The Journal of Molecular Diagnostics, Vol. 18, No. 1, January 2016

ELSEVIER

the Journal of Molecular Diagnostics
jmd.amjpathol.org

Characterization of Deletions of the *HBA* and *HBB* Loci by Array Comparative Genomic Hybridization

Daniel E. Sabath,* Michael A. Bender,[†‡] Vijay G. Sankaran,[§] Esther Vamos,[¶] Alex Kentsis,[‖] Hye-Son Yi,* and Harvey A. Greisman*

From the Departments of Laboratory Medicine and Pediatrics,[†] University of Washington, Seattle, Washington; the Clinical Research Division,[‡] Fred Hutchinson Cancer Research Center, Seattle, Washington; the Division of Hematology/Oncology,[§] Boston Children's Hospital and Department of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts; the Departments of Pediatrics and Genetics,[¶] Université Libre de Bruxelles, Brussels, Belgium; and the Department of Pediatrics,[‖] Memorial Sloan Kettering Cancer Center, New York, New York*

Accepted for publication July 28, 2015.

Address correspondence to Daniel E. Sabath, M.D., Ph.D., Department of Laboratory Medicine, University of Washington, Box 359743, Harborview Medical Center, 3NJ-345.1, 325 Ninth Ave., Seattle, WA 98104; or Harvey A. Greisman, M.D., Ph.D., E-mail: dsabath@uw.edu or hagreisman@gmail.com.

Thalassemia is among the most common genetic diseases worldwide. α-Thalassemia is usually caused by deletion of one or more of the duplicated *HBA* genes on chromosome 16. In contrast, most β-thalassemia results from point mutations that decrease or eliminate expression of the *HBB* gene on chromosome 11. Deletions within the *HBB* locus result in thalassemia or hereditary persistence of fetal Hb. Although routine diagnostic testing cannot distinguish thalassemia deletions from point mutations, deletional hereditary persistence of fetal Hb is notable for having an elevated HbF level with a normal mean corpuscular volume. A small number of deletions accounts for most α-thalassemias; in contrast, there are no predominant *HBB* deletions causing β-thalassemia. To facilitate the identification and characterization of deletions of the *HBA* and *HBB* globin loci, we performed array-based comparative genomic hybridization using a custom oligonucleotide microarray. We accurately mapped the breakpoints of known and previously uncharacterized *HBB* deletions defining previously uncharacterized deletion breakpoints by PCR amplification and sequencing. The array also successfully identified the common *HBA* deletions $-^{SEA}$ and $-^{FIL}$. In summary, comparative genomic hybridization can be used to characterize deletions of the *HBA* and *HBB* loci, allowing high-resolution characterization of novel deletions that are not readily detected by PCR-based methods. *(J Mol Diagn 2016, 18: 92—99; http://dx.doi.org/10.1016/j.jmoldx.2015.07.011)*

Thalassemia is one of the most common genetic diseases and is responsible for significant morbidity worldwide. α-Thalassemia is most commonly caused by deletion of one or more of the duplicated *HBA* genes on chromosome 16. In contrast, most β-thalassemia results from point mutations that cripple or completely inactivate expression from the *HBB* gene on chromosome 11. These mutations commonly result in a decrease in HBB and a relative increase in HBD protein expression, resulting in an elevated HbA_2 level. There is frequently a relative increase in HBG1/2 expression as well, with a corresponding increase in HbF. Deletions within the *HBB* locus occur and can eliminate HBB expression. If the *HBD* and *HBG* genes are deleted along with the *HBB* gene, the characteristic increases in HbA_2 and HbF will not be seen, making it impossible to distinguish from α-thalassemia on the basis of HbA_2 quantification alone. Deletions resulting in levels of HBD and HBG polypeptide production insufficient to balance that from the *HBA* locus result in δβ- or γδβ-thalassemia with a decreased mean corpuscular volume (MCV). In contrast, some deletions result in elevated levels of HBG sufficient to balance that from the *HBA* locus, resulting in hereditary persistence of fetal Hb (HPFH), the hallmark of which is a normal MCV. In both of these cases, an *HBB* locus deletion is assumed. *HBB* deletions causing thalassemia may be more common than generally thought because they may not be recognized during routine thalassemia diagnostic procedures, especially in the rare deletions that do not result in HbF elevation.

Supported by NIH grant DK44746 (M.A.B.) and funds from the Department of Laboratory Medicine, University of Washington.

Disclosures: None declared.

