Vona et al. BMC Res Notes (2018) 11:391
https://doi.org/10.1186/s13104-018-3466-7

BMC Research Notes

DATA NOTE

Open Access

Hereditary hearing loss SNP-microarray pilot study

Barbara Vona[1*], Michaela A. H. Hofrichter[1], Jörg Schröder[1], Wafaa Shehata-Dieler[2], Indrajit Nanda[1] and Thomas Haaf[1]

Abstract

Objectives: Despite recent advancements in diagnostic tools, the genomic landscape of hereditary hearing loss remains largely uncharacterized. One strategy to understand genome-wide aberrations includes the analysis of copy number variation that can be mapped using SNP-microarray technology. A growing collection of literature has begun to uncover the importance of copy number variation in hereditary hearing loss. This pilot study underpins a larger effort that involves the stage-wise analysis of hearing loss patients, many of whom have advanced to high-throughput sequencing analysis.

Data description: Our data originate from the Infinium HumanOmni1-Quad v1.0 SNP-microarrays (Illumina) that provide useful markers for genome-wide association studies and copy number variation analysis. This dataset comprises a cohort of 108 individuals (99 with hearing loss, 9 normal hearing family members) for the purpose of understanding the genetic contribution of copy number variations to hereditary hearing loss. These anonymized SNP-microarray data have been uploaded to the NCBI Gene Expression Omnibus and are intended to benefit other investigators interested in aggregating platform-matched array patient datasets or as part of a supporting reference tool for other laboratories to better understand recurring copy number variations in other genetic disorders.

Keywords: Copy number variation, Genotyping arrays, Hereditary hearing loss, Infinium HumanOmni1-Quad, Illumina, SNP-microarray

Objective

Copy number variations (CNVs) are a well-recognized cause of genetic disease through the disruption of gene dosage and/or expression. However, their contribution to hereditary hearing loss (HL) has long been underestimated and remains an important question. Recently, a better appreciation of the CNV-burden in HL patients has emerged, with one study estimating that CNVs are implicated in up to 18.7% of patients in whom a genetic cause of HL was identified [1]. Ongoing efforts in the field are underway to not only diagnose patients, but also to identify genes underlying HL [2–4]. The high frequency of CNVs in Mendelian phenotypes such as HL support a SNP-microarray analysis strategy for the purpose of identifying chromosomal aberrations in known and candidate genes [5].

In this project, we ascertained HL patients in whom a molecular genetic diagnosis could not be determined from exclusionary *GJB2* (DFNB1A) screening to assess the contribution of CNVs to the diagnostic rate of HL. Our analysis established *STRC* (DFNB16) as a frequent cause of congenital HL [6] and identified a rare syndromic form of HL caused by a de novo deletion in the chromosome 4q35.1q35.2 region [7]. The data have also identified patients with inconspicuous SNP-microarray array findings who have advanced to projects that utilize high-throughput sequencing and bioinformatics analysis [8, 9]. The most interesting and impactful results from this work have been published. We have subsequently shifted our research efforts to employ whole exome sequencing in our HL cohort. However, we believe these SNP-microarray data may be of retrospective interest and offer valued information to the scientific community.

*Correspondence: barbara.vona@uni-wuerzburg.de
[1] Institute of Human Genetics, Julius Maximilians University, Würzburg, Germany
Full list of author information is available at the end of the article

© The Author(s) 2018. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Table 1 Overview of data files/data sets

Label	Name of data file/data set	File types (file extension)	Data repository and identifier (DOI or accession number)
Data set 1	Infinium HumanOmni1-Quad v1.0 SNP-microarray data	This dataset contains the raw intensity data files (Grn.idat and Red.idat) of each patient, as well as Matrix Signal Intensities (.txt) and Matrix Processed data (.txt)	NCBI Gene Expression Omnibus Data series accession: GSE111131 Data identifiers: GSM3022603, GSM3022604, GSM3022605, GSM3022606, GSM3022607, GSM3022608, GSM3022609, GSM3022610, GSM3022611, GSM3022612, GSM3022613, GSM3022614, GSM3022615, GSM3022616, GSM3022617, GSM3022618, GSM3022619, GSM3022620, GSM3022621, GSM3022622, GSM3022623, GSM3022624, GSM3022625, GSM3022626, GSM3022627, GSM3022628, GSM3022629, GSM3022630, GSM3022631, GSM3022632, GSM3022633, GSM3022634, GSM3022635, GSM3022636, GSM3022637, GSM3022638, GSM3022639, GSM3022640, GSM3022641, GSM3022642, GSM3022643, GSM3022644, GSM3022645, GSM3022646, GSM3022647, GSM3022648, GSM3022649, GSM3022650, GSM3022651, GSM3022652, GSM3022653, GSM3022654, GSM3022655, GSM3022656, GSM3022657, GSM3022658, GSM3022659, GSM3022660, GSM3022661, GSM3022662, GSM3022663, GSM3022664, GSM3022665, GSM3022666, GSM3022667, GSM3022668, GSM3022669, GSM3022670, GSM3022671, GSM3022672, GSM3022673, GSM3022674, GSM3022675, GSM3022676, GSM3022677, GSM3022678, GSM3022680, GSM3022681, GSM3022682, GSM3022683, GSM3022684, GSM3022685, GSM3022686, GSM3022687, GSM3022688, GSM3022689, GSM3022690, GSM3022691, GSM3022692, GSM3022693, GSM3022694, GSM3022695, GSM3022696, GSM3022697, GSM3022698, GSM3022699, GSM3022700, GSM3022701, GSM3022702, GSM3022703, GSM3022704, GSM3022705, GSM3022706, GSM3022707, GSM3022708, GSM3022709, GSM3022710, GSM3022711 GSE111131_Martix_signal_intensities.txt.gz GSE111131_Matrix_Processed.txt.gz
Data file 1	Data file 1: clinical overview Details: this table provides an overview of the audiological and clinical characteristics of each anonymized patient, as well as patient population background and familial relationships, if present	This file is available as an excel (.xls) table	NCBI Gene Expression Omnibus Data series accession: GSE111131 Data identifier: GSE111131_Data_file_1.xls.gz
Data file 2	Data file 2: sample sheet all Details: this table contains the sample sheet that includes the patient sex and anonymized ID with the beadchip position and barcode information and parental relationships, if present	This file is available as a comma separated variables (.csv) table	NCBI Gene Expression Omnibus Data series accession: GSE111131 Data identifier: GSE111131_Data_file_2.csv.gz

Data description

Patient recruitment

We studied the genomic DNA extracted from whole blood of 99 consecutively recruited patients with suspected hereditary HL and 9 unaffected family members between February 2011 and May 2013. Index patients with suspected environmental forms of HL were excluded. Family members were included, when possible, to enhance data analysis. Prior to investigation in this research setting, the patients had undergone routine diagnostic *GJB2* screening that included Sanger sequencing and duplication/deletion analysis

using a multiplex ligation-dependent probe amplification approach. Patients with homozygous or compound heterozygous pathogenic *GJB2* variants were excluded from the study. In parallel, clinical records were collected and reviewed that are summarized in Data File 1 listed in Table 1. Data File 1 also includes familial relationships, if available.

Experimental protocols

The Illumina Infinium HD assay was performed according to manufacturer's instructions using 200 ng genomic DNA. The Infinium HumanOmni1-Quad v1.0 SNP-microarrays (Illumina) were scanned using the BeadArray Reader and the iScan that are included in the last column of Data File 1 (Table 1).

Data analysis

Unprocessed raw intensity data (.idat files) shown in Data Set 1 of Table 1 were generated. Additionally, raw and normalized green and red intensities (GSE111131_Matrix_signal_intensities.txt.gz), as well as matrix processed data (GSE111131_Matrix_Processed.txt.gz) were assembled. For our study, data were loaded into GenomeStudio v.2011.1 software and the B allele frequency and log R ratio were analyzed using Manifest H, cnvPartition 3.2.0, and QuantiSNP 2.2 [10]. The sample sheet that contains the necessary information to match the patient IDs with the sub-array data for this analysis are included in Data File 2 (Table 1).

Limitations

This study was undertaken to initiate screening of a cohort of diagnostically unresolved HL patients. Of particular interest was obtaining a greater understanding of the contribution of CNVs to hereditary HL, which was underappreciated at the time of study initiation. As this was a pilot study, our intention was to screen a small cohort of 99 patients to gain insight into our primary research aims and then publish the most interesting findings separately [6, 7]. As our work has advanced to include high-throughput sequencing of genes involved in HL, it became evident that many clinically-relevant mutations reside beyond the resolution of the SNP-microarrays [8].

One further limitation relates to the clinical overview of the patients (Data File 1, Table 1). As this study was conducted between 2011 and 2013, any subsequent progression of HL or syndromes that may have manifested in patients after HL was diagnosed and clinical chart review occurred are not included. Thus, these data may not be well-suited for genome-wide association studies, but can nonetheless be included in data collections investigating other disorders with the disclaimer that these disorders, especially adult-onset disorders in patients who were recruited as children, cannot be conclusively excluded.

Technical limitations well-known to SNP-microarrays involve the inability to detect balanced translocations, copy-neutral alterations, and inversions that may nonetheless be relevant [11] for the clinical diagnosis of HL [12, 13].

Abbreviations

CNV: Copy number variation; HL: Hearing loss.

Authors' contributions

BV, IN, and MAHH generated SNP-microarray data. BV, JS, and MAHH collected clinical information. JS and WSD were involved in patient recruitment and clinical information acquisition. BV, IN, and TH wrote the manuscript. TH designed the study. All authors read and approved the final manuscript.

Author details

[1] Institute of Human Genetics, Julius Maximilians University, Würzburg, Germany. [2] Department of ORL, Plastic, Aesthetic and Reconstructive Head and Neck Surgery, Comprehensive Hearing Center, Würzburg, Germany.

Acknowledgements

The authors would like to express their sincere gratitude to the family for their participation in this study. We acknowledge the University of Wuerzburg Open Access Publishing programme for funding this publication.

Competing interests

The authors declare that they have no competing interests.

Availability of data materials

The data described in this Data Note can be freely and openly accessed on NCBI Gene Expression Omnibus under Series GSE111131. Please see table 1 and reference list for details and links to the data.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The study was approved by the Ethics Committee of the University of Würzburg under the reference IDs 205/11 and 46/15. Full written informed parental consent was obtained prior to initiating our investigation.

Funding

T.H. was funded by the German Research Foundation (HA 1374/7-2).

Data citation

Vona B., Haaf T. Hereditary Hearing Loss SNP-Microarray Pilot Study. Gene Expression Omnibus GSE111131 https://identifiers.org/ncbi/GEO:GSE111131 (2018).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in publishedmaps and institutional affiliations.

Received: 8 March 2018 Accepted: 1 June 2018
Published online: 14 June 2018

References

1. Shearer AE, Kolbe DL, Azaiez H, Sloan CM, Frees KL, Weaver AE, Clark ET, Nishimura CJ, Black-Ziegelbein EA, Smith RJ. Copy number variants are a common cause of non-syndromic hearing loss. Genome Med. 2014;6:37.
2. Vona B, Nanda I, Hofrichter MA, Shehata-Dieler W, Haaf T. Non-syndromic hearing loss gene identification: a brief history and glimpse into the future. Mol Cell Probes. 2015;29:260–70.
3. Sloan-Heggen CM, Smith RJ. Navigating genetic diagnostics in patients with hearing loss. Curr Opin Pediatr. 2016;28:705–12.
4. Bademci G, Diaz-Horta O, Guo S, Duman D, Van Booven D, Foster J 2nd, Cengiz FB, Blanton S, Tekin M. Identification of copy number variants through whole-exome sequencing in autosomal recessive nonsyndromic hearing loss. Genet Test Mol Biomarkers. 2014;18:658–61.
5. Mehta D, Noon SE, Schwartz E, Wilkens A, Bedoukian EC, Scarano I, Crenshaw EB 3rd, Krantz ID. Outcomes of evaluation and testing of 660 individuals with hearing loss in a pediatric genetics of hearing loss clinic. Am J Med Genet A. 2016;170:2523–30.
6. Vona B, Hofrichter MA, Neuner C, Schröder J, Gehrig A, Hennermann JB, Kraus F, Shehata-Dieler W, Klopocki E, Nanda I, et al. DFNB16 is a frequent cause of congenital hearing impairment: implementation of STRC mutation analysis in routine diagnostics. Clin Genet. 2015;87:49–55.
7. Vona B, Nanda I, Neuner C, Schröder J, Kalscheuer VM, Shehata-Dieler W, Haaf T. Terminal chromosome 4q deletion syndrome in an infant with hearing impairment and moderate syndromic features: review of literature. BMC Med Genet. 2014;15:72.
8. Vona B, Müller T, Nanda I, Neuner C, Hofrichter MA, Schröder J, Bartsch O, Läßig A, Keilmann A, Schraven S, et al. Targeted next-generation sequencing of deafness genes in hearing-impaired individuals uncovers informative mutations. Genet Med. 2014;16:945–53.
9. Vona B, Nanda I, Neuner C, Müller T, Haaf T. Confirmation of GRHL2 as the gene for the DFNA28 locus. Am J Med Genet A. 2013;161:2060–5.
10. Colella S, Yau C, Taylor JM, Mirza G, Butler H, Clouston P, Bassett AS, Seller A, Holmes CC, Ragoussis J. QuantiSNP: an objective Bayes hidden-Markov model to detect and accurately map copy number variation using SNP genotyping data. Nucleic Acids Res. 2007;35:2013–25.
11. Coughlin CR 2nd, Scharer GH, Shaikh TH. Clinical impact of copy number variation analysis using high-resolution microarray technologies: advantages, limitations and concerns. Genome Med. 2012;4:80.
12. Schneider E, Märker T, Daser A, Frey-Mahn G, Beyer V, Farcas R, Schneider-Rätzke B, Kohlschmidt N, Grossmann B, Bauss K, 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. 2009;18:655–66.
13. Vona B, Neuner C, El Hajj N, Schneider E, Farcas R, Beyer V, Zechner U, Keilmann A, Poot M, Bartsch O, et al. Disruption of the ATE1 and SLC12A1 genes by balanced translocation in a boy with non-syndromic hearing loss. Mol Syndromol. 2014;5:3–10.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions