Accepted Manuscript

Targeted next generation sequencing approach identifies nineteen new candidate genes in normosmic hypogonadotropic hypogonadism and Kallmann Syndrome

Samuel D. Quaynor, Maggie E. Bosley, Christina G. Duckworth, Kelsey R. Porter, Soo-Hyun Kim, Hyung-Goo Kim, Lynn P. Chorich, Megan E. Sullivan, Jeong-Hyeon Choi, Richard S. Cameron, Lawrence C. Layman

PII: S0303-7207(16)30297-0

DOI: 10.1016/j.mce.2016.08.007

Reference: MCE 9601

To appear in: *Molecular and Cellular Endocrinology*

Received Date: 12 April 2016

Revised Date: 3 August 2016

Accepted Date: 3 August 2016

Please cite this article as: Quaynor, S.D., Bosley, M.E., Duckworth, C.G., Porter, K.R., Kim, S.-H., Kim, H.-G., Chorich, L.P., Sullivan, M.E., Choi, J.-H., Cameron, R.S., Layman, L.C., Targeted next generation sequencing approach identifies nineteen new candidate genes in normosmic hypogonadotropic hypogonadism and Kallmann Syndrome, *Molecular and Cellular Endocrinology* (2016), doi: 10.1016/j.mce.2016.08.007.

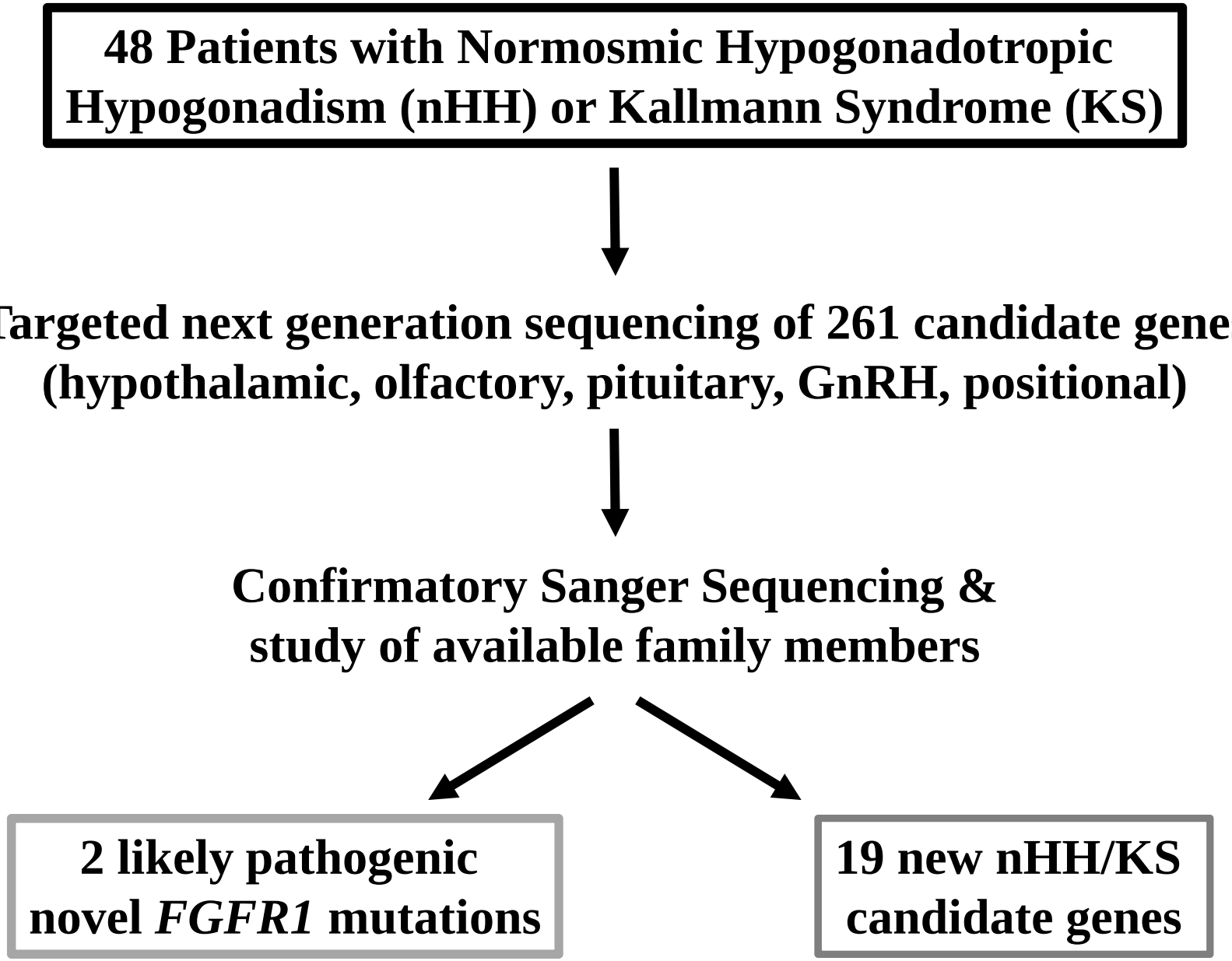
This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Georgia Online Research Archive

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

Targeted Next Generation Sequencing Approach Identifies Nineteen New Candidate Genes in Normosmic Hypogonadotropic Hypogonadism and Kallmann Syndrome

Samuel D. Quaynor, M.D./Ph.D.[1,2]

Maggie E. Bosley, B.S.[1*]

Christina G. Duckworth, B.S.[1*]

Kelsey R. Porter, .M.S.[1]

Soo-Hyun Kim, Ph.D. [3]

Hyung-Goo Kim, Ph.D.[1,4]

Lynn P. Chorich, M.S.[1,4]

Megan E. Sullivan, B.S. [1,4]

Jeong-Hyeon Choi, Ph.D.[5]

Richard S. Cameron, Ph.D.[4,5]

Lawrence C. Layman, M.D. [1,4,6]

[1] *Section of Reproductive Endocrinology, Infertility, & Genetics, Department of Obstetrics & Gynecology, Medical College of Georgia, Augusta University, Augusta, GA*

[2]*University of Chicago Department of Neurology, Chicago, IL*

[3]*Molecular Cell Sciences Research Centre, St. George's Medical School, University of London, London, UK*

[4] *Department of Neuroscience and Regenerative Medicine; Medical College of Georgia, Augusta University, Augusta, GA*

[5]*Department of Medicine, Medical College of Georgia, Medical College of Georgia, Augusta University, Augusta, GA*

[6]*Department of Physiology, Medical College of Georgia, Augusta University, Augusta, GA*

ACCEPTED MANUSCRIPT

*Samuel Quaynor and Maggie Bosley contributed equally

Corresponding author: Lawrence C. Layman, MD.

Robert B. Greenblatt, M.D. Distinguished Chair in Endocrinology

Section of Reproductive Endocrinology, Infertility, & Genetics

Department of Obstetrics & Gynecology

Department of Neuroscience & Regenerative Medicine

Department of Physiology

Medical College of Georgia at Augusta University

1120 15th Street, Augusta, GA 30912

lalayman@augusta.edu

Phone (706)721-7591; Fax (706)721-8685

Grants: Funding by NIH grant HD033004 & Medical College of Georgia Bridge Foundation (L.C.L.) This work was presented in part at the Society for Reproductive Investigation in Florence, Italy, March 2014.

DISCLOSURE STATEMENT: The authors have nothing to disclose. We would like to acknowledge students who contributed to this project-- Luke V. Lee, Irene Falk, ME Jett, Julianne Jett, and Alli Falkenstrom.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Abstract

The genetic basis is unknown for ~60% of normosmic hypogonadotropic hypogonadism (nHH)/Kallmann syndrome (KS). DNAs from (17 male and 31 female) nHH/KS patients were analyzed by targeted next generation sequencing (NGS) of 261 genes involved in hypothalamic, pituitary, and/or olfactory pathways, or suggested by chromosome rearrangements. Selected variants were subjected to Sanger DNA sequencing, the gold standard. The frequency of Sanger-confirmed variants was determined using the ExAC database. Variants were classified as likely pathogenic (frameshift, nonsense, and splice site) or predicted pathogenic (nonsynonymous missense). Two novel *FGFR1* mutations were identified, as were 19 new candidate genes including: *AMN1, CCKBR, CRY1, CXCR4, FGF13, GAP43, GLI3, JAG1, NOS1, MASTL, NOTCH1, NRP2, PALM2, PDE3A, PLEKHA5, RD3, TMTC1,* and *TRAPPC9,* and *TSPAN11.* Digenic and trigenic variants were found in 8/48 (16.7%) and 1/48 (2.1%) patients, respectively. NGS with confirmation by Sanger sequencing resulted in the identification of new causative *FGFR1* gene mutations and suggested 19 new candidate genes in nHH/KS.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Key Words

Next generation DNA sequencing

Kallmann syndrome

Hypogonadotropic hypogonadism

Delayed puberty

GnRH deficiency

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

1. Introduction

The development of reproductive function in humans is regulated by complex signaling interactions of the hypothalamic-pituitary-gonadal (HPG) axis. Disruption of any component of this system may result in delayed puberty and infertility (Layman, 2013). Central nervous system defects may impair GnRH action and function resulting in GnRH deficiency, also known as hypogonadotropic hypogonadism, which may be manifested clinically by low sex steroids and low or inappropriately normal gonadotropins. To date, two principle conditions constitute deficiency in GnRH signaling, namely normosmic hypogonadotropic hypogonadism (nHH) and Kallmann Syndrome (KS). Patients with KS present with hypogonadotropic hypogonadism and a lack of smell due to the impaired migration of GnRH and olfactory neurons. Apart from pubertal and reproductive disturbances, other associated anomalies such as renal agenesis, midfacial defects, neurologic defects, and cardiac anomalies have been reported (Layman, 2013,Bhagavath et al., 2006).

Even with improved knowledge of GnRH development and function, only 40% of all nHH/KS cases can be explained by reported mutations in more than 30 genes. Some mutations occur in ligand/receptors such as *KAL1/FGF8/FGFR1/HS6ST1, LEP/LEPR, GNRH1/GNRHR, PROK2/PROKR2, KISS1/KISS1R,* and *TAC3/TACR3*.(Layman, 2013) Other identified genes include *NR0B1*, *CHD7*, *NELF*, *WDR11*, *SEMA3A*, *SOX10, FGF17, IL17RD, DUSP6, SPRY4, FLRT3, FEZF1, STUB1, HESX1, PCSK1, RNF216,* and *OTUD4* (Layman, 2013,Hanchate et al., 2012,Tornberg et al., 2011,Margolin et al., 2013). These causative genes were identified through linkage analysis, candidate gene approaches, and positional cloning (Layman, 2013,Kim et al., 2010,Kim et al., 2008). The inheritance pattern of nHH/KS includes X-linked recessive, autosomal dominant, autosomal recessive, sporadic, and at least several percent appear to be digenic/oligogenic (Quaynor et al., 2011,Sykiotis et al., 2010). Molecular studies performed by both *in vitro* and *in vivo* analysis of the genes indicate that either the development/migration or signaling of GnRH is altered or impaired (Layman et al., 1998,Quaynor et al., 2013).

ACCEPTED MANUSCRIPT

The majority of molecular causes of nHH/KS are yet to be characterized. This gap in our understanding may be narrowed with advancements in massively parallel deep resequencing methods, otherwise known as next generation sequencing (NGS), which includes targeted NGS, whole exome sequencing, and whole genome sequencing. More expedient, less expensive, and theoretically more accurate results should be able to be obtained using NGS (Metzker, 2010). The goal of the present study was to identify new nHH/KS candidate genes using targeted NGS of potentially relevant known genes that are involved in GnRH and olfactory neuron development, migration and signaling.

2. Materials and Methods

2.1 Genes Selected for NGS

The list of 261 genes was gathered from the literature to identify important genes involved in hypothalamic or pituitary development, GnRH or olfactory neuron specification, migration, or regulation, nHH/KS known pathway genes, or genes located near derivative chromosome breakpoints in nHH/KS patients with chromosome translocations (genes shown in Table 1; references for gene selection shown in Supplemental Table 1). All known nHH/KS genes at the time of the study initiation in 2011 were also included.(Layman, 2013) For most genes, 95-100% of the desired sequence was covered—protein coding exons and splice junctions, but it was not 100% of all candidate genes. For 7 genes, *FAM60A, FGF20, FGF7, GSTM1, HS6ST1, TUBB2A,* and *TUBB2B,* the coverage ranged from 61-89.2% (Supplemental Table 2).

2.2 Patients

Hypogonadotropic hypogonadism was defined as the absence of puberty (age 17 in females; 18 in males) with low/normal FSH and LH levels (Bhagavath et al., 2006). Affected males had serum testosterone levels less than 100ng/dL (normal: 300-1100ng/dL), whereas females had absent breast development and

ACCEPTED MANUSCRIPT

hypoestrogenic amenorrhea (Bhagavath et al., 2006). MRI of the brain and pituitary was normal. Serum prolactin, thyroid stimulating hormone (TSH), thyroxine (T4), and cortisol were normal (Bhagavath et al., 2006). Patients characterized as nHH exhibited a normal sense of smell by the Smell Identification Test or by history, while those with KS had absent or decreased sense of smell (Quaynor et al., 2011).

Seventeen males (10 with KS, 5 with nHH, and 2 with unknown sense of smell) and 31 females (12 with KS, 17 with nHH, and 2 with unknown sense of smell) were studied by NGS. No patients had mutations in any known nHH/KS genes; and none had undergone NGS previously. This study was approved by the Augusta University Human Assurance Committee, and all patients consented to participate in the study.

2.3 Targeted Next Generation Sequencing

Genomic DNAs from 48 patients were extracted from blood using phenol-chloroform and ethanol precipitation (Xu et al., 2011) and treated with RNAse to remove residual RNA. DNA (2-3ug) was submitted to Otogenetics Corporation (Norcross, GA USA) for selected target enrichment and sequencing. Control DNA was provided by Otogenetics Corporation. Genomic DNA was subjected to agarose gel and OD ratio tests to confirm the purity and concentration prior to quantification by the Qubit assay (Life Technologies, Grand Island, NY USA). DNA fragmentation using restriction enzyme digestion and target enrichment and Illumina ready libraries were made using HaloPlex target enrichment kits (Agilent, Santa Clara, CA USA, Catalog# G9901B) following the manufacturer's instructions. Enriched libraries were tested for enrichment, size distribution, and concentration by an Agilent 2200 TapeStation. For the selected genes, 261 target IDs were resolved to 285 targets comprising 3099 regions of the genome.

DNA sequencing was performed on an Illumina HiSeq2000 to generate paired-end reads of 100 nucleotides and $\geq$30X coverage. Burrows-Wheeler Aligner (BWA) was used to map sequence reads to the genome; Genome Analysis Toolkit (GATK) was utilized to call exome-wide variants; and SnpEff was used to annotate variants (missense, nonsense, etc.). Sequence was filtered to exclude synonymous changes and variants in SNP

ACCEPTED MANUSCRIPT

database v132. Selected variants confirmed by Sanger sequencing were also assessed for frequency in the ExAC database (http://exac.broadinstitute.org/).

Binary alignment map (BAM) files of selected variants were manually evaluated to determine if the variant was identified close to the ends of the PCR fragments, which could be more problematic to map since A's are placed at the ends of amplicons during PCR. These were classified in 10 percent increments (0-10%, 11-20%, 21-30%, etc.) based upon the percentage of reads within 10 bp from the end of fragments out of total reads (referred to as end reads) in BAM files. Variants were classified according to guidelines from the American College of Medical Genetics and Genomics and Association for Molecular Pathology as class 4—likely pathogenic (frameshift, nonsense, and canonical splice site) or class 3 (nonsynonymous missense), which were predicted deleterious *in silico*).(Richards et al., 2015)

The 319 sequence variants identified by GATK were filtered out with the following criteria: 1) depth < 20, 2) the number of patients with alternative allele > 5, 3) known variants, and 4) non-coding. The minimum, average, and maximum depths (DP) of the final variants are 95, 8425, and 12000, respectively. The minimum, average, and maximum quality scores (QUAL) of the final variants are 21, 1970, and 15995, respectively.

2.4 Sanger Sequencing of Variants

Sanger sequencing is considered the gold standard to confirm or refute nucleotide changes found by NGS. Any variant found by NGS is only preliminary—it must be confirmed by Sanger DNA sequencing to consider it as a bonafide variant. Only genes in which variants found by NGS and subsequently verified by Sanger sequencing were considered as being new candidate genes. Those variants identified by NGS, but not confirmed by Sanger sequencing, were considered to be false positives.

Selected class 4 variants were subjected to Sanger sequencing and available family members were studied for the presence or absence of the variant (Quaynor et al., 2011). To prioritize missense variants, three different bioinformatic programs were used: SIFT (Sorting Intolerant from Tolerant) (Ng et al., 2003), Mutation Taster

(Schwarz et al., 2010), and PolyPhen2 (Polymorphism Phenotyping) (Adzhubei et al., 2010). Missense variants were selected for PCR-based Sanger sequencing if ≥2 programs predicted that the variant would be deleterious. The prevalence of the variant was ascertained using the ExAC Database (http://exac.broadinstitute.org/), which consists of sequence data from ~61,000 individuals. Sanger sequencing was performed on these variants to determine if targeted capture sequencing was correct (Layman et al., 1997). Available family members were then studied to assess segregation of the variant with the phenotype.

2.5 The Reproducibility of Targeted NGS

There was concern regarding the high prevalence of false positive results by targeted NGS. As part of this project, we wanted to determine if the location of the variant near the end of the PCR product, as viewed in BAM files, affected the reliability of the targeted NGS. Therefore, 90 of our 107 variants (combination of all variants) were categorized based upon the percentage of reads within 10bp of the ends of fragments (end reads). As can be seen in Supplemental Table 3, when ≤ 70% of the reads were within 10bp of the ends of fragments, correct variant calls by Sanger sequencing occurred in 21/63 (33.3%), range of 0-42.9%. However, for 71-100% end read calls, only 1/27 (3.7%), range of 0-16.7%, were confirmed by Sanger sequencing.

Of the class 4 variants called by NGS that were subjected to Sanger sequencing, only 10/17 (58.8%) were confirmed. Of the class 3 variants, 28/90 (31.1%) were confirmed by Sanger sequencing. Therefore, only 38 of 107 (35.5%) variants were confirmed when both class 3 and 4 variants were grouped together. It was only these Sanger-confirmed variants that were considered to be true variants that could have a potential role in nHH/KS.

2.6 Colony PCR/Sequencing and RT-PCR

For frameshift variants, PCR was performed, confirmed by agarose gel electrophoresis, and then cloned using the TA Cloning Kit according to manufacturers' instructions. At least 10 colonies were selected for Sanger sequencing as described previously (Layman et al., 1997). For variants predicted to affect splicing,

reverse transcription PCR (RT-PCR) was performed as described previously (Kim et al., 2008). RNA was extracted from lymphoblasts and subjected to RT-PCR. PCR products were then sequenced to determine if exon skipping or insertion was present.

3. Results

319 variants from 261 genes were identified by targeted NGS in 48 nHH/KS patients. Of the 319 variants identified, 56 were class 4 (frameshift, nonsense, and splice site) and 153 were class 3 (missense) variants. The remaining 110 variants were located in the 5'UTR, 3'UTR or >10 bp from splice sites in introns, and were not studied further. It should be noted that only true variants in genes identified by NGS that were confirmed by Sanger sequencing were considered important for nHH/KS (Table 2).

3.1 Class 4 Variants

Thirty-four class 4 variants found by targeted NGS were excluded from consideration for Sanger sequencing confirmation because the reads were seen almost exclusively within 10 bp of the ends of the PCR fragment on BAM files and therefore, not likely to become confirmed by Sanger sequencing. Of the 22 variants remaining, they were identified in the following genes: *AXL, CCKBR, CHAT, FGF3, FGFR1, GAS6, GPRIN2, IGF1, MASTL, NRP2, PTCH2, SDR16C5, SOS1, SP8, TRAPPC9* and *CHGA*. *CHGA* was not studied further as it was mistakenly included in our initial 261 genes selected, with the intention of studying CGA. These variants in candidate genes were subjected to Sanger sequencing, which was necessary to confirm the variant observed by NGS (Supplemental Table 4). The variants included start gain, which produce new ATGs or CTGs thereby initiating translation (which was unexpectedly called by SnpEff), frameshift, splice site, and nonsense. Three other variants in three genes (*GAS6, IGF1* and *SDR16C5*) were not subjected to Sanger sequencing because they were classified as start gain, but did not create a new ATG or CTG (start gain variants produce new ATGs

or CTGs which could initiate translation) (Dever, 2012). In total, 10/17 (59%) class 4 variants were positively confirmed by Sanger sequencing.

3.2 *FGFR1* Mutations

FGFR1 has been shown to be involved in ~10% of IHH/KS patients (Layman, 2013). Three previously unreported *FGFR1* variants, all of which were heterozygous and absent in the ExAC database, were identified and confirmed by Sanger sequencing. These variants included one intronic insertion 3bp from the exon 16 splice donor site region, one nonsense, and one frameshift variant (Table 2).

In the first patient, an insertion near a splice donor site (c.2180+3insT) of *FGFR1* yielded a single wild type band from RT-PCR from lymphoblast RNA prior to sequencing, arguing against alternative splicing. Sanger sequencing revealed that the unaffected father and unaffected sister were also heterozygotes, suggesting that it is a polymorphism (Table 2; pedigree not shown). In the second patient, a c.313C>T (p.Q72X) variant created a premature stop codon within the linker region of FGFR1, which was predicted to result in protein truncation. By Sanger sequencing, the mother was unaffected, but the father was not available for study (Table 2; Figure 1A). This patient also had a heterozygous missense variant in a second gene—*GLI3* (see below), which was not observed in her mother (Table 3; Figure 1A). In the third patient, a frameshift mutation in *FGFR1* (c.2248delG, p.E750Kfs*9K) resulted in 8 altered amino acids following the deletion before a premature stop codon. DNA from the father was not available, but her mother did not have the variant. This patient also had a missense variant in *AXL* that was absent in her mother (Tables 2 & 3; Figure 1B). Sanger sequencing chromatograms are only shown for both *FGFR1* class 4 mutations in Supplemental Figure 3.

3.3 Other Class 4 Variants

Complex variants (heterozygous 9bp insertion and missense variants) were observed in *GPRIN2* and a start gain variant was seen in *MASTL* and in *AXL*. However, the *GPRIN2* variants were seen in unaffected family

members. The *MASTL* and *AXL* variants did produce a new ATG and CTG, respectively, which are both potential new protein translation start sites (not shown). For both *MASTL* and *AXL* variants, available unaffected family members did not possess the variant (Table 2). In total, only two class 4 variants, both in *FGFR1*, are more definitively contributive to the disease phenotype (Table 2). However, it is possible that the start gains in *MASTL* and/or *AXL* could be pathogenic, but these will require further *in vitro* translation studies (suggesting they are better characterized as class 3 variants).

3.4 Class 3 Variants

One hundred fifty three missense variants were identified by NGS in our cohort of 48 nHH/KS patients, and 90 variants were predicted to be deleterious by ≥2 protein prediction programs. These 90 variants were subjected to Sanger DNA sequencing, but only 30 (33.3%) of the variants were confirmed (Tables 2& 3), all of which were heterozygous. The frequency of each confirmed variant in the ExAC database is shown in Table 2 (only confirmed variants by Sanger sequencing are shown). Heterozygous missense variants were identified in five known nHH/KS genes—*FGFR1*, *GNRH1*, *GNRHR*, *IL17RD*, and *AXL*, all of which were predicted to be deleterious by ≥2 of 3 programs. Additional missense variants were found in the following genes—*AMN1*, *CCKBR*, *CRY1*, *CXCR4*, *FGF13*, *FGFR3*, *GADL1*, *GAP43*, *GLI3*, *GPRIN1*, *JAG1*, *NOS1*, *NOTCH1*, *NRP2*, *PALM2*, *PDE3A*, *POMC*, *RD3*, *SEMA4D*, *SMO*, *TMTC1*, *TRAPPC9*, and *TSPAN11* (Table 2A & B). When available family members were studied, some of these variants were seen in unaffected family members, consistent with probable polymorphisms (Table 2B;). Therefore, 19 novel candidate genes were identified and include: *AMN1*, *CCKBR*, *CRY1*, *CXCR4*, *FGF13*, *GAP43*, *GLI3*, *JAG1*, *NOS1*, *MASTL*, *NOTCH1*, *NRP2*, *PALM2*, *PDE3A*, *PLEKHA5*, *RD3*, *TMTC1*, and *TRAPPC9*, and *TSPAN11*. (Table 2A). Thirteen genes were physiologic candidates, while six genes (*AMN1*, *MASTL*, *PDE3A*, *PLEKHA5*, *TMTC1*, and *TSPAN11*) were positional candidates from reported translocations.(Bhagavath et al., 2006,Kim et al., 2008).

ACCEPTED MANUSCRIPT

Initially, nine patients had digenic and two patients had trigenic variants (Table 4). Segregation patterns are shown in Figure 1. By chance, half of the pedigrees would be expected not to segregate with the variant if they were polymorphisms. Although not all parents were available, nine of eleven pedigrees were consistent with the possibility of digenic or trigenic segregation (Figure 1). In most cases, the proband was heterozygous for both variants and they were absent in the one parent studied. Alternatively, the proband had two variants, but no one unaffected had both (although they could be heterozygous for one variant). Two pedigrees suggested a polymorphism because the proband and an unaffected relative were heterozygous for trigenic or digenic variants (Figure 2).

Although not digenic, the *CCKBR* gene missense variant c.500C>T (p.T167M) was observed in a large family, and segregated with both nHH and KS suggesting that it could potentially be detrimental (Table 2; Figure 1J). Three family members with KS and one with nHH were heterozygous for this *CCKBR* missense variant. Interestingly, one relative heterozygous for the variant had a brain glioma, and two unaffected individuals were wild type.

4. Discussion

The molecular basis of nHH/KS has been elucidated for 30-40% of all patients. When there is a family history of nHH/KS, particularly if there are associated anomalies, targeted sequencing of known genes may be warranted.(Layman, 2013). However, should results yield no positive findings, further analysis by targeted NGS could be considered. Despite false positive variants, the Sanger sequencing confirmed variants found by targeted NGS revealed meaningful results in both known and previously unsuspected nHH/Ks genes. The prevalence of mutations in known genes in our cohort of 48 nHH/KS is not inconsistent with that in the literature, as *FGFR1* mutations have been reported in 6-10% of patients (Pitteloud et al., 2006,Xu et al., 2007). In the current study, two of 48 (~4%) had deleterious (class 4) mutations, both of which were in *FGFR1*.

Neither mutation was found in the ExAC database or reported in the literature. In addition, there was one other *FGFR1* variant, a missense predicted deleterious, totaling 3/48 (6.3%). Other heterozygous variants in known genes *GNRH1*, *GNRHR*, *IL17RD*, and *AXL* were also found in 7/48 (14.5%) of patients. The prevalence of mutations in most genes involved in nHH/KS is only 1-2%, so it is not surprising that variants in other known nHH/KS genes were not identified (Layman, 2013).

In the current study, targeted NGS with Sanger confirmation also resulted in the identification of variants in 19 new nHH/KS candidate genes. These variants are predicted to be deleterious; and their available pedigree analyses do not indicate polymorphisms. The new candidate genes include: *AMN1*, *CCKBR, CRY1, CXCR4, FGF13, GAP43*, *GLI3, JAG1*, *NOS1*, *MASTL, NOTCH1*, *NRP2, PALM2*, *PDE3A, PLEKHA5, RD3*, *TMTC1*, and *TRAPPC9*, and *TSPAN11*. Because targeted sequencing of known genes was performed, mutations in unknown genes (i.e., genes not on the panel) would not be identified using our targeted approach as it could be with whole exome sequencing (Yang et al., 2013). All 261 genes were selected because of possible involvement in GnRH or olfactory nerve specification, migration, and/or function, hypothalamic or pituitary development, so they all have potential relevance to nHH/KS. Variants located close to translocation breakpoints may or may not have direct relevance by proposed physiology, but could be involved by a position effect upon other genes (Kleinjan et al., 2005).

Class 4 variants, which consist of nonsense, frameshift, and canonical splice site variants, are generally considered as likely pathogenic. Although class 3 (nonsynonymous missense) variants may be predicted by protein programs to impair function, they will require future study to determine if they impair protein function *in vitro*. This is similarly true for the start-gain variants. Nevertheless, our identification and confirmation of variants in 19 new candidate genes suggests that this small subset of genes could be important in nHH/KS. Nearly all variants in these new candidate genes are monoallelic, so it is not known if further study of additional patients would yield any biallelic variants. As reported by Sarfati et al,(Sarfati et al., 2010), the phenotype could

be more severe as with biallelic variants in *PROK2* or *PROKR2*. However, mutations in more than one could also contribute to the phenotype.

We identified nine possible digenic and trigenic pedigrees (~19%), but this will require further characterization of a large number of cohorts, animal studies, and cell-based assays (Schaffer, 2013). Digenic or trigenic variants have been identified in several percent of nHH/KS patients, but in most cases it is extremely difficult to discern whether both variants are sufficient and/or contributing to the phenotype (Quaynor et al., 2011,Sykiotis et al., 2010). Variants involved in digenic disease could affect the phenotype by synergistic heterozygosity of genes in the same pathway, whereas heterozygosity for either variant alone may have no or minimal clinical impact upon phenotype (Quaynor et al., 2011).

Confirmed missense variants, such as those found in *CXCR4,* as well as the other candidates, will be studied in our large cohort of nHH/KS patients by sequencing the entire protein coding exons and splice site of each candidate gene for mutations (Table 5). CXCR4 protein affects sensory axon extension and is known to be important in GnRH neuron migration.(Schwarting et al., 2006) Additionally, it may play a role in directional migration, proliferation and survival of primordial germ cells (Gu et al., 2009). *CCKBR* is a particularly promising new candidate gene because of its segregation pattern in our large family with nHH/KS. *CCKBR,* encoding the cholecystokinin B receptor, is a G-protein coupled receptor neuropeptide expressed in the paraventricular nucleus of the hypothalamus, which is involved in binding cholecystokinin and gastrin (Mohammad et al., 2012).Whether it is associated with glioma formation requires further study. Relevance to nHH/KS is also available for the other candidate genes— *AMN1* (Bhagavath et al., 2006,Wang et al., 2003), *CRY1* (Wunderer et al., 2013), *FGF13* (Itoh et al., 2004), *GAP43* (Strittmatter et al., 1995), *GLI3* (Balmer et al., 2004), *JAG1* (Ratie et al., 2013), *MASTL* (Bhagavath et al., 2006,Johnson et al., 2009), *NOS1* (Leshan et al., 2012,Garrel et al., 2010), *NOTCH1* (Ratie et al., 2013), *NRP2* (Cariboni et al., 2011), *PALM2* (Panza et al., 2007), *PDE3A (Masciarelli et al., 2004), PLEKHA5 (Dowler et al., 2000), RD3* (Friedman et al., 2006),

TMTC1(Sunryd et al., 2014), *TRAPPC9* (Lorenc et al., 2014), and *TSPAN11*(Bassani et al., 2012). All of these genes represent reasonable physiologic or positional candidates (Table 5).

Six of the candidate genes were selected for the sequencing panel because they were near translocation breakpoints: *AMN1, MASTL, PDE3A, PLEKHA5, TMTCI,* and *TSPAN11* (Table 5).(Bhagavath et al., 2006,Kim et al., 2008) However, these positional candidates may also have some plausible physiologic roles in nHH/KS. *TSPAN11* regulates cell adhesion, motility, and synapse formation;(Bassani et al., 2012) *PLEKHA5* is expressed in the brain and interacts with phosphatidylinositol trisphosphate, an important second messenger in G-protein coupled receptors;(Dowler et al., 2000) and *TMTC1* is involved in calcium homeostasis.(Sunryd et al., 2014) The other positional candidates are involved in cell cycle exiting (*AMN1*),(Wang et al., 2003) cell division (*MASTL*),(Johnson et al., 2009) and meiosis resumption (*PDE3A*).(Masciarelli et al., 2004)

Of the remaining 13 new candidate genes, two would be predicted to be involved in nHH/KS—*FGF13* is related to other FGF/FGFR genes *FGFR1* and *FGF8,* which are genes known to have mutations in nHH/KS patients.(Itoh et al., 2004) Similarly, known KS gene *SEMA3A* encoding semaphorin 3A, binds to one of its receptors—neuropilin 2, encoded by *NRP2* (Table 5).(Cariboni et al., 2011) *PALM2* is expressed in the olfactory bulb and was proposed to be a KS gene as one patient had a translocation disrupting this gene.(Panza et al., 2007). Two candidates are ligand and receptor--*JAG1* encoding Jagged 1, which binds to its receptor, NOTCH, encoded by *NOTCH1,* which are involved in cell differentiation and morphogenesis.(Ratie et al., 2013). *GLI3* is a very viable candidate gene because it is involved in the sonic hedgehog pathway, which regulates multiple body processes including induction, patterning, and differentiation. It affects olfactory exons and their directed outgrowth to the telencephalon.(Balmer et al., 2004) *NOS1* has well known vascular functions, but it is also known that GnRH increases *NOS1* expression in pituitary gonadotropes. *CRY1* is pituitary clock gene thought to regulate the feedback loop of CLOCK-BMAL;(Wunderer et al., 2013); *GAP43* regulates growth of axons and moderates the formation of new neuronal connections(Strittmatter et al., 1995); *RD3* is highly expressed in the retina and is involved in transcription and splicing;(Friedman et al., 2006); and

TRAPPC9, within the Prader-Willi region, is highly expressed in the brain and appears to regulate the NF-kappaB signaling pathway.(Lorenc et al., 2014) These candidates are summarized in Table 5.

The goal of targeted capture sequencing is to rapidly and accurately identify disease-causing variants. In this study, many variants were identified requiring confirmation by Sanger sequencing; however, a majority were false positives. Possible explanations for this high rate of false positives could have to do with the procedure in which the DNA was sheared by restriction enzyme digestion rather than a random method (such as sonication), depth of reads, quality assessment, read alignment, variant identification, and/or annotation (Pabinger et al., 2014). The degree to which variants were missed—the false negatives—is not able to be addressed in this study. However, the confirmed variants do provide exciting new possibilities that require future study.

In conclusion, we have identified two new likely pathogenic *FGFR1* mutations and 19 new candidate nHH/KS genes. These additional candidates will require future study in a large cohort of nHH/KS patients to determine their role in the phenotype. Although targeted capture sequencing permits the study of a large number of genes simultaneously, technical improvements of NGS are necessary before widespread use can be considered for routine diagnostic utility. Since the cost of whole exome sequencing is now competitive with targeted NGS done in our study, it is likely this method will be utilized for our future studies. Nevertheless, those variants confirmed by Sanger sequencing in this study yield important new genes for consideration in the pathogenesis of nHH/KS and a role in normal puberty.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Figure 1.

Digenic and trigenic pedigrees are shown (A-I), as well as an informative *CCKBR* family (J). Each pedigree is labeled with a letter A to J, which corresponds to the specific variant of the same letter in Table 4. Only variants identified by targeted NGS that were positively confirmed by Sanger sequencing are included in the analysis.

Figure 2

Two families of possible digenic or trigenic disease are shown, which are not informative. Since the variants do not segregate with nHH/KS, it is likely they are polymorphisms. Note for Figure 2A, the genotype of the father for *SMO* was not able to be done due to insufficient DNA.

Table 1. 261 genes selected for targeted sequencing are shown along with the their proposed relevance to nHH/KS. Some genes may be listed in more than one column. References for gene selection are shown in Supplemental Table 1.

Table 2. (A) Variants identified by targeted NGS in 48 nHH/KS, all of which were confirmed by Sanger sequencing, are shown. WT = wild type; HET = heterozygous; KS = Kallmann syndrome; nHH = normosmic hypogonadotropic hypogonadism. Findings in available family members are also shown. Note that although variants in some genes may not appear to segregate with the phenotype in a monogenic fashion (indicated by an asterisk), when evaluated in the context of digenic inheritance, these are still viable potentially pathogenic variants (see Figure 1). (B) Shown are the variants that are probable polymorphisms as they do not segregate with nHH/KS in a monogenic or digenic fashion.

Table 3. Class 3 missense variants are shown that were confirmed by Sanger sequencing. They were predicted deleterious if $\geq$2 or 3 programs (SIFT, PolyPhen2, MutationTaster) indicated so.

ACCEPTED MANUSCRIPT

Table 4. Digenic and trigenic pedigrees are shown. They are labelled A-I and each variant corresponds to the same variant indicated by the same letter as in Figure 1. Only variants confirmed by Sanger sequencing (class 3 or 4) are included.

Table 5. The 19 new candidate genes are shown, and where relevant, human phenotypes in OMIM are referenced.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

References

[1] Layman, L.C., 2013. Clinical genetic testing for kallmann syndrome, J Clin Endocrinol Metab. 98, 1860-2.

[2] Layman, L.C., 2013. The genetic basis of female reproductive disorders: Etiology and clinical testing, Mol Cell Endocrinol. 370, 138-48.

[3] Bhagavath, B., Podolsky, R.H., Ozata, M. et al., 2006. Clinical and molecular characterization of a large sample of patients with hypogonadotropic hypogonadism, Fertil Steril. 85, 706-13.

[4] Hanchate, N.K., Giacobini, P., Lhuillier, P. et al., 2012. SEMA3A, a gene involved in axonal pathfinding, is mutated in patients with Kallmann syndrome, PLoS Genet. 8, e1002896.

[5] Tornberg, J., Sykiotis, G.P., Keefe, K. et al., 2011. Heparan sulfate 6-O-sulfotransferase 1, a gene involved in extracellular sugar modifications, is mutated in patients with idiopathic hypogonadotrophic hypogonadism, Proc Natl Acad Sci U S A. 108, 11524-9.

[6] Margolin, D.H., Kousi, M., Chan, Y.M. et al., 2013. Ataxia, dementia, and hypogonadotropism caused by disordered ubiquitination, N Engl J Med. 368, 1992-2003.

[7] Kim, H.G., Ahn, J.W., Kurth, I. 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, 465-79.

[8] Kim, H.G., Kurth, I., Lan, F. et al., 2008. Mutations in CHD7, encoding a chromatin-remodeling protein, cause idiopathic hypogonadotropic hypogonadism and Kallmann syndrome, Am J Hum Genet. 83, 511-9.

[9] Quaynor, S.D., Kim, H.G., Cappello, E.M. et al., 2011. The prevalence of digenic mutations in patients with normosmic hypogonadotropic hypogonadism and Kallmann syndrome, Fertil Steril. 96, 1424-1430 e6.

[10] Sykiotis, G.P., Plummer, L., Hughes, V.A. et al., 2010. Oligogenic basis of isolated gonadotropin-releasing hormone deficiency, Proc Natl Acad Sci U S A. 107, 15140-4.

[11] Layman, L.C., Cohen, D.P., Jin, M. et al., 1998. Mutations in gonadotropin-releasing hormone receptor gene cause hypogonadotropic hypogonadism, Nat Genet. 18, 14-5.

[12] Quaynor, S.D., Stradtman, E.W., Jr., Kim, H.G. et al., 2013. Delayed puberty and estrogen resistance in a woman with estrogen receptor alpha variant, N Engl J Med. 369, 164-71.

[13] Metzker, M.L., 2010. Sequencing technologies - the next generation, Nat Rev Genet. 11, 31-46.

[14] Xu, N., Kim, H.G., Bhagavath, B. et al., 2011. Nasal embryonic LHRH factor (NELF) mutations in patients with normosmic hypogonadotropic hypogonadism and Kallmann syndrome, Fertil Steril. 95, 1613-20 e1-7.

[15] Richards, S., Aziz, N., Bale, S. et al., 2015. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology, Genet Med. 17, 405-24.

[16] Ng, P.C. and Henikoff, S., 2003. SIFT: Predicting amino acid changes that affect protein function, Nucleic Acids Res. 31, 3812-4.

[17] Schwarz, J.M., Rodelsperger, C., Schuelke, M. et al., 2010. MutationTaster evaluates disease-causing potential of sequence alterations, Nat Methods. 7, 575-6.

[18] Adzhubei, I.A., Schmidt, S., Peshkin, L. et al., 2010. A method and server for predicting damaging missense mutations, Nat Methods. 7, 248-9.

[19] Layman, L.C., Lee, E.J., Peak, D.B. et al., 1997. Delayed puberty and hypogonadism caused by mutations in the follicle-stimulating hormone beta-subunit gene, N Engl J Med. 337, 607-11.

[20] Dever, T.E., 2012. Molecular biology. A new start for protein synthesis, Science. 336, 1645-6.

[21] Kim, H.G., Bhagavath, B. and Layman, L.C., 2008. Clinical Manifestations of Impaired GnRH Neuron Development and Function, Neurosignals. 16, 165-182.

ACCEPTED MANUSCRIPT

[22] Pitteloud, N., Acierno, J.S., Jr., Meysing, A. et al., 2006. 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.

[23] Xu, N., Qin, Y., Reindollar, R.H. et al., 2007. A mutation in the fibroblast growth factor receptor 1 gene causes fully penetrant normosmic isolated hypogonadotropic hypogonadism, J Clin Endocrinol Metab. 92, 1155-8.

[24] Yang, Y., Muzny, D.M., Reid, J.G. et al., 2013. Clinical whole-exome sequencing for the diagnosis of mendelian disorders, N Engl J Med. 369, 1502-11.

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

[26] Sarfati, J., Guiochon-Mantel, A., Rondard, P. et al., 2010. A comparative phenotypic study of kallmann syndrome patients carrying monoallelic and biallelic mutations in the prokineticin 2 or prokineticin receptor 2 genes, J Clin Endocrinol Metab. 95, 659-69.

[27] Schaffer, A.A., 2013. Digenic inheritance in medical genetics, J Med Genet. 50, 641-52.

[28] Sykiotis, G.P., Hoang, X.H., Avbelj, M. et al., 2010. Congenital idiopathic hypogonadotropic hypogonadism: evidence of defects in the hypothalamus, pituitary, and testes, J Clin Endocrinol Metab. 95, 3019-27.

[29] Schwarting, G.A., Henion, T.R., Nugent, J.D. et al., 2006. Stromal cell-derived factor-1 (chemokine C-X-C motif ligand 12) and chemokine C-X-C motif receptor 4 are required for migration of gonadotropin-releasing hormone neurons to the forebrain, J Neurosci. 26, 6834-40.

[30] Gu, Y., Runyan, C., Shoemaker, A. et al., 2009. Steel factor controls primordial germ cell survival and motility from the time of their specification in the allantois, and provides a continuous niche throughout their migration, Development. 136, 1295-303.

[31] Mohammad, S., Ozaki, T., Takeuchi, K. et al., 2012. Functional compensation between cholecystokinin-1 and -2 receptors in murine paraventricular nucleus neurons, J Biol Chem. 287, 39391-401.

[32] Wang, Y., Shirogane, T., Liu, D. et al., 2003. Exit from exit: resetting the cell cycle through Amn1 inhibition of G protein signaling, Cell. 112, 697-709.

[33] Wunderer, F., Kuhne, S., Jilg, A. et al., 2013. Clock gene expression in the human pituitary gland, Endocrinology. 154, 2046-57.

[34] Itoh, N. and Ornitz, D.M., 2004. Evolution of the Fgf and Fgfr gene families, Trends Genet. 20, 563-9.

[35] Strittmatter, S.M., Fankhauser, C., Huang, P.L. et al., 1995. Neuronal pathfinding is abnormal in mice lacking the neuronal growth cone protein GAP-43, Cell. 80, 445-52.

[36] Balmer, C.W. and LaMantia, A.S., 2004. Loss of Gli3 and Shh function disrupts olfactory axon trajectories, J Comp Neurol. 472, 292-307.

[37] Ratie, L., Ware, M., Barloy-Hubler, F. et al., 2013. Novel genes upregulated when NOTCH signalling is disrupted during hypothalamic development, Neural Dev. 8, 25.

[38] Johnson, H.J., Gandhi, M.J., Shafizadeh, E. et al., 2009. In vivo inactivation of MASTL kinase results in thrombocytopenia, Exp Hematol. 37, 901-8.

[39] Leshan, R.L., Greenwald-Yarnell, M., Patterson, C.M. et al., 2012. Leptin action through hypothalamic nitric oxide synthase-1-expressing neurons controls energy balance, Nat Med. 18, 820-3.

[40] Garrel, G., Simon, V., Thieulant, M.L. et al., 2010. Sustained gonadotropin-releasing hormone stimulation mobilizes the cAMP/PKA pathway to induce nitric oxide synthase type 1 expression in rat pituitary cells in vitro and in vivo at proestrus, Biol Reprod. 82, 1170-9.

[41] Cariboni, A., Davidson, K., Rakic, S. et al., 2011. Defective gonadotropin-releasing hormone neuron migration in mice lacking SEMA3A signalling through NRP1 and NRP2: implications for the aetiology of hypogonadotropic hypogonadism, Hum Mol Genet. 20, 336-44.

[42] Panza, E., Gimelli, G., Passalacqua, M. et al., 2007. The breakpoint identified in a balanced de novo translocation t(7;9)(p14.1;q31.3) disrupts the A-kinase (PRKA) anchor protein 2 gene (AKAP2) on chromosome 9 in a patient with Kallmann syndrome and bone anomalies, Int J Mol Med. 19, 429-35.

[43] Masciarelli, S., Horner, K., Liu, C. et al., 2004. Cyclic nucleotide phosphodiesterase 3A-deficient mice as a model of female infertility, J Clin Invest. 114, 196-205.

ACCEPTED MANUSCRIPT

[44] Dowler, S., Currie, R.A., Campbell, D.G. et al., 2000. Identification of pleckstrin-homology-domain-containing proteins with novel phosphoinositide-binding specificities, Biochem J. 351, 19-31.

[45] Friedman, J.S., Chang, B., Kannabiran, C. et al., 2006. Premature truncation of a novel protein, RD3, exhibiting subnuclear localization is associated with retinal degeneration, Am J Hum Genet. 79, 1059-70.

[46] Sunryd, J.C., Cheon, B., Graham, J.B. et al., 2014. TMTC1 and TMTC2 are novel endoplasmic reticulum tetratricopeptide repeat-containing adapter proteins involved in calcium homeostasis, J Biol Chem. 289, 16085-99.

[47] Lorenc, A., Linnenbrink, M., Montero, I. et al., 2014. Genetic differentiation of hypothalamus parentally biased transcripts in populations of the house mouse implicate the Prader-Willi syndrome imprinted region as a possible source of behavioral divergence, Mol Biol Evol. 31, 3240-9.

[48] Bassani, S. and Cingolani, L.A., 2012. Tetraspanins: Interactions and interplay with integrins, Int J Biochem Cell Biol. 44, 703-8.

[49] Pabinger, S., Dander, A., Fischer, M. et al., 2014. A survey of tools for variant analysis of next-generation genome sequencing data, Brief Bioinform. 15, 256-78.

ACCEPTED MANUSCRIPT

Olfactory Bulb Development	Hypothalamic Development/ Patterning	GnRH Regulation	GnRH Neuron Development/ Migration	Chromosomal Rearrangements	Known Pathway Genes	Molecules on GnRH Neuron Surfaces & in Cytoplasm	Pituitary Related
ACSM4	*ALDH1A1*	*BRN2*	*AKAP2*	*AEBP2*	*AKAP1*	*BDNF*	*CGA*
BPIFB4	*ALDH1A2*	*CRY1*	*AXL*	*AMN1*	*AKAP3*	*CACNA1B*	*FSHB*
CHD7	*BMP2*	*CRY2*	*BRN2*	*ANKRD26*	*AKAP4*	*CCKAR*	*HESX1*
CTXN3	*BMP4*	*DLX1*	*CAS*	*ARNTL2*	*CCDC141*	*CCKBR*	*LHB*
DLX1	*BMP7*	*DLX2*	*CHAT*	*ASUN*	*FGF1*	*CNTN2*	*LHX3*
EMX2	*DHH*	*DLX5*	*CXCL12*	*ATE1*	*FGF10*	*DCC*	*LHX4*
FSTL5	*EMX1*	*EGFR*	*CXCR4*	*C12ORF70*	*FGF11*	*EFNA5*	*PIT1*
GADL1	*EMX2*	*EGR1*	*DCC*	*C12ORF71*	*FGF12*	*GABABR1*	*PITX2*
GPX6	*FOXH1*	*FGF8*	*DLX1*	*CADM2*	*FGF13*	*GAL*	*PROP1*
GSTM1	*GDF1*	*FGFR1*	*EBF2*	*CAPRIN2*	*FGF14*	*GALR1*	*PTX1*
MASH1	*GLI1*	*GATA2*	*GAS6*	*CCDC91*	*FGF16*	*GALR2*	*RUNX1*
MATH4A	*GLI2*	*GATA3*	*GATA4*	*CMAS*	*FGF17*	*GALR3*	*RUNX2*
NCAM	*GLI3*	*GATA4*	*HGF*	*ERGIC2*	*FGF18*	*GAP43*	*RUNX3*
NEUROD	*LHX5*	*GNRH1*	*IL17RD*	*ETNK1*	*FGF19*	*GAS6*	*SOX2*
NEUROG2	*LIM1*	*GNRHR*	*JAG1*	*FAM60A*	*FGF2*	*GDNF*	*SOX3*
NOTCH1	*NODAL*	*GRG*	*KAL1*	*FAR2*	*FGF20*	*GRG4*	
NRP1	*NROB1*	*HCRT*	*MAP1L*	*FGFR2*	*FGF21*	*GRG5*	
NTN1	*PAX6*	*HCRTR1*	*MAPK14*	*H3F3C*	*FGF22*	*GRIN1*	
OBP2B	*PAX7*	*HS6ST1*	*MSX1*	*IPO8*	*FGF3*	*GRIN2*	
OMP	*PCSK1*	*KISS1*	*MYCN*	*KIAA0528*	*FGF4*	*HGF*	
P75NTR	*POMC*	*KISS1R*	*NCAM*	*KLHDC5*	*FGF5*	*IGF1*	
PAX6	*PROKR1*	*KLF6*	*NELF*	*MANSC4*	*FGF6*	*IGF1R*	
PCDH8	*PTC1*	*LEP*	*NRP1*	*MASTL*	*FGF7*	*IL17RD*	
PLXNA1	*PTCH1*	*LEPR*	*NRP2*	*MCMBP*	*FGF8*	*LRP8*	
PLXNB1	*SFRP1*	*MEIS1*	*NSCL2*	*MED21*	*FGF9*	*MET*	
PROK2	*SFRP5*	*MEOX1*	*NTN1*	*METTL20*	*FGFR1*	*NCAM*	
PROKR2	*SHH*	*MSX1*	*OCT1*	*MRPS35*	*FGFR1OP2*	*NDN*	
RTP1	*SIX3*	*MSX2*	*PKNOX1*	*NSMCE4A*	*FGFR3*	*NELF*	
RTP2	*SMAD5*	*NDN*	*PKNOX2*	*OVCH1*	*IFT57*	*NTN*	
SCGB1c1	*SMO*	*NF1*	*PLXNB1*	*PDE3A*	*KCNK9*	*PTHLH*	
SEC14L3	*TBX2*	*NK2*	*RAC*	*PENK*	*KCTD11*	*SPRY4*	
SEMA3A	*TGFB*	*OTX2*	*RELN*	*PLEKHA5*	*NOS1*	*STMN1*	
SLC17A6	*WNT8b*	*OCT1*	*SCIP*	*PPAPDC1A*	*PTCH2*	*TAG-1*	
SP8		*POU1F1*	*SEMA3A*	*PPFIBP1*	*RD3*	*TGFB*	
TBR1		*POU5F1*	*SEMA4D*	*REP15*	*SHH*		
UMODL1		*PREP1*	*SH2D3C*	*SDR16C5*	*STIL*		
WDR11		*SIX6*	*TBR1*	*SEC23IP*	*SUFU*		
		SMAD3	*TBX2*	*SLCO1C1*	*SYCP1*		
		SMAD4	*TUBB2A*	*ST8SIA1*	*TRAPPC9*		
		TAC3	*TUBB2B*	*STK38L*	*TRIM25*		
		TACR3	*TUBB3*	*TM7SF3*	*ZFP423*		
		WNT11		*TMTC1*			
				TSPAN11			
				WDR11			
				YME1L1			

ACCEPTED MANUSCRIPT

IHH	Gene	Accession Number	Variant	# Subjects	ExAC Frequency	% ExAC Frequency	Subject Phenotype	Family History
Class 4		-	-	-	-	-	-	-
225	***FGFR1***	NM_001174063	c.214C>T, p.Q72X	1	0	0	nHH	mother WT, unaffected
262	***FGFR1***	NM_001174063	c.2248delG, p.E750Kfs*9	1	0	0	nHH	mother WT, unaffected
Class 3								
234	***AMN1***	NM_001113402	c.656G>A, p.C219Y	1	0	0	nHH	father HET, unaffected; sister1 HET, IHH; sister2 HET, unaffected
061, 302	***AXL***	NM_001699	c.-40C>T, p.L1extM-13	2	39/17382	0.2244	subject C: nHH subject D: nHH	subject C: father WT, unaffected; subject D: mother WT, unaffected
262	***AXL***	NM_001699	c.568C>T, p.R190C	1	1/57944	0.001726	nHH	mother WT, unaffected
210	***CCKBR***	NM_176875	c.500C>T, p.T167M	1	3/118692	0.002528	KS, bilateral hearing aids, blind left eye, anosmia, benign ovarian tumor, no breasts	father (no DNA) undescended testes; brother1 HET, IHH hyposmia, undescended testes, hearing loss, anger issues, kidney and stomach cancer; brother2 HET, normosmic, deceased, brain glioma; sister HET, KS, hearing loss, anosmia, bilateral amblyopia, sister-in-law WT, unaffected; nephew1 HET, hyposmic with undescended testes; nephew2 WT, unaffected
302	***CRY1***	NM_004075	c.1292G>A,p.R431H	1	29/119292	0.02431	nHH	mother WT, unaffected
61	***CXCR4***	NM_003467	c.236A>C, p.H79P	1	0	0	nHH	father WT, unaffected
217	***FGF13***	NM_033642	c.529T>C, p.S177P	1	0	0	KS	mother WT, unaffected; father WT, unaffected; brother WT unaffected
235	***FGFR1***	NM_001174063	c.2053G>A, p.G685R	1	0	0	nHH	mother WT, unaffected
217	***GAP43***	NM_002045	c.502G>A , p.A168T	1	0	0	KS	father HET, unaffected; mother WT, unaffected; brother HET, unaffected
225	***GLI3***	NM_000168	c.341G>A, p.R114K	1	304/101458	0.2996	nHH	mother WT, unaffected
245	***GLI3***	NM_000168	c.245G>A, p.R82K	1	2/117682	0.00169	nHH	brother 1 HET nHH brother 2 HET unaffected
224	***GNRH1***	NM_000825	c.103C>T, p.R35C	1	0	0	nHH	father HET, unaffected; mother WT, unaffected; sister WT, unaffected

ACCEPTED MANUSCRIPT

245	***GNRHR***	NM_000406	c.662T>A, p.I221N	1	0	0	nHH	brother 1 HET nHH brother 2 WT unaffected
233	***IL17RD***	NM_017563	c.965G>A, p.R322H	1	0	0	pituitary tumor	mother HET, unaffected
234	***JAG1***	NM_000214	c.323A>T, p.N108I	1	0	0	nHH	father WT, unaffected; sister WT, unaffected; sister HET, nHH
302	***MASTL***	NM_001172303	c.-170C>T, p.M1extM-57	1	0	0	nHH	mother WT, unaffected
216	***NOS1***	NM_001204213	c.1007G>A, p.R336H	1	0	0	nHH	sister WT, unaffected
119	***NOTCH1***	NM_017617	c.2333C>T, p.T778I	1	1/118624	0.000843	nHH	mother WT, unaffected
310	***NRP2***	NM_003872	c.2057T>G, p.F686C	1	0	0	nHH, 1˚ amenorrhea	mother WT, unaffected; grandmother WT, unaffected
258	***PALM2***	NM_001037293	c.451A>G, p.T151A	1	0	0	KS, 1˚ amenorrhea, Graves disease	none
128	***PDE3A***	NM_001244683	c.1477G>A, p.G493R	1	0	0	nHH	father HET, unaffected; mother WT, unaffected sister HET, unaffected
247	***PLEKHA5***	NM_001143821	c.884A>T, p.N295Y	1	0	0	Unknown sense of smell	mother HET unaffected
216	***RD3***	NM_001164688	c.509C>T, p.S170F	1	3/105638	0.00284	nHH	sister WT, unaffected
307	***TMTC1***	NM_175861	c.1978G>A, p.D660N	1	0	0	KS, bilateral cryptorchidism, left orchiopexy, myopia, gynecomastia	father HET, unaffected mother WT, unaffected
128	***TRAPPC9***	NM_001160372	c.1532C>T, p.T511M	1	0	0	nHH	father WT, unaffected; mother HET, unaffected sister WT, unaffected
AO2	***TSPAN11***	NM_001080509	c.203G>A, p.G68D	1	0	0	KS, *NELF* mutation	father HET, hyposmic, blepharospasm; mother WT, unaffected

ACCEPTED MANUSCRIPT

IHH	Gene	Accession Number	Variant	# Subjects	ExAC Frequency	% ExAC Frequency	Subject Phenotype	Family History
Class 3	_	_	_	_	_	_	_	_
165	***FGFR1****	NM_001174063	c.2180+3insT	1	0	0	nHH	father HET unaffected; mother WT, unaffected; sister HET, nHH
208	***FGFR1****	NM_001174063	c.301T>G, p.C101G	1	0	0	KS	father HET, unaffected; mother WT, unaffected; brother HET, anosmic
6	***FGFR3****	NM_022965	c.962C>T, p.S321F	1	0	0	KS, midline high palate	brother1 HET, 45,X/46,XY; brother2 WT, 46,XY, unaffected
224	***GADL1****	NM_207359	c.1480C>T, p.R494C	1	2/121188	0.00165	nHH	father HET, unaffected; mother WT, unaffected; sister WT, unaffected
123	***GPRIN1****	NM_052899	c.2884C>T, p.R962C	1	0	0	KS, unilateral undescended testes	father WT, unaffected; mother HOM, unaffected
165	***POMC****	NM_000939	c.777C>G, p.I259M	1	0	0	nHH	father WT, unaffected; mother HET, unaffected; sister HET, nHH
123	***SEMA4D****	NM_006378	c.2278C>A, p.L760M	1	0	0	KS	father WT, unaffected; mother HET, unaffected
123	***SMO****	NM_005631	c.1789G>A, p.D597N	1	0	0	KS	mother HET, unaffected

ACCEPTED MANUSCRIPT

Gene	Accession Number	Variant	SIFT	PolyPhen2	MutationTaster
AXL	NM_001699	c.568C>T, p.R190C	Not tolerated; p=0.03	Possibly damaging with a score of 0.909	polymorphism
CCKBR	NM_176875	c.500C>T, p.T167M	Not tolerated; p=0.00	Probably damaging with a score of 1.000	disease causing
CRY1	NM_004075	c.1292G>A,p.R431H	Not tolerated; p=0.00	Benign with a score of 0.047	disease causing
CXCR4	NM_003467	c.236A>C, p.H79P	Not tolerated; p=0.00	Probably damaging with a score of 1.000	disease causing
FGF13	NM_033642	c.529T>C, p.S177P	Not tolerated; p=0.00	Error	disease causing
FGFR1	NM_001174063	c.2053G>A, p.G685R	Not tolerated; p=0.00	Error	disease causing
FGFR1*	NM_001174063	c.301T>G, p.C101G	Not tolerated; p=0.00	Probably damaging with a score of 1.000	disease causing
FGFR3*	NM_022965	c.962C>T, p.S321F	Not tolerated; p=0.00	Error	disease causing
GADL1*	NM_207359	c.1480C>T, p.R494C	Not tolerated; p=0.00	Probably damaging with a score of 1.000	disease causing
GAP43	NM_002045	c.502G>A , p.A168T	Tolerated; p=0.30	Probably damaging with a score of 1.000	disease causing
GLI3	NM_000168	c.341G>A, p.R114K	Not tolerated; p=0.02	Probably damaging with a score of 0.995	disease causing
GLI3	NM_000168	c.245G>A, p.R82K	Not tolerated; p=0.00	Probably damaging with a score of 0.994	disease causing
GNRH1	NM_000825	c.103C>T, p.R35C	Not tolerated; p=0.00	Probably damaging with a score of 1.000	disease causing
GNRHR	NM_000406	c.662T>A, p.I221N	Not tolerated; p=0.00	Possibly damaging with a score of 0.912	disease causing
GPRIN1*	NM_052899	c.2884C>T, p.R962C	Not tolerated; p=0.03	Benign with a score of 0.046	disease causing
IL17RD	NM_017563	c.965G>A, p.R322H	Not tolerated; p=0.03	Probably damaging with a score of 1.000	disease causing
JAG1	NM_000214	c.323A>T, p.N108I	Not tolerated; p=0.04	Possibly damaging with a score of 0.645	disease causing
NOS1	NM_001204213	c.1007G>A, p.R336H	Not tolerated; p=0.03	Error	disease causing
NOTCH1	NM_017617	c.2333C>T, p.T778I	Not tolerated; p=0.04	Probably damaging with a score of 0.963	disease causing
NRP2	NM_003872	c.2057T>G, p.F686C	Not tolerated; p= 0.00	Error	disease causing
PALM2	NM_001037293	c.451A>G, p.T151A	Not tolerated; p=0.00	Possibly damaging with a score of 0.760	disease causing
PDE3A	NM_001244683	c.1477G>A, p.G493R	Not tolerated; p=0.04	Error	disease causing
PLEKHA5	NM_001143821	c.884A>T, p.N295Y	Not tolerated; p=0.00	Probably damaging with a score of 1.000	disease causing
POMC*	NM_000939	c.777C>G, p.I259M	Not tolerated; p=0.04	Probably damaging with a score of 0.970	disease causing
RD3	NM_001164688	c.509C>T, p.S170F	Not tolerated; p=0.00	Probably damaging with a score of 1.000	disease causing
SEMA4D*	NM_006378	c.2278C>A, p.L760M	Not tolerated; p=0.00	Probably damaging with a score of 1.000	disease causing
SMO*	NM_005631	c.1789G>A, p.D597N	Not tolerated; p=0.02	Possibly damaging with a score of 0.566	disease causing
TMTC1	NM_175861	c.1978G>A, p.D660N	Not tolerated; p=0.04	Error	polymorphism
TRAPPC9	NM_001160372	c.1532C>T, p.T511M	Not tolerated; p=0.00	Probably damaging with a score of 1.000	disease causing
TSPAN11	NM_001080509	c.203G>A, p.G68D	Not tolerated; p=0.00	Probably damaging with a score of 1.000	disease causing

ACCEPTED MANUSCRIPT

nHH/KS Digenic and Trigenic Variants

Subject	Variant 1	Variant 2	Variant 3
A, 225	***FGFR1*** c.313C>T, p.Q72X	***GLI3*** c.341G>A, p.R114K	
B, 262	***AXL*** c.568C>T, p.R190C	***FGFR1*** c.2248delG, p. E750Kfs*9	
C, 061	***AXL*** c.-40C>T p.L1extM-13,	***CXCR4*** c.236A>C, p.H79P	
D, 302	***AXL*** c.-40C>T, p.L1extM-13	***CRY1*** c.1292G>A, p.R431H	***MASTL*** c.-170C>T, p.M1extM-57
E 216	***NOS1*** c.1007G>A, p.R336H	***RD3*** c.509C>T, p.S170F	
F, 128	***TRAPPC9*** c.1532C>T, p.T511M	***PDE3A*** c.1477G>A, p.G493R	
G, 217	***FGF13*** c.529T>C, p.S177P	***GAP43*** c.502G>A, p.A168T	
H, 234	***AMN1*** c.656G>A, p.C219Y	***JAG1*** c.323A>T, p.N108I	
I, 245	***GNRHR*** c.662T>A, p.I221N	***GLI3*** c.245G>A, p.R82K 662T>A,	

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Gene	Name	Locus	Function	Reference
1. *AMN1*	Antagonist of mitotic exit network 1 homolog	12p11.21	Positional; also part of daughter-specific switch helping cells exit and reset cell cycle via G-protein inhibition	Bhagavath (3) Wang (32)
2. *CCKBR*	Cholecystokinin B receptor	11p15.4	G-protein coupled receptor neuropeptide for gastrin and cholecystokinin, both of which are regulatory peptides of the brain and gastrointestinal tract; expressed in paraventricular nucleus of hypothalamus	Mohammad (31)
3. *CRY1*	Cytochrome circadian clock 1	12q23.3	Pituitary clock gene, binds as multiprotein complex in circadian feedback loop of CLOCK-BMAL	Wunderer (33)
4. *CXCR4*	C-X-C motif chemokine receptor 4	2q21	Affects sensory axon extension and GnRH1 neuron migration; maintains GnRH1 neuronal expression as the cells move away from nasal pit region	Schwarting (29)
5. *FGF13*	Fibroblast growth factor 13	Xq26.3-q27.1	FGFs known in nHH/KS pathway expressed in developing & adult nervous systems (predominantly in brain-cerebellum & cortex)	Itoh (34)
6. *GAP43*	Growth associated protein 43	3q13.31	Linked to synaptosomal membrane; regulates growth of axons & moderates formation of new connections	Strittmatter (35)
7. *GLI3*	GLI family zinc finger 3	7p14.1	Transcriptional regulator of sonic hedgehog involved in induction, patterning, & differentiation of body regions including face & forebrain; affects olfactory receptor axons & directed outgrowth to telecephalon; haploinsufficiency results in Greig Cephalopolysyndactyly (MIM 175700) & Pallister Hall syndrome (MIM 146510); (PHS includes hypothalamic hamartomas)	Balmer (36)

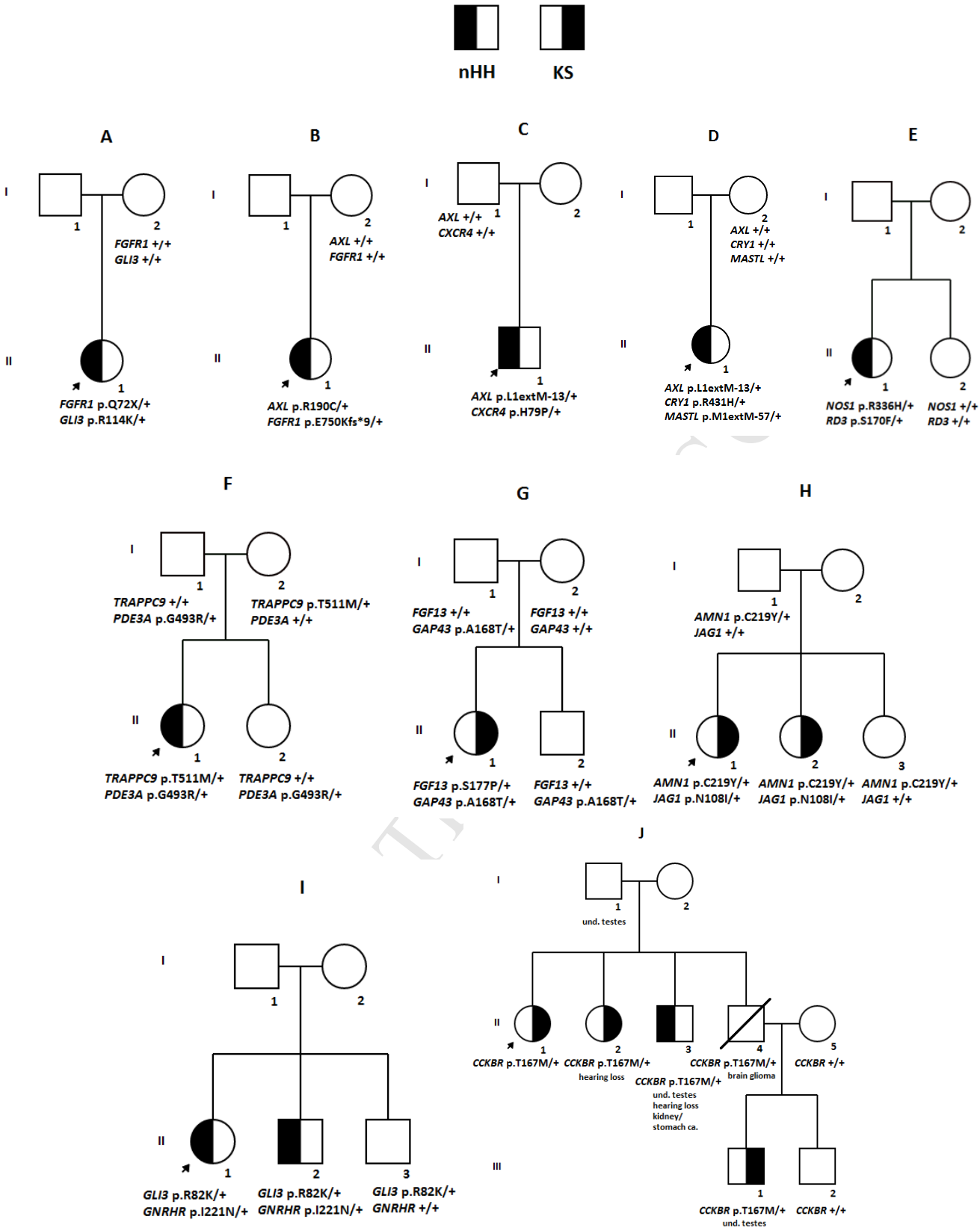
ACCEPTED MANUSCRIPT

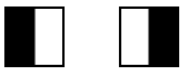
8. *JAG1*	Jagged 1	20p12.2	Ligand of Notch receptor, which is increased by WNT signaling molecules (beta-catennin & TCF); activates transcription factors important in cell differentiation & morphogenesis; mutations in Alagille syndrome (MIM 118450)-facial, cardiac, ocular, & skeletal defects; cholestasis	Ratie (37)
9. *MASTL*	Microtubule-associated serine/ threonine kinase like	10p12.1	Positional candidate Thrombocytopenia (MIM 608221); mutations in kinase domain in Drosophila disrupt cell division	Bhagavath (3) Johnson (38)
10. *NOS1*	Nitric oxide synthase 1	12q24.22	Molecular messenger with diverse brain & peripheral nervous system functions; GnRH regulates (increases) NOS1 in pituitary gonadotropes	Garrel (40)
11. *NOTCH1*	Notch, Drosophila Homologue of, 1	9q34.3	See JAG1 above; mutations in Adams-Oliver syndrome (MIM 616028); aortic valve disease 1 (MIM 109730)	Ratie (37)
12. *NRP2*	Neuropilin 2	2q33.3	Neuropilins are receptors for semaphorins (*SEMA3A* is known KS gene); expressed on vomeronasal nerve, which serves as guide for GnRH neuron migration	Cariboni (41)
13. *PALM2*	Paralemmin 2	9q31.3	Disrupted in KS patient with translocation of 9q31.3;also known as A-kinase anchor protein 2; expressed in olfactory bulb & cartilaginous structures of embryo	Panza (42)
14. *PDE3A*	Phosphodiesterase 3A	12p12	Positional candidate Mutations in brachydactyly & hypertension syndrome (MIM 112410); PDE3A is responsible for hydrolysis of cAMP activity of oocyte (needed for resumption meiosis in egg)	Bhagavath (3) Masciarelli (43)
15. *PLEKHA5*	Plekstrin homology domain-containing protein, family A, member 5	12p12.3	Positional candidate Expressed at highest levels in ovary, but also in brain; interacts with phosphatidylinositol 3,4,5-trisphosphate (important 2nd messenger in G-protein coupled	Bhagavath (3) Dowler (44)

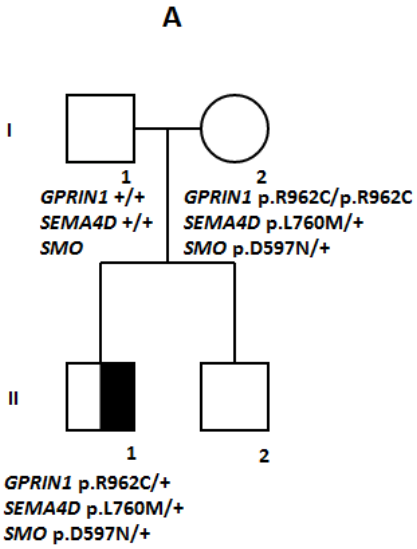
ACCEPTED MANUSCRIPT

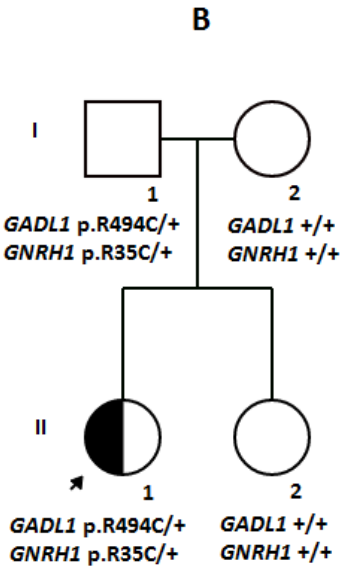
			receptors (nHH/KS genes)	
16. *RD3*	Retinal degeneration 3 Mouse, homolog of	1q32.3	Highly expressed in retina; part of subnuclear protein complex involved in diverse processes, such as transcription and splicing; Mutations in Leber congenital amaurosis-12 (MIM 610612)	Friedman (45)
17. *TMTC1*	Transmembrane and tetratricopeptide repeat domains-containing protein 1	12p11.22	Positional candidate; also endoplasmic reticulum protein involved in calcium homeostasis	Bhagavath (3) Sunryd (46)
18. *TRAPPC9*	Trafficking protein particle complex, subunit 9	8q24.3	Expressed in neurons of the cerebral cortex, hippocampus, and deep gray matter; protein binds NF-kappa-B inducing kinase (NIK) and inhibitor of kappa light chain gene enhancer in B cells, kinase of (IKK-beta); involved in NF-kappa signaling pathway; paternally expressed & implicated in Prader-Willi; mutations in mental retardation, autosomal recessive (MIM 613192)	Lorenc (47)
19. *TSPAN11*	Tetraspanin 11	12p11.21	Positional candidate; also integral membrane protein regulating cell adhesion, motility, and synapse formation; interacts with integrins	Bhagavath (3) Bassani (48)

ACCEPTED MANUSCRIPT


nHH
KS
A
I
1
2
FGFR1 +/+
GLI3 +/+
II
1
FGFR1 p.Q72X/+
GLI3 p.R114K/+
B
I
1
2
AXL +/+
FGFR1 +/+
II
1
AXL p.R190C/+
FGFR1 p.E750Kfs*9/+
C
I
AXL +/+
CXCR4 +/+
1
2
II
1
AXL p.L1extM-13/+
CXCR4 p.H79P/+
D
I
1
2
AXL +/+
CRY1 +/+
MASTL +/+
II
1
AXL p.L1extM-13/+
CRY1 p.R431H/+
MASTL p.M1extM-57/+
E
I
1
2
II
1
2
NOS1 p.R336H/+
RD3 p.S170F/+
NOS1 +/+
RD3 +/+
F
I
1
2
TRAPPC9 +/+
PDE3A p.G493R/+
TRAPPC9 p.T511M/+
PDE3A +/+
II
1
2
TRAPPC9 p.T511M/+
PDE3A p.G493R/+
TRAPPC9 +/+
PDE3A p.G493R/+
G
I
1
2
FGF13 +/+
GAP43 p.A168T/+
FGF13 +/+
GAP43 +/+
II
1
2
FGF13 p.S177P/+
GAP43 p.A168T/+
FGF13 +/+
GAP43 p.A168T/+
H
I
1
2
AMN1 p.C219Y/+
JAG1 +/+
II
1
2
3
AMN1 p.C219Y/+
JAG1 p.N108I/+
AMN1 p.C219Y/+
JAG1 p.N108I/+
AMN1 p.C219Y/+
JAG1 +/+
I
I
1
2
II
1
2
3
GLI3 p.R82K/+
GNRHR p.I221N/+
GLI3 p.R82K/+
GNRHR p.I221N/+
GLI3 p.R82K/+
GNRHR +/+
J
I
1
2
und. testes
II
1
2
3
4
5
CCKBR p.T167M/+
CCKBR p.T167M/+
hearing loss
CCKBR p.T167M/+
brain glioma
CCKBR +/+
CCKBR p.T167M/+
und. testes
hearing loss
kidney/
stomach ca.
III
1
2
CCKBR p.T167M/+
und. testes
CCKBR +/+







ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Highlights—Quaynor et al.

- Targeted and Sanger DNA sequencing of 261 genes was performed in 48 patients with hypogonadotropic hypogonadism
- Two novel *FGFR1* likely pathogenic mutations were found
- Nineteen new candidate genes for hypogonadotropic hypogonadism were identified
- Eight possible digenic and one possible trigenic families were identified

ACCEPTED MANUSCRIPT