The Journal of Molecular Diagnostics, Vol. 19, No. 6, November 2017

ELSEVIER

the Journal of Molecular Diagnostics
jmd.amjpathol.org

Clinical Validation of Copy Number Variant Detection from Targeted Next-Generation Sequencing Panels

Jennifer Kerkhof,* Laila C. Schenkel,† Jack Reilly,* Sheri McRobbie,* Erfan Aref-Eshghi,† Alan Stuart,* C. Anthony Rupar,†‡ Paul Adams,§ Robert A. Hegele,¶‖ Hanxin Lin,*† David Rodenhiser,**††‡‡§§ Joan Knoll,*† Peter J. Ainsworth,*† and Bekim Sadikovic*†

From the Molecular Genetics Laboratory and the Biochemical Genetics Laboratory,‡ Molecular Diagnostics Division, London Health Sciences Centre, London, Ontario; the Departments of Pathology and Laboratory Medicine,† Gastroenterology,§ Medicine,¶ Biochemistry,** Paediatrics,†† and Oncology‡‡ and the Robarts Research Institute,‖ Western University, London, Ontario; and the London Regional Cancer Center Program,§§ the Children's Health Research Institute, London, Ontario, Canada*

Accepted for publication July 31, 2017.

Address correspondence to Bekim Sadikovic, Ph.D., DABMGG, FACMG, Department of Pathology and Laboratory Medicine, Victoria Hospital, London Health Sciences Centre, 800 Commissioner's Rd. E., B10-104, London, ON, Canada N6A 5W9. E-mail: bekim.sadikovic@lhsc.on.ca.

Next-generation sequencing (NGS) technology has rapidly replaced Sanger sequencing in the assessment of sequence variations in clinical genetics laboratories. One major limitation of current NGS approaches is the ability to detect copy number variations (CNVs) approximately >50 bp. Because these represent a major mutational burden in many genetic disorders, parallel CNV assessment using alternate supplemental methods, along with the NGS analysis, is normally required, resulting in increased labor, costs, and turnaround times. The objective of this study was to clinically validate a novel CNV detection algorithm using targeted clinical NGS gene panel data. We have applied this approach in a retrospective cohort of 391 samples and a prospective cohort of 2375 samples and found a 100% sensitivity (95% CI, 89%—100%) for 37 unique events and a high degree of specificity to detect CNVs across nine distinct targeted NGS gene panels. This NGS CNV pipeline enables stand-alone first-tier assessment for CNV and sequence variants in a clinical laboratory setting, dispensing with the need for parallel CNV analysis using classic techniques, such as microarray, long-range PCR, or multiplex ligation—dependent probe amplification. This NGS CNV pipeline can also be applied to the assessment of complex genomic regions, including pseudogenic DNA sequences, such as the *PMS2CL* gene, and to mitochondrial genome heteroplasmy detection. *(J Mol Diagn 2017, 19: 905—920; http://dx.doi.org/10.1016/j.jmoldx.2017.07.004)*

Potentially deleterious changes in DNA sequences involve single-nucleotide polymorphisms (SNPs) and/or structural variants of rearrangements that affect >50 bp, including small insertions and deletions, copy number variations (CNVs), and large structural variants. The ability to detect CNVs with a high degree of sensitivity and specificity is fundamental to a comprehensive gene analysis by a modern clinical laboratory. Inherited and somatically acquired CNVs account for a substantial proportion of genetic variation in the human genome and have been associated with a significant number of human disorders.[1–5] By definition, CNV refers to an intermediate scale structural variant, with copy number changes ranging from 1 Kb to 5 Mb of DNA[6]; however, clinically significant structural variants can range from nucleotide-level insertions/deletions to entire chromosomes and are therefore included in our definition of a CNV. Large deletions and duplications that involve dosage-sensitive developmental genes are also known to be related to the presentation of well-characterized microdeletion and microduplication syndromes, such as Charcot-Marie-Tooth and DiGeorge syndromes.[6]

Several approaches have been developed for CNV assessment, including fluorescent *in situ* hybridization,[7] multiplex ligation—dependent probe amplification (MLPA),[8] comparative

Supported by the London Health Sciences Molecular Genetics Laboratory research and development fund.

Disclosures: None declared.

http://dx.doi.org/10.1016/j.jmoldx.2017.07.004

Table 1 Next-Generation Sequencing Target Panels and Respective Genes

Assorted gene panel			
ACADM	*MECP2*	*RET*	*TP53*
GJB2	*MEN1*	*SCN4A*	*TTR*
GJB6	*NOTCH3*	*SPTLC1*	
BRCA*			
BRCA1	*BRCA2*		
Cancer			
APC	*CDH1*	*MSH6*	*RAD51D*
ATM	*CDK4*	*MUTYH*	*SMAD4*
BARD1	*CDKN2A*	*NBN*	*STK11*
BMPR1A	*CHEK2*	*PALB2*	*TP53*
BRCA1	*EPCAM*	*PMS2*	
BRCA2	*MLH1*	*PTEN*	
BRIP1	*MSH2*	*RAD51C*	
Charcot-Marie-Tooth syndrome			
EGR2	*HSPB1*	*MPZ*	*SH3TC2*
FIG4	*HSPB8*	*NEFL*	*TRPV4*
GARS	*LITAF*	*PMP22*	
GDAP1	*LMNA*	*PRX*	
GJB1	*MFN2*	*RAB7A*	
Dyslipidemia			
ABCA1	*CAV1*	*LDLR*	*PAX4*
ABCC8	*CEL*	*LDLRAP1*	*PCSK4*
ABCG5	*CETP*	*LEP*	*PDX1*
ABCG8	*CIDEC*	*LEPR*	*PIK3R1*
ADIPOQ	*DYRK1B*	*LIPA*	*PLIN1*
AGPAT2	*FBN1*	*LIPC*	*POLD1*
AKT2	*FTO*	*LIPE*	*POMC*
ANGPTL3	*GCK*	*LIPG*	*PPARG*
APOA1	*GPD1*	*LMF1*	*PSMB8*
APOA5	*GPIHBP1*	*LMNA*	*PTRF*
APOB	*HNF1A*	*LMNB2*	*RXRG*
APOC2	*HNF1B*	*LPIN1*	*SAR1B*
APOC3	*HNF4A*	*LPL*	*SCARB1*
APOE	*INS*	*MC3R*	*STAP1*
ATF6	*KCNJ11*	*MC4R*	*TBC1D4*
BLK	*KCNJ6*	*MTTP*	*USF1*
BSCL2	*KLF11*	*NEUROD1*	*WRN*
C5AR2	*LCAT*	*OSBPL10*	*ZMPSTE24*
Epilepsy			
ALDH7A1	*GATM*	*NECAP1*	*SCN8A*
AMT	*GLDC*	*NEU1*	*SCN9A*
ARX	*GOSR2*	*NHLRC1*	*SLC2A1*
ASAH1	*GRIN2A*	*NRXN1*	*SLC6A8*
ATP1A2	*GRIN2B*	*PCDH19*	*SLC9A6*
ATP1A3	*HCN1*	*PHGDH*	*SPTAN1*
CDKL5	*KCNC1*	*PLCB1*	*STXBP1*
CERS1	*KCNJ10*	*PNKP*	*SUOX*
CHD2	*KCNJ11*	*PNPO*	*SYNGAP1*
CHRNA7	*KCNQ2*	*POLG*	*TBC1D24*
CNTNAP2	*KCNQ3*	*PRICKLE2*	*TCF4*
CSTB	*KCNT1*	*PRRT2*	*TSC1*
DNM1	*KCTD7*	*PSAT1*	*TSC2*
DOCK7	*LMNB2*	*PSPH*	*UBE3A*
EPM2A	*MBD5*	*SCARB2*	*ZEB2*
FOLR1	*MECP2*	*SCN1A*	
FOXG1	*MEF2C*	*SCN1B*	
GAMT	*MOCS1*	*SCN2A*	
Hyperferritinemia			
ALAS2	*FTH1*	*HFE2*	*STEAP3*
B2M	*FTL*	*SEC23B*	*TF*
CDAN1	*HAMP*	*SLC25A38*	*TFR2*
CP	*HFE*	*SLC40A1*	
Lysosomal storage/urea cycle disorder			
AGA	*CTSK*	*GUSB*	*NEU1*
ARG1	*DNAJC5*	*HEXA*	*NPC1*
ARSA	*FUCA1*	*HEXB*	*NPC2*
ARSB	*GAA*	*HGSNAT*	*OTC*
ASAH1	*GALC*	*HYAL1*	*PPT1*
ASL	*GALNS*	*IDS*	*PSAP*
ASS1	*GBA*	*IDUA*	*SGSH*
CA5A	*GLA*	*LAMP2*	*SLC17A5*
CLN3	*GLB1*	*LIPA*	*SLC24A2*
CLN5	*GLUD1*	*MAN2B1*	*SLC25A13*
CLN6	*GLUL*	*MANBA*	*SLC25A15*
CLN8	*GM2A*	*MCOLN1*	*SLC7A7*
CPS1	*GNPTAB*	*MFSD8*	*SMPD1*
CTNS	*GNPTG*	*NAGA*	*SUMF1*
CTSA	*GNS*	*NAGLU*	*TPP1*
CTSD	*GRN*	*NAGS*	
Mitochondrial DNA			
MT-ATP6	*MT-ND4L*	*MT-TI*	*MT-TS2*
MT-ATP8	*MT-ND5*	*MT-TK*	*MT-TT*
MT-CO1	*MT-ND6*	*MT-TL1*	*MT-TV*
MT-CO2	*MT-TA*	*MT-TL2*	*MT-TW*
MT-CO3	*MT-TC*	*MT-TM*	*MT-TY*
MT-CYB	*MT-TD*	*MT-TN*	*MT-RNR1*
MT-ND1	*MT-TE*	*MT-TP*	*MT-RNR2*
MT-ND2	*MT-TF*	*MT-TQ*	
MT-ND3	*MT-TG*	*MT-TR*	
MT-ND4	*MT-TH*	*MT-TS1*	

*BRCA panel description given by Schenkel et al.[12]

genomic hybridization microarrays, and SNP arrays.[9] Although large chromosomal and segmental rearrangements are detectable by fluorescent *in situ* hybridization, most CNVs identified in the human genome are below the resolution of current fluorescent *in situ* hybridization technology. To date, MLPA and microarray-based technologies have been the most reliable and effective methods for discovering copy number alterations. Although both comparative genomic hybridization and SNP array techniques provide a genome-wide CNV screening capability, and SNP arrays also allow allelotyping, generally small deletions/duplications need confirmation by an alternate method or are not detectable.[10] MLPA is a semiquantitative PCR-based technique that can detect deletions and duplications for up to 50 genetic loci in one assay.[11] Because of its low cost, high sensitivity and specificity, and medium throughput, the MLPA technique has become the gold standard diagnostic tool. However, drawbacks of MLPA include inability to provide information regarding the exact location of a duplicated sequence or its orientation, lack of sensitivity for regions not directly encompassed by the