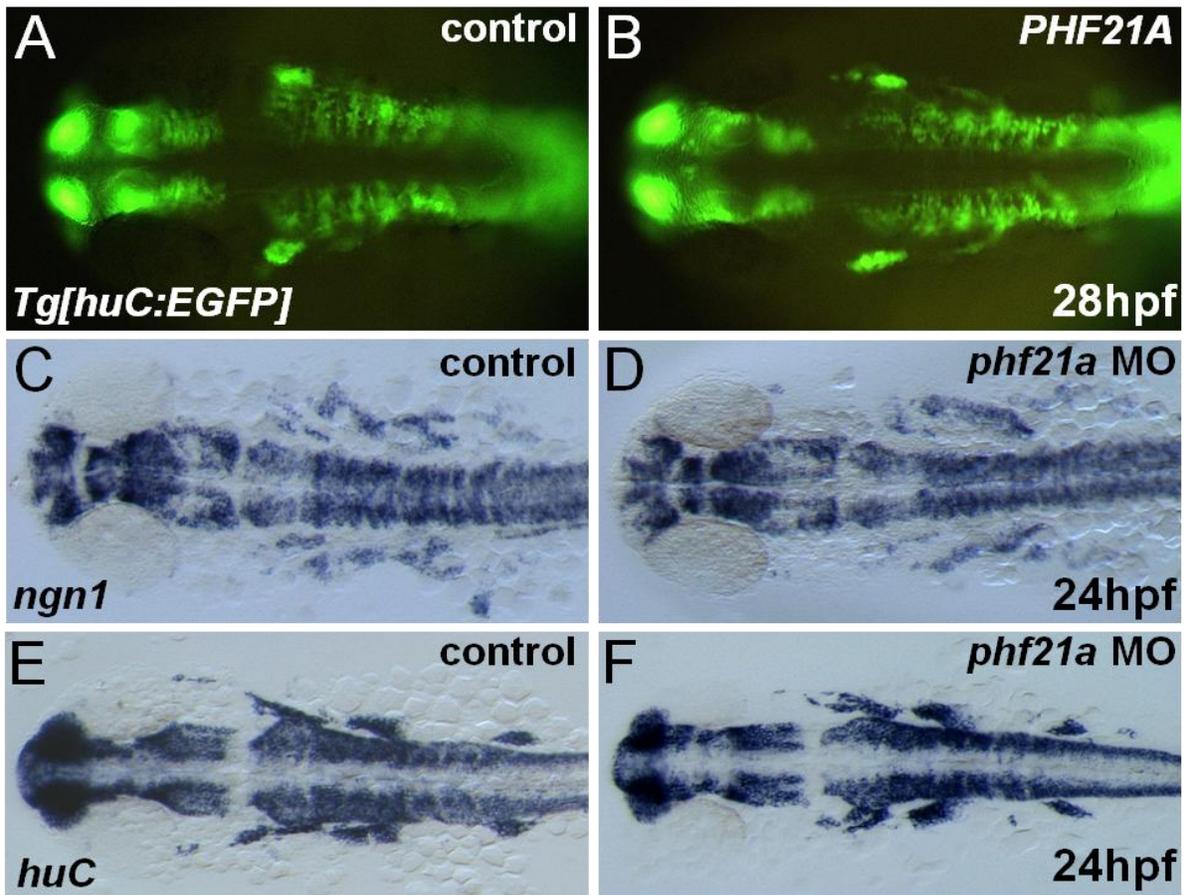
(F) In lateral view, *phf21a* expression in the head region is increased from the late somitogenesis stage.

(G) At 24 hpf, *phf21a* transcripts are detected both in head and trunk regions.

(H) In lateral view, *phf21a* is expressed at high levels in the head region at 48 hpf.

(I) In dorsal view at 24 hpf, transcripts of *phf21a* are detected in otic vesicle and the CNS including retina, forebrain, midbrain, hindbrain, and spinal cord.

(J) In dorsal view at 48 hpf, *phf21a* is strongly expressed in brain regions, pectoral fin, and branchial arches. Abbreviations: ba, branchial arches; f, forebrain; h, hindbrain; m, midbrain; ov, otic vesicle; pf, pectoral fin; ret, retina; sp, spinal cord. Scale bars = 200  $\mu$ m.



**Figure S5. Gain or Loss of *phf21a* Function in Zebrafish Does Not Affect Development of the Central Nervous System**

(A and B) Dorsal view of head region in *huC:EGFP* transgenic line at 28 hpf. CNS development visualized by GFP was similar between uninjected controls (A) and human *PHF21A* mRNA-injected embryos (B).

(C–F) Normal neurogenesis in *phf21a* knock-down embryo. Neurogenesis was confirmed as normal with both early neurogenic marker, *ngn1* (C and D), and late neuronal differentiation marker, *huC* (E and F), in both control (C and E) and *phf21a*-MO injected embryos (D and F).

**Table S1. BACs and a Cosmid Used for FISH Analysis of DGAP012 with t(11;19)(p11.2;p13.2)****BACs from Chromosome 11** (coordinates from NCBI Build 36/hg18 assembly)

No.	BAC ID	Start (bp)	End (bp)	Telomeric	Spanning	Centromeric
1.	RP11-36H11	36,353,273	36,513,773	X		
2.	RP11-187A08	36,634,404	36,634,544	X		
3.	RP11-79G4	41,558,233	41,710,577	X		
4.	RP11-12D19	43,681,668	43,842,756	X		
5.	RP11-150D18	41,858,282	42,020,207	X		
6.	RP11-18D13	45,034,241	45,034,572	X		
7.	RP11-102E22	44,098,956	44,271,172	X		
8.	RP11-12C11	44,879,169	45,046,093	X		
9.	RP11-206I1	45,047,400	45,224,812	X		
10.	RP11-10C6	45,120,194	45,272,892	X		
11.	RP11-703E9	45,198,333	45,345,575	X		
12.	RP11-430H10	45,341,690	45,521,415	X		
13.	RP11-586E3	45,403,601	45,571,439	X		
14.	RP11-958J22	45,406,042	45,589,724	X		
15.	RP11-495O11	45,552,909	45,753,243	X		
16.	CTD-2210P24	45,664,882	45,807,468	X		
17.	CTD-2580L7	45,752,305	45,911,861	X		
18.	RP11-618K13	45,804,176	45,975,298		X	
19.	CTD-2254P23	45,834,963	45,941,943		X	
20.	RP11-142O14	45,845,429	46,038,614		X	
21.	CTD-2353H9	45,925,745	46,063,695			X
22.	RP11-702F3	46,021,151	46,213,090			X
23.	RP11-425L10	46,359,782	46,429,924			X
24.	RP11-29O22	46,582,988	46,583,429			X
25.	RP11-39G12	46,582,998	46,737,312			X
26.	RP11-722K13	47,479,616	47,963,280			X
27.	RP11-56E13	47,989,033	47,989,377			X

**BACs and a Cosmid from Chromosome 19** (coordinates from NCBI Build 36/hg18 assembly)

No.	BAC ID	Start (bp)	End (bp)	Telomeric	Spanning	Centromeric
1.	RP11-75H6	902,642	1,095,485	X		
2.	RP11-49M3	2,418,857	2,419,340	X		
3.	RP11-54G9	3,006,444	3,006,786	X		
4.	RP11-211I3	4,889,356	4,889,698	X		
5.	CTB-25J19	6,970,231	7,076,170	X		
6.	CTD-2558K15	7,320,276	7,376,729	X		
7.	CTD-3214H19	7,540,075	7,718,036	X		
8.	CTD-2102F19	7,667,777	7,811,388	X		
9.	RP11-42J18	7,733,047	7,917,986	X		
10.	CTD-2305E7	7,742,144	7,870,622	X		
11.	CTD-3120P2	7,765,345	7,905,189	X		
12.	RP11-84C17	7,777,647	7,778,013	X		
13.	CTD-3193O13	7,789,674	7,950,444		X	
14.	LLNLF-138H2 (cosmid)	7,939,101	7,978,680			X
15.	CTB-2325M2	7,955,128	8,064,573			X
16.	CTD-3020H12	8,029,107	8,255,828			X
17.	CTD-2547N9	8,490,674	8,519,982			X
18.	CTB-2369P2	10,141,242	10,297,745			X
19.	CTC-510F12	11,174,082	11,296,402			X
20.	RP11-19I2	11,640,413	11,816,303			X
21.	CTD-3105H18	12,264,186	12,488,854			X
22.	RP11-31N2	12,553,162	12,553,279			X
23.	CTC-250I14	13,100,859	13,216,789			X
24.	RP11-201F4	13,686,126	13,686,376			X

**Table S2. BACs and Fosmids Used for FISH Analysis of MCN1762 with t(1;11)(p21.1;p11.2)****BACs from Chromosome 1** (coordinates from NCBI Build 36/hg18 assembly)

No.	BAC ID	Start (bp)	End (bp)	Telomeric	Spanning	Centromeric
1.	RP11-30M11	84,608,787	84,775,072	X		
2.	RP11-82O2	102,084,529	102,257,779	X		
3.	RP11-34I24	103,233,251	103,423,196	X		
4.	RP5-1108M17	103,893,839	104,012,549	X		
5.	RP5-947P14	106,370,147	106,477,206	X		
6.	RP4-669H10	106,562,075	106,683,105	X		
7.	RP5-916A15	106,759,849	106,884,451		X	
8.	RP11-294K17	107,172,477	107,359,937			X
9.	RP11-356N1	108,329,970	108,513,006			X
10.	RP11-297O4	109,590,896	109,760,136			X
11.	RP11-389O22	113,371,752	113,560,563			X
12.	RP4-787H6	116,580,484	116,707,534			X

**BACs and Fosmids from Chromosome 11** (coordinates from NCBI Build 36/hg18 assembly)

No.	BAC ID	Start (bp)	End (bp)	Telomeric	Spanning	Centromeric
1.	RP11-992G23	45,743,060	45,953,917	X		
2.	RP11-110P15	45,790,852	45,990,620	X		
3.	RP11-618K13	45,804,176	45,975,298	X		
4.	G248P88405C7 (fosmid)	45,972,774	46,012,368	X		
5.	RP11-177H1	45,928,255	46,103,134		X	
6.	G248P88722D5 (fosmid)	45,995,837	46,032,977		X	
7.	RP11-702F3	46,021,151	46,213,090			X
8.	RP11-425L10	46,359,782	46,429,924			X

**Table S3. Primers Used for Mapping and Cloning Breakpoints**

<b>1. DGAP012 with t(11;19)(p11.2;p13.2)</b>		
<b>Name</b>	<b>Forward Primer (5'→3')</b>	<b>Reverse Primer (5'→3')</b>
D012-A	CCAATGAGCTGGTGATCCTT	GTTTCCTTCCTTAGACGTCG
D012-B	CAAAGCTATCTCCACCACAG	CTCTCCGTTTGGATTGTCTG
D012-C	CCTGGAAACTCAGAAGAAAGC	TGGTATATAGGTCGTCCTCC
D012-D	CTTGCATACTGTTCTCTAAGG	TCAAGTTCCTTTGTAGCCTTG
D012-E	GCTGTGCATGCAATGTGTGG	GTCAAGTTTGGTGAGCCTGC
D012-F	GAAAACGCACACTCAGCATC	CGTCACCAATGTGAAAGTGATC
D012-G	GAGAACCTGGCAGACAGAG	TTGGAAAAAGGAGTTCAGCC
Adaptor 1	5'-CTAATACGACTCACTATAGGGCTCGAGCGGCCCGCCCGGGCAGGT-3'	
Adaptor 2	5'-ACCTGCCCGG-NH <sub>2</sub> -3'	
Suppression boost PCR	GACACCTGCATGGGTCAAAAC (chr19 forward 1)	CCTAATACGACTCACTATAGG (AP1-a)
Suppression nested PCR	CTCCGTTGAAGGTAGAAATGC (chr19 forward 2)	CTATAGGGCTCGAGCGGC (AP2-a)
der(19) junction fragment	Boost PCR	
	GGATGGTCAATGACAAGCTC (chr11 forward 1)	CTGAGAGAAGAGGAGGAAGA (chr19 reverse 1)
	Nested PCR	
	AGACGTGATCACTTTTCTCCAT (chr11 forward 2)	GGAGGAAGAGAGGAGAAGAT (chr19 reverse2)
der(11) junction fragment	Boost PCR	
	GACACCTGCATGGGTCAAAAC (chr19 forward 1)	CTCTCCGTTTGGATTGTCTG (chr11 reverse 1)
	Nested PCR	
	CTCCGTTGAAGGTAGAAATGC (chr19 forward 2)	CGTGTGTCAGTCTCTTCTGT (chr11 reverse 2)
<b>2. MCN1762 with t(1;11)(p21.1;p11.2)</b>		
der(1) junction fragment	CTTCTCCCATGAACCA-TCTGAAAGTACGCCAG (chr11 forward)	CAGCTGTAGTCACTTA-TAGCTATAACGTCCCT (chr1 reverse)
der(11) junction fragment	CAGTGGCAGGATGACGTTGAGACATAGAG (chr1 forward)	GTTTAAAATGGCCAAA-TACCTGTCCAGCAATC (chr11 reverse)

**Table S4. Primers Used for Mutation Screening of *PHF21A* (NM\_001101802.1)**

<b>Location</b>	<b>Forward Primer (5'→3')</b>	<b>Reverse Primer (5'→3')</b>
exon 3	TCTTGCCCTGCGTTAGAGAATCAG	GATGGACTTGTTTACTAATCTTAGC C
exon 4	GGTCCAAACCTGATTAGCCATAG	CTCAAACCAATTAATATGACACC
exon 5	CATAGTCTTCCTGCTACTATTCTG	GTTCTCTTAGAAGCCTCGAGCAGAC
exon 6	CATATTTCTGTGCAGCTGATAGAGC AG	GACAGGACCCAGTTCTATATGAC
exon 7	CAGCATATTGTTTAGTGGGATTTAG	CTATACAAATGCGTATCCTAACC
exon 8 of NM_016621 *	GTCTGATCCAAAGAATAGTGATGC	GCTACTTAGAAAGAGCCCATGAGG
exon 9	TGGAAGATTTGCTTTTATAGGAC	CGGGTCGTCGGTATACGGAGTTGC
exon 10	CTATTAACCTTCTGATTGCTTCATCC	CATCATTTCAGAGAGCAGGTGCC
exon 11	CCTACTAGGAACTGGCAATTGGC	CATCATTTTCAGGAACAAAGCTG
exon 12	CCTGATTCTGATTTCTTATGG	CAGAGGGGAGGCACTGAACAGCTC
exon 13	GTCTTTAGCTTGTGCCAGTATTAG	GGAGAAGAAACACACACACAAAGC
exon 14	CCATGCTAGCCAATGAGGAACAC	GGTGCAAGTCGATAAGGAGACAAC
exon 14 of NM_016621	GCTTCTCTTCCTTCCTTTGG	GGAGGACGACCTATATACCAAGAG
exon 15	GGAACGAGGCAGCTGATTGTCC	GGAATTCCTCCTGATGGCCGTGTC
exon 16	GCTAGCTGTTACATAGAGGATGTG	CGTTTGTTTCCTCCAATAGTTAAGG GC
exon 17	GCAGGCATTCTGAGCACAGAGGAC	GTATATTGGAAAGGCCTAGGTCTC
exon 18	GAATCCCCGACGGTCCGGACAATC	GAATTCTGCACTTTCAGAAATCCG G

\* exon 8 of NM\_016621.3 has additional nucleotides CAG at the very beginning of its counterpart exon 8 of NM\_001101802.1

**Table S5. Primers Used for the *SCN3A* Promoter Target**

<b>Location</b>	<b>Forward Primer (5'→3')</b>	<b>Reverse Primer (5'→3')</b>
#1	CAGCCATCTCTGTTCTGCTG	GTGGCTCCAGCTTACCTTTG
#2	CCATATTTCCAGCCATGTG	GGAACGGGATCATACTCTGC
#3	CTCTGTACAGGGAGGAAAG	AGACTAGAGCAGGCCACAAG
#4	TCCGCTTCCTGTTCTGAGAT	TGACTACCCATGCTTCATGC
#5	GCCTGTCTGGAAATGGTGAT	GCATGGCCTCTCTAAAGCTG
<b>Control oligos designed in an intergenic region between <i>ACTG1</i> and <i>FSCN2</i></b>		
<b>Name</b>	<b>Forward primer (5'→3')</b>	<b>Reverse primer (5'→3')</b>
Control oligos	CAATTGACAGGCAATGATGG	ACAAGAGAGGCCTTGGGAAT

**Table S6. Primers Used for Mouse *In Situ* Hybridization Analysis**

<b>Probe #</b>	<b>Location in NM_138755.2</b>	<b>Forward Primer (5'→3')</b>	<b>Reverse Primer (5'→3')</b>
#1	2059 to 2538	TGTAACCAGGGGAAGAGACT A	AAGTTTAGGAACAGGAAAGCC C
#2	307-667	AGCCCACTACCTCATAAGCTG	GTAGTTTGTTCAGACTGTGA C
#3	1026-1322	CATCCACCCTGTCCGTGTTG	TGGTCATGTGTTACCAACCC

**Table S7. Primers Used for Zebrafish *phf21a* Constructs**

<b>Name</b>	<b>Forward Primer (5'→3')</b>	<b>Reverse Primer (5'→3')</b>
<b>RT-PCR</b>	CCATCGATCCTTATTCCAGCAGCAG GGAGT	CTATTTGGGGTTGTCCACGCT GCT
<b><i>phf21a-RFP</i> fusion reporter</b>	CCATCGATCCTTATTCCAGCAGCAG GGAGT	CCATCGATGAGGAGAGGATG AAGGTGGTGT