Current address of H.A.G.: LabCorp, Seattle. WA.

http://dx.doi.org/10.1016/j.jmoldx.2015.07.011

Table 1 PCR and Sequencing Primers Used in This Study

Subject No.	Forward PCR primer	Reverse PCR primer	Additional sequencing primers, if used
4	5′-TAGATCCCTTTGCCATTATG-3′	5′-TTGGGTTTCTGATAGGCACTG-3′	
5	5′-AGCCTCATGGTAGCAGAATC-3′	5′-TGGTATCTGCAGCAGTTGCC-3′	
7*	5′-CCTTTTTCTTGTGGTAAATG-CTT-3′	5′-TTTCCTTGTGTTTGAAAGTGCT-3′	F: 5′-ACCCTTTGAGTAATAGTTTCCTGA-3′ R: 5′-GCAAATAAGCACACATATATTCCAA-3′
	5′-ATGCCAGTGCTCTCCACAAT-3′	5′-ATAATAAGCCTGCGCCCTTC-3′	R: 5′-AGTTTCTTGTTTACTCTGGA-3′ R: 5′-GGCCATATAGGGTAACTTTCTGAC-3′ R: 5′-AAAAGTGTGCCATGGTTTTAATG-3′
8	5′-CATCCACCACTTTCTGATAGG-3′	5′-TAGCATGCATGAGCAAATTAAGA-3′	
9	5′-GCCACATGGTATGGGAGGTA-3′	5′-TGTTACTTGTCTGGTGTGGCTAA-3′	

*Upper line shows sequences for telomeric breakpoint; lower line shows sequences for centromeric breakpoint.
F, forward; R, reverse.

Several well-characterized deletions account for most α-thalassemia; thus, gap-PCR–based methods are routinely used to identify *HBA* gene deletions.[1] In contrast, *HBB* deletions, when they occur, are diverse, so routine PCR-based methods are not efficient or routinely available. This is especially true in the United States, where no set of mutations predominates because of the ethnic diversity of this population. To provide a rapid inexpensive and efficient platform to identify common, previously described, and unique deletions of either the *HBA* or *HBB* loci, we developed a microarray-based comparative genomic hybridization (CGH) assay. This method can be used to genotype patients with deletional forms of α- or β-thalassemia, HPFH, or complex mixed *HBA* and *HBB* globin deletions, whose genotype is not apparent from protein-based studies, such as electrophoresis and high-performance liquid chromatography (HPLC). We used this array to characterize both previously characterized and novel deletions of the *HBB* locus. This array has been used in patients with apparent thalassemia trait but no evidence of α-thalassemia and no increase in HbA_2, providing diagnostic information that allows accurate genetic counseling.

Materials and Methods

Patient Specimens and Diagnostic Studies

Patient specimens were sent to the University of Washington Red Cell Disorders laboratory for routine clinical testing. Clinical testing at the University of Washington included qualitative Hb analysis by isoelectric focusing (Perkin-Elmer Multiphor II; Perkin-Elmer, Waltham MA) and HPLC (Bio-Rad Variant II; Bio-Rad, Hercules, CA). HbF and HbA_2 were quantified by HPLC. Detection of HbH inclusions was performed by brilliant cresyl blue staining as previously described.[2] Iron deficiency was assessed by determining the zinc protoporphyrin/heme ratio[3]; serum iron, ferritin, total iron binding capacity, and transferrin saturation were determined in a subset of cases. For samples collected at the University of Washington, DNA was isolated using a Corbett automated DNA extractor (Qiagen, Valencia, CA). PCR to detect *HBA* deletions was performed as previously described.[4] Additional specimens were obtained from Boston Children's Hospital. All studies were approved by the University of Washington Human Subjects Division.

Microarray Design

Custom arrays were designed essentially as described, including the software used for probe design and repetitive sequence masking.[5] Genome coordinates used throughout this article are from the human hg38 assembly (December 2013). For chromosome 11, probes were tiled at a median density of 60 bp across an approximately 450-kbp region surrounding the *HBB* locus (coordinates chr11:5,022,271-5,472,133), and at a median density of 57 to 62 kbp on either side of this region for the remainder of chromosome 11 (coordinates chr11:218,364-5,022,271 and chr11:5,472,133-134,998,453). For chromosome 16, probes were tiled at a median density of 64 bp across an approximately 400-kbp region surrounding the *HBA* loci (coordinates chr16:11,177-414,936) and a median density of 56,261 bp for the remainder of chromosome 16 (chr16:429,413-90,081,925). An additional 883 probes representing the other chromosomes (median spacing of 1.8 Mbp) were included for the purpose of signal normalization. The array design is accessible in the National Center for Biotechnology Information Gene Expression Omnibus (*http://www.ncbi.nlm.nih.gov/geo*; accession number GPL20136).

Array Hybridization, Scanning, and Data Analysis

Subject genomic DNA was labeled with Cy5-dUTP, and normal human genomic DNA (Promega, Madison, WI) was labeled with Cy3-dUTP using the Agilent Genomic DNA Labeling kit PLUS (Agilent Technologies, Santa Clara, CA). Array hybridization, scanning, and data analysis were performed as described previously.[5] The raw data are accessible at the National Center for Biotechnology Information Gene Expression Omnibus (*http://www.ncbi.nlm.nih.gov/geo*; accession number GSE68587).