Human Mutation
OFFICIAL JOURNAL
HGVS
HUMAN GENOME VARIATION SOCIETY
www.hgvs.org

Simultaneous Mutation and Copy Number Variation (CNV) Detection by Multiplex PCR–Based GS-FLX Sequencing

Dirk Goossens,[1,3] Lotte N. Moens,[1,3] Eva Nelis,[2,3] An-Sofie Lenaerts,[1,3] Wim Glassee,[1,3] Andreas Kalbe,[4] Bruno Frey,[4] Guido Kopal,[4] Peter De Jonghe,[2,3] Peter De Rijk,[1,3] and Jurgen Del-Favero[1,3]*

[1]*Applied Molecular Genomics Group, Department of Molecular Genetics, Flanders Interuniversity Institute for Biotechnology (VIB), Belgium*
[2]*Neurogenetics Group, Department of Molecular Genetics, Flanders Interuniversity Institute for Biotechnology (VIB), Belgium*
[3]*University of Antwerp (UA), Antwerpen, Belgium*
[4]*Roche Applied Science, Penzberg, Germany*

Communicated by Mats Nilsson

Received 10 March 2008; accepted revised manuscript 10 July 2008.
Published online 4 December 2008 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/humu.20873

ABSTRACT: We evaluated multiplex PCR amplification as a front-end for high-throughput sequencing, to widen the applicability of massive parallel sequencers for the detailed analysis of complex genomes. Using multiplex PCR reactions, we sequenced the complete coding regions of seven genes implicated in peripheral neuropathies in 40 individuals on a GS-FLX genome sequencer (Roche). The resulting dataset showed highly specific and uniform amplification. Comparison of the GS-FLX sequencing data with the dataset generated by Sanger sequencing confirmed the detection of all variants present and proved the sensitivity of the method for mutation detection. In addition, we showed that we could exploit the multiplexed PCR amplicons to determine individual copy number variation (CNV), increasing the spectrum of detected variations to both genetic and genomic variants. We conclude that our straightforward procedure substantially expands the applicability of the massive parallel sequencers for sequencing projects of a moderate number of amplicons (50–500) with typical applications in resequencing exons in positional or functional candidate regions and molecular genetic diagnostics.
Hum Mutat 30, 472–476, 2009.
© 2008 Wiley-Liss, Inc.

KEY WORDS: multiplex PCR; GS-FLX sequencing; mutation detection; CNV detection; diagnostic sequencing

Introduction

For years, Sanger sequencing of PCR products has been the gold standard for targeted variation detection in general, and in mutation screening in particular. But when the focus of human genetics research moved from monogenic diseases toward more complex, polygenic phenotypes, the number and size of the candidate regions and thereby the number of candidate genes increased dramatically. Likewise, the number of identified causative sequence variants and genes for a given disease is on the rise. Accordingly, both in research and diagnostics, the need to parallelize gene sequencing is growing rapidly. One example of a disease with a large number of identified causative genes is Charcot-Marie-Tooth Disease (CMT), the most common inherited disorder of the peripheral nervous system. To date, about 50 loci and 42 genes have been identified for CMT and related disorders, while the majority of dominant mutations are found in seven genes: *PMP22* (MIM# 601097), *MPZ* (MIM# 159440), *MFN2* (MIM# 608507), *LITAF* (MIM# 603795), *GJB1* (MIM# 304040), *EGR2* (MIM# 129010), and *NEFL* (MIM# 162280) [Timmerman et al., 2006].

With an increasing number of candidate genes, the traditional Sanger PCR sequencing paradigm will find a competitor in the recently developed technologies for massive parallel sequencing [Margulies et al., 2005; Shendure et al., 2005]. Still, these next-generation sequencers were conceived for sequencing of complete genomes and they perform extremely well on bacterial and viral genome sequencing. On the other hand, the operational costs remain too high for routine resequencing of complete human genomes. Therefore, their application to human genetics calls for "front-end" methods to specifically amplify a targeted subset of our complex genome.

In view of the vast amount of sequences that are generated, from 0.1 to 2.0 Gb depending on the platform, the use of standard simplex PCR amplification as front-end is questionable. Indeed, since 0.1 Gb of sequencing equals 40-fold coverage of 2,500 amplicons of 1 kb, a simplex PCR front-end would present a bottleneck due to logistic problems in generating and handling the large number of PCR products. For this reason, a number of alternative front-end methods have been recently published. These methods make use of simultaneous amplification after a targeted circularization [Dahl et al., 2007; Fredriksson et al., 2007; Porreca et al., 2007], enrichment strategies through selective hybridization to oligonucleotide microarrays [Albert et al., 2007; Okou et al.,

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Ir. Jurgen Del-Favero, PhD, Prof. Dr., Applied Molecular Genomics Group, VIB Department of Molecular Genetics, University of Antwerp — Campus CDE, Parking P4, Building V, Room 1.14, Universiteitsplein 1, BE-2610 Antwerpen, Belgium. E-mail: jurgen.delfavero@ua.ac.be

Contract grant sponsors: Industrial Research Fund of the University of Antwerp (IOF) Research Foundation–Flanders (FWO) Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT).

© 2008 WILEY-LISS, INC.

2007], or partially surface-based multiplex PCR amplification [Meuzelaar et al., 2007]. They all consist of multistep protocols with mild to high complexity, as opposed to the simultaneous amplification of numerous targets in a multiplex PCR reaction.

Although it is generally believed that multiplex PCR is limited to the amplification of only a few amplicons, due to spurious interactions of the primers in the reaction mixture leading to PCR failure or PCR artifacts [Edwards and Gibbs, 1994; Markoulatos et al., 2002], several multiplex primer design tools are available that take these interactions into account [Rachlin et al., 2005; Yamada et al., 2006]. Moreover, the successful multiplex amplification and genotyping of over 1,000 randomly selected SNPs with primers designed accordingly has been reported [Wang et al., 2005]. Previously, we showed specific and quantitative multiplex amplification of 21 to 37 amplicons using standard PCR conditions with primers designed by software based on the above principle [Gijselinck et al., epub 2007; Sleegers et al., 2006]. Here we evaluate multiplex PCR as a front-end method for massive parallel sequencing and its application to simultaneous variation and copy number variation (CNV) detection.

Materials and Methods

Multiplex Primer Design

An optimal set of targets was defined to cover all coding exons. For every target, a list of potential primers was generated using Primer3 [Rozen and Skaletsky, 2000] based on the following sequences: NC_000001.9 (MFN2 and MPZ), NC_000008.9 (NEFL), NC_000010.9 (EGR2), NC_000016.8 (LITAF), NC_000017.9 (PMP22), and NC_000023.9 (GJB1). Multiplex primer sets were selected from these lists, taking into account previously reported considerations [Elnifro et al., 2000; Wang et al., 2005]. Most importantly, all primers were checked for primer-dimer interactions, both self-dimers and cross-dimers, to prevent template independent primer extension. For this, fewer than four consecutively complementary bases between the 3′ end of any primer, and anywhere in all the others, was allowed. Furthermore, even low-affinity annealing of primers on any formed PCR amplicon was prevented since it can result in the formation of spurious PCR amplicons. This was implemented by calculating the maximum melting temperature of a primer on any PCR amplicon sequence already present in the multiplex using nearest-neighbor thermodynamic parameters as described [SantaLucia, 1998]. The cutoff for this maximum melting temperature was 35°C below the melting temperature of the primer itself.

Multiplex PCR Reaction

The resulting primers were ordered from Eurogentec (Liège, Belgium) and tested in simplex PCR reactions on 20 ng of genomic DNA using 10 pmol per primer; the other parameters were equal to those of the multiplex PCR. The multiplex PCR reactions were performed on 100 ng genomic DNA in a 25-µl reaction containing 1 × Titanium™ Taq PCR buffer (Clontech, Palo Alto, CA) with a final concentration of 0.25 mM for each dNTP (Invitrogen, Carlsbad, CA) and a total of 0.125 µl of Titanium™ Taq DNA Polymerase (Clontech). Primer concentrations were optimized and varied between 0.05 pmol/µl and 0.2 pmol/µl final concentration. PCR cycle conditions were 5 minutes at 98°C, followed by 30 cycles of 45 sat 95°C, 45 sat 62°C, and 2 minutes at 68°C. After a final extension step of 10 minutes at 72°C, samples were cooled to 8°C.

Pooling of Multiplex PCR Reactions

In a preparatory step, the seven multiplex PCR reactions of two samples were separately purified on a QIAquick PCR Purification column (Qiagen GmbH, Hilden, Germany) and the concentration of each multiplex PCR reaction eluate was measured using a Nanodrop spectrophotometer (NanoDrop Technologies, Wilmington, DE). A pooling scheme was defined based on the average concentration over both samples of each multiplex PCR reaction and the number of amplicons per multiplex to obtain an equal representation of every amplicon in the final mixture.

For all remaining individuals, the multiplex PCR reactions were pooled according to the pooling scheme and subsequently purified on a QIAquick PCR purification column (Qiagen).

GS-FLX Sequencing

The pooled and purified multiplex PCR reactions were interrogated with the standard amplicon sequencing protocol on a GS-FLX genome sequencing system (Roche Applied Science, Penzberg, Germany) according to the manufacturer's instructions. For every sample, two lanes of a 16-lane Bead Loading gasket on a 70 × 75 mm PicoTiterPlate (Roche Applied Science, Penzberg, Germany) were loaded, one for sequencing with the forward primer and one for reverse sequencing. The generated standard flow files (SFF) were analyzed with the GS Amplicon Variant Analyzer software (version 1.1.02; Roche Applied Science) and the putative variants that were identified by the software were curated manually.

Sanger Sequencing

Sanger sequencing was performed on 10 ng of genomic DNA with 10 pmol of each primer (primer sequences available on request) in a standard PCR reaction, followed by ExoSAPit treatment (Amersham Biosciences, Uppsala, Sweden) and subsequently sequenced using the Big Dye terminator cycle sequencing kit (v3.1; (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. Sequencing reactions were run on an Applied Biosystems 3730XL DNA Analyzer (Applied Biosystems) and the resulting trace files were analyzed using novoSNP [Weckx et al., 2005].

Dosage Quotient Calculation

To determine the dosage quotient (DQ) of an amplicon in a given individual, the total absolute read count per specific multiplex was determined as the sum of the read counts for all amplicons of that specific multiplex, except the ones of GJB1 because of their location on the X chromosome and the resulting difference between males and females. Per individual, the relative read count was determined for every amplicon as the ratio of the read count for that amplicon over the total absolute read count of the specific multiplex to which the amplicon belongs. The ratio of the relative read count of an amplicon of a test individual over the average relative read counts in the reference individuals results in the DQ for that amplicon in that individual. In short: $DQ^{i}_{test} = (RC^{i}/RC^{all})_{test}/(RC^{i}/RC^{all})_{ref}$ with "RC" = read count; "i" = amplicon i; "all" = all amplicons; "test" = test individual; and "ref" = reference individual.

Results

Multiplex PCR

We designed multiplex PCR primers to amplify all 37 coding exons of seven selected genes (*PMP22*, *MPZ*, *MFN2*, *LITAF*, *GJB1*,

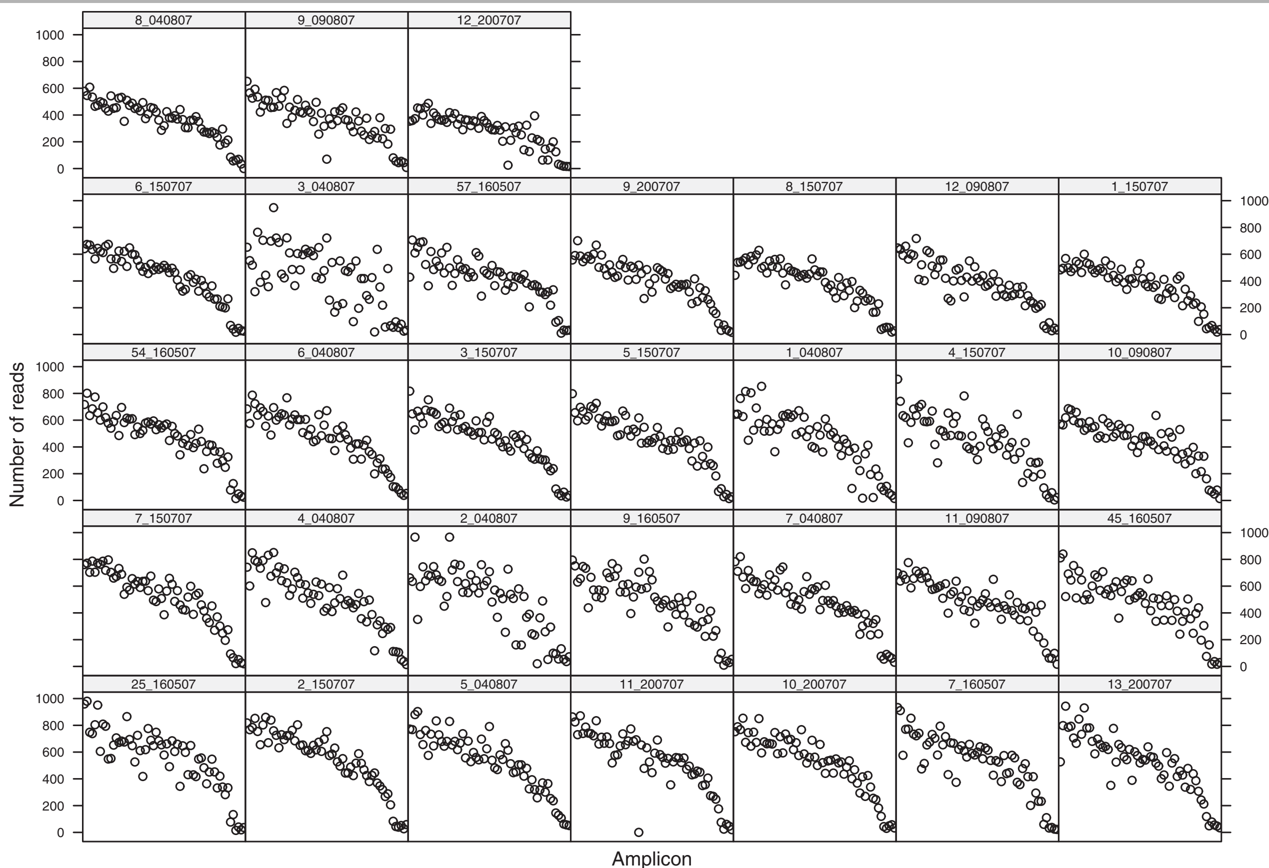

Figure 1. Trellis graph showing read counts per amplicon. Each subgraph plots the number of reads per amplicon for one sample. In all subgraphs, amplicons are sorted according to the average read count calculated over all individuals.

EGR2, and *NEFL*) in which mutations can cause CMT. The ~8 kb of coding sequence was covered by 63 amplicons, totaling 14.4 kb of sequence, amplified in seven multiplex PCR reactions (Supplementary Table S1; available online at http://www.interscience.wiley.com/jpages/1059-7794/suppmat). All primer pairs encompassed fragments ranging from 150 to 250 bp and had the A and B adaptor sequences needed for the GS-FLX amplicon sequencing protocol attached to the forward and reverse primer, respectively [Margulies et al., 2005]. When tested in a simplex PCR reaction, all except two of the primer pairs amplified the expected fragment, resulting in a conversion rate >96.5%. The two failed amplicons were omitted from further experiments.

GS-FLX Sequencing

To test the designed assays, we PCR-amplified the DNA from 31 CMT patients with known mutations and interrogated the products with the amplicon sequencing protocol on a GS-FLX sequencer. Amplicons were sequenced from both directions, each in one lane of a 16-lane gasket. On average, a total of 6.6 Mb was generated per sample, corresponding to over 29,000 reads with an average length of 227 bp (Supplementary Table S2). The specificity of the PCR amplification was very high, with on average 96% of quality filtered reads that mapped back to the expected sequence. Also, it can be appreciated from Figure 1 that our method results in reproducible read counts. The overall observed read count of >90% of the amplicons falls within a 15-fold range. However, when we exclude the six amplicons with the lowest average read count, the observed read counts of >90% of the remaining amplicons fall within a five-fold range in all samples, demonstrating uniformity of amplification.

Mutation detection

To determine the performance of the combined multiplex PCR/GS FLX sequencing method in detecting variations, we compared it to traditional PCR-based Sanger sequencing. Sanger-based resequencing of the coding sequences of the seven genes in all 31 selected CMT patients resulted in the detection of a total of 47 variants located in the amplicons used for GS-FLX sequencing, 31 of which were previously diagnosed as the causative mutations (Supplementary Table S3). Analysis of the generated GS-FLX dataset with the GS Amplicon variant analyzer software (version 1.1.02) confirmed the presence of all 47 variants and did not result in the detection of any additional variant. In view of the supposed difficulties of the 454 sequencing technology with homopolymers, it is important to note that one mutation was a A/C substitution flanking a [C]5-stretch, while another mutation was an insertion of a G in a [G]6-stretch (data not shown).

Analysis of all heterozygous variants showed relatively unbiased representation of both alleles with 44.9±6.5% of the number of reads containing the variant alleles, again confirming the uniformity of amplification.

Since the average multiplicity of coverage for this proof-of-concept experiment was 475×, we performed simulation studies to determine the minimal multiplicity of coverage required for detecting all sequence variants. Hereto, we reduced the actual multiplicity of coverage by subsampling a fixed number of

random reads for each individual and verified whether the variants were still detected. For every multiplicity of coverage sampled, 100 random subsets were generated and analyzed. This resulted in a simulated detection sensitivity of 99.5% for a 40 × coverage, excluding the variants in the six amplicons with low average read count (Supplementary Fig. S1A).

CNV detection

In addition to specific and accurate variant detection, for diagnostic purposes it is also important that simultaneous CNV analysis can be performed within the same multiplex PCR/GS-FLX experiment. As demonstrated by methods such as multiplex amplicon quantification (MAQ) [Sleegers et al., 2006] and quantitative multiplex PCR of short fluorescent fragments (QMPSF) [Casilli et al., 2002], multiplex PCR reactions can be used to determine the copy number of the amplified targets by calculating their DQ. This DQ results from the comparison of normalized amplicon amounts between a test individual and control individuals as detailed in Materials and Methods. To identify CNVs with the multiplex PCR/GS-FLX sequencing method, we used GS-FLX read counts per amplicon as readout. Hereto, the CMT PCR multiplex assay was used to interrogate the DNA of nine individuals (two males, seven females) harboring one (n = 3), two (n = 2), or three (n = 4) copies of the 1.4-Mb CMT1A region at 17p12, encompassing the *PMP22* gene. Duplication of this region is the most common cause of CMT1A (MIM# 118220), while the reciprocal deletion results in hereditary neuropathy with liability to pressure palsies (HNPP; MIM# 162500).

A prerequisite for CNV detection using the multiplex PCR/GS-FLX method is that the normalized read counts are stable over different samples. Indeed, the normalized read count data from amplicons not derived from the CMT1A CNV or chromosome X are stable (Supplementary Table S4). Based on this observation we calculated the average DQs of the six *GJB1* amplicons located on chromosome X using female DNA samples as a reference. This analysis showed a clear-cut distinction between males with a DQ of 0.5 (range, 0.45–0.49) and females with a DQ of 1 (range, 0.86–1.06) corresponding to one and two X-chromosomes, respectively (Fig. 2; Supplementary Fig. S2). Next, calculation of the average DQs of three *PMP22* amplicons using the two wild-type (WT) individuals as a reference, showed three separated ranges at 0.5 (range, 0.47–0.58), 1.0 (range, 0.89–1.11), and 1.5 (range, 1.34–1.49), corresponding to the number of copies present in the individuals (Fig. 2; Supplementary Fig. S3). Subsampling experiments simulating different coverages indicated that for GJB1 50 × coverage is required to reach a sensitivity of >99.5% (Supplementary Fig. S1B). With this coverage we obtained a detection sensitivity of 92.4% for PMP22, probably due to the low number of PMP22 amplicons (Supplementary Fig. S1B). This experiment proves the feasibility of CNV analysis with the multiplex PCR/GS-FLX sequencing method, a result further supported by the observation that the average DQs in five autosomal genes with normal copy number do not deviate from 1.0 (Fig. 2; Supplementary Figs. S4–S8).

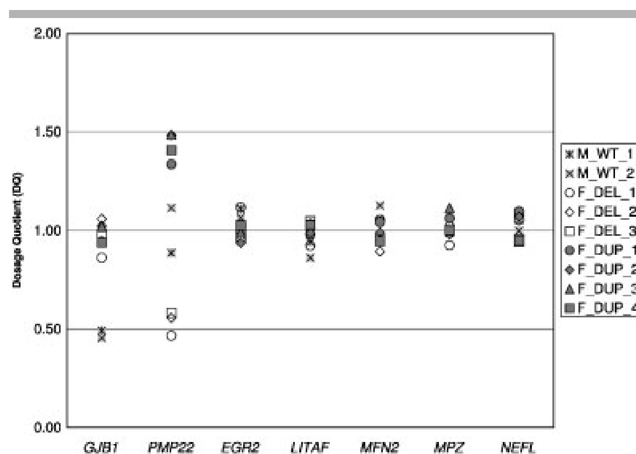

Figure 2. The overall dosage plot. Every dot is the average of the DQs of the amplicons of one gene in one individual. For the calculation of the DQs in *GJB1*, the seven female samples were used as reference; for *PMP22*, the two individuals with two copies of the CMT1A CNV were used as reference; and for the other genes, all samples were used as reference. Crosses indicate males with one copy of the CMT1A CNV, open symbols indicate females with one copy of the CMT1A CNV, and filled symbols indicate females with three copies of the CMT1A CNV.

Discussion

We evaluated multiplex PCR as a front-end to massive parallel sequencing methods. When assessing multiplex targeting methods, important aspects are the specificity, uniformity, and the degree of multiplexing [Porreca et al., 2007]. The specificity is defined as the fraction of amplified sequences that originate from the targets. Uniformity is a measure for the relative abundances of targeted amplification products. The degree of multiplexing refers to the number of targets amplified in one reaction.

With over 95% of quality filtered reads originating from the targets, the specificity of the method is clearly satisfactory and not noticeably different from results obtained with standard, simplex amplicon sequencing experiments and clearly indicates the absence of spurious byproducts of the multiplex PCR amplification step.

With on average nine amplicons per multiplex PCR reaction, the number of PCR reactions required to amplify the target sequences is greatly reduced. Still, there is room for improvement of the multiplexing degree. Since the A and B adaptor sequences are joined to the PCR primers, the degrees of freedom in the primer design process are diminished considerably. From this point of view, it would be interesting to investigate the effects of adaptor ligation after PCR amplification as opposed to the use of primer-adaptor fusion primers.

In the current multiplex PCR assay, six amplicons showed a consistent lower read count in all individuals, influencing two important parameters: the final read count uniformity and minimal required coverage for mutation detection. Since these lower read counts can be improved by simply increasing the primer concentrations for these amplicons in the multiplex PCR reactions, the impact of excluding these amplicons on these parameters was investigated. As shown by our analyses, the final read count uniformity improved substantially from a 15-fold to a five-fold range. Also the minimal required coverage for mutation detection dropped from 120 × to 40 × to obtain a mutation detection sensitivity of 99.5%.

Recently, other front-end approaches for massive parallel sequencing were published and two of these use high-density microarrays to enrich human genome–specific sequences for high-throughput sequencing of genes or selected genomic regions [Albert et al., 2007; Okou et al., 2007]. The number of targeted segments obtained is currently 100-fold higher compared to our multiplex PCR front-end amplification step. However, the sample throughput of the microarray-based enrichment is an order of magnitude lower. Also the current microarray enrichment technique shows low segment specificity ranging from 35 to 76% [Albert et al., 2007]. Two other recent publications [Dahl

et al., 2007; Meuzelaar et al., 2007] use highly multiplexed amplification methods for the isolation of specific DNA sequences as a front-end for massive parallel sequencing applications. The published data show that the coverage depth varies significantly between different amplicons and even between alleles of an amplicon (nonuniform amplification). The lack of amplicon uniformity observed with the current protocols of both methods will hamper cost-effective massive parallel sequencing. An approach combining a microarray-based enrichment method with targeted circularization was also published recently [Porreca et al., 2007]. This labor-intensive method shows a major problem with nonuniform capturing of sequences with both a low (480 exons) and high (55,000 exons) number of targets. Also, a high dropout rate (80%) was observed with 55,000 targeted sequences, rendering the current process only suitable for specific amplification of a moderate number of sequences.

With the sequencing of seven genes in 31 individuals, this is the biggest dataset generated by a massive parallel sequencer for variation detection in complex genomes. The independent detection of all 47 variants that were discovered by Sanger sequencing confirms the sensitivity, while simulations studies indicate that an average coverage of $50\times$ is sufficient to perform variant detection with a sensitivity of 99.8%. This opens perspectives to increase the cost-efficiency of the method. Indeed, as mentioned earlier, the uniformity of amplification could be improved by adapting the primer concentrations. More uniform amplification would then allow decreasing the average sequencing depth to $100\times$ or even $50\times$, as opposed to $475\times$ in the current setup. This increase of throughput can be achieved by expanding the target regions in number or size or by using barcoding sequence tags to differentiate the reads of pooled samples [Meyer et al., 2007].

Since CNVs are important determinants in the etiology of (complex) diseases, we designed a proof-of-concept experiment to evaluate the feasibility of CNV detection using a combination of multiplex PCR and GS-FLX sequencing. With the results of our high-coverage dataset, we could unambiguously determine the copy number of both the GJB1 and PMP22 genes. Furthermore, we showed that with a simulated $50\times$ coverage we had a sensitivity of $>99.5\%$ in detecting the copy number of *GJB1* (see Supplementary Fig. S1B). However, with the same coverage we only achieved a detection sensitivity of 92.4% analyzing the copy number of *PMP22*. The main difference between the analyses of both genes is the number of amplicons used to calculate the DQ. This indicates that the assays should contain a sufficient amount of amplicons with a stable normalized read count per CNV to achieve a sensitivity of $>99.5\%$.

With this proof-of-concept experiment we show the quantitative nature of the multiplex amplification and massive parallel sequencing method and we present the first method allowing simultaneous detection of both SNPs and CNVs in a defined set of target genes. This, in combination with the highly-specific, uniform, and reproducible manner of amplification, leads us to believe that the presented multiplex PCR front-end is the method of choice for parallelizable sequencing projects of a moderate number of amplicons (50–500) with typical applications in resequencing exons in positional candidate regions and molecular genetic diagnostics.

Acknowledgments

We thank the VIB Genetic Service Facility (www.vibgeneticservicefacility.be) for the Sanger sequencing analyses. The research described in this paper was supported by the Industrial Research Fund of the University of Antwerp (IOF), the Research Foundation–Flanders (FWO), and the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT).

References

Albert TJ, Molla MN, Muzny DM, Nazareth L, Wheeler D, Song X, Richmond TA, Middle CM, Rodesch MJ, Packard CJ, Weinstock GM, Gibbs RA. 2007. Direct selection of human genomic loci by microarray hybridization. Nat Methods 4:903–905.

Casilli F, Di Rocco ZC, Gad S, Tournier I, Stoppa-Lyonnet D, Frebourg T, Tosi M. 2002. Rapid detection of novel BRCA1 rearrangements in high-risk breast-ovarian cancer families using multiplex PCR of short fluorescent fragments. Hum Mutat 20:218–226.

Dahl F, Stenberg J, Fredriksson S, Welch K, Zhang M, Nilsson M, Bicknell D, Bodmer WF, Davis RW, Ji H. 2007. Multigene amplification and massively parallel sequencing for cancer mutation discovery. Proc Natl Acad Sci USA 104:9387–9392.

Edwards MC, Gibbs RA. 1994. Multiplex PCR: advantages, development, and applications. PCR Methods Appl 3:S65–S75.

Elnifro EM, Ashshi AM, Cooper RJ, Klapper PE. 2000. Multiplex PCR: optimization and application in diagnostic virology. Clin Microbiol Rev 13:559–570.

Fredriksson S, Baner J, Dahl F, Chu A, Ji H, Welch K, Davis RW. 2007. Multiplex amplification of all coding sequences within 10 cancer genes by Gene-Collector. Nucleic Acids Res 35:e47.

Gijselinck I, Sleegers K, Engelborghs S, Robberecht W, Martin JJ, Vandenberghe R, Sciot R, Dermaut B, Goossens D, van der ZJ, De Pooter T, Del Favero J, Santens P, De Jonghe P, De Deyn PP, Van Broeckhoven C, Cruts M. Neuronal inclusion protein TDP-43 has no primary genetic role in FTD and ALS. Neurobiol Aging epub 10 December 2007.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380.

Markoulatos P, Siafakas N, Moncany M. 2002. Multiplex polymerