American Journal of Obstetrics and Gynecology (2005) **192**, 1274–84

ELSEVIER

American Journal of OBSTETRICS and GYNECOLOGY

www.ajog.org

Screening candidate genes for mutations in patients with hypogonadotropic hypogonadism using custom genome resequencing microarrays

Ning Xu, MB,[a,b] Robert H. Podolsky, PhD,[c] Pranav Chudgar, BS,[a,b] Lynn P. Chorich, MS,[a,b] Chunmei Liu, MS,[d] Paul G. McDonough, MD,[a] Janet A. Warrington, PhD,[d] Lawrence C. Layman, MD[a,b,*]

Section of Reproductive Endocrinology, Infertility, and Genetics,[a] Department of Obstetrics and Gynecology, Neurodevelopmental Biology Program,[b] The Institute of Molecular Medicine and Genetics, Department of Biostatistics and Bioinformatics,[c] The Medical College of Georgia, Augusta, Ga, and Affymetrix, Inc, Santa Clara, Calif [d]

Received for publication December 2, 2004; revised December 10, 2004; accepted December 10, 2004

KEY WORDS
Idiopathic hypogonadotropic hypogonadism
Genomic chip
Microarray
Custom genome resequencing microarray

Objective: The purpose of this study was to determine the consistency of calling single nucleotide polymorphisms (SNPs) by custom genome resequencing microarrays compared with capillary DNA sequencing.

Study design: Amplified genomic DNA from 23 patients with hypogonadotropic hypogonadism was hybridized to microarrays containing 30 kilobases of sequence from 6 different candidate genes. Capillary DNA sequencing was performed in 10 patients.

Results: For 10 patients with $\geq$90% of bases called, 49 SNPs in 5 of 6 genes were identified. Of the 490 bases, 75 were ambiguous (read as "N"), and 415 were able to be called an A, C, G, or T. Of 415 called, 401 (96.6%) sequences were confirmed by DNA sequencing. All homozygotes (285/285) were called identically, while sequence from 89.2% (116/130) of heterozygotes agreed by both methods. The level of agreement between microarray calls and capillary DNA sequencing demonstrated substantial accuracy.

Conclusion: Custom genome resequencing microarrays are highly consistent with capillary sequencing in calling individual bases in genomic DNA from patients with human disease.

Idiopathic hypogonadotropic hypogonadism (IHH) is a clinical entity comprised of absent puberty with subsequent infertility, low serum levels of the gonadotropins follicle stimulating hormone (FSH) and luteinizing hormone (LH), and the absence of a pituitary tumor.[1,2] The etiology of IHH is due to impaired gonadotropin releasing hormone (GnRH) or gonadotropin secretion and/or action. Kallmann syndrome consists of IHH and anosmia, as well as a variety of

Supported by NIH grants HD33004 & HD040287 (L.C.L.).

Presented at the 23rd Annual Meeting of the American Gynecological and Obstetrical Society, September 9-11, 2004, Bolton Landing, NY.

* Reprint requests: Lawrence C. Layman, MD, Section of Reproductive Endocrinology, Infertility, and Genetics, Department of Obstetrics and Gynecology, The Medical College of Georgia, 1120 15th St, Augusta, GA. 30912-3360.

E-mail: Llayman@mcg.edu

0002-9378/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ajog.2004.12.066

neurologic findings, midfacial defects, and renal agenesis.[1,2] Mutations in genes expressed in either the hypothalamus or pituitary, including KAL1, KAL2, LEP, LEPR, GNRHR, NROB1, and GPR54, account for the molecular basis of approximately 10% to 20% of IHH patients.[3-5] However, the molecular basis for most cases of IHH remains unknown.

Traditional methods to detect gene mutations include polymerase chain reaction (PCR)-based techniques such as single strand conformation polymorphism (SSCP)[6] or denaturing gradient gel electrophoresis (DGGE),[7] which are then confirmed by DNA sequencing. We have successfully used the techniques of DGGE and DNA sequencing to characterize human gene mutations in IHH patients.[8,9] However, the process may be laborious, and the fragment size to be screened is generally small (100-500 bp). The use of custom genome resequencing microarrays, previously known as variation detection arrays, could offer a high throughput technology of mutation detection compared to more traditional approaches.[10-14]

Microarrays have been used extensively to study gene expression,[14,15] but there are limited data regarding microarrays hybridized to genomic DNA for the purpose of gene mutation detection in humans.[10-14] Both microarrays and capillary DNA sequencing are DNA sequencing techniques. Microarrays are a hybridization-based DNA sequencing method, while the more commonly used technique is a capillary-based DNA sequencing method. Although custom genome resequencing microarrays have been shown in a previous study to be highly consistent with other methods of DNA sequencing when used to call single base differences in genomic DNA[13], the efficacy of this method to detect gene mutations in individual laboratories interested in human disease has not been reported. In this study, we wanted to determine the feasibility of using custom genome resequencing microarrays to screen IHH patients for mutations in 6 candidate genes, either proposed to cause IHH or known to cause IHH. All of the identified single base differences, hereafter called single nucleotide polymorphisms (SNPs), were then subjected to confirmation by capillary DNA sequencing, which served as the gold standard. The purpose of the present study was to evaluate the consistency of single base changes identified using microarrays when compared with capillary-based DNA sequencing. If these microarrays could be shown to be consistent with capillary sequencing methods, then they would be useful in research laboratories for screening for human gene mutations.

Material and methods

IHH patient characterization

IHH was defined as absent puberty ($\geq$17 years in females; $\geq$18 years in males), low serum gonadotropins, and normal imaging of the hypothalamus-pituitary region. Other pituitary hormones (TSH, prolactin, growth hormone) and cortisol were normal.[1] Males had a testosterone less than 100 ng/dL (300-1100), while females had absent breast development and primary amenorrhea. This study was approved by the Human Assurance Committee of the Medical College of Georgia.

Custom genome resequencing microarray design

Custom genome resequencing microarrays (CustomSeq™) were produced by Affymetrix (Santa Clara, Calif) using photolithography and solid-phase DNA synthesis.[16] Each microarray contained 30 kb of unique genomic sequence (both sense and antisense) of 6 candidate genes: GNRH1 (gonadotropin releasing hormone 1), GNRHR (gonadotropin releasing hormone receptor), FSHB (follicle stimulating hormone-beta), HESX1, POU3F2, and NELF (nasal embryonic LHRH factor). Each microarray contains 240,000 features that cover all of the specific sequences of genes to be tested, with each feature consisting of $\sim 10^6$ copies of a 25 bp specific probe.[13] To query each single base of the candidate genes, 4 features, which were identical except for the 13th base of the 25-mer (A, C, G, or T), were tiled on the microarray. Four additional features were used to query each base of the antisense strand of each gene (8 features for each of 30,000 nucleotides = 240,000 features).

These candidate genes were selected because of evidence of a principle role in the regulation of GnRH or gonadotropins: (1) a GNRH1 gene deletion has been observed in the hypogonadal mouse[17]; (2) GNRHR mutations have been demonstrated to cause IHH in humans[9]; (3) FSHB mutations cause isolated FSH deficiency, which has a similar phenotype to IHH[8] in humans, and Fshb knockout mice have low FSH and normal LH,[18] not unlike some human IHH patients; (4) HESX1 mutations cause pituitary failure in patients with septo-optic dysplasia[19]; (5) POU3F2 regulates GnRH expression[20]; and (6) nasal embryonic LHRH factor (NELF) is proposed to influence GnRH neuron migration in mice.[21] Any of these genes could potentially cause IHH in humans.

The sequence for each candidate gene was obtained from GenBank or the Human Genome database, then subjected to 2 programs to remove repetitive sequences: Repeat Masker (Institute for Systems Biology, Seattle, Wash) to identify repeat regions unsuitable for analysis (ie, SINE, LINE, ALU, etc)[13]; and Micropeats (Bioinformatics Computational Core Laboratories, Virginia Commonwealth University, Richmond, Va) to mark the repeats (overlap) among the sequences for the same chip design.[13] The size of initial and masked gene sequences included on each microarray are shown in Table I. Before microarray production, long-range PCR conditions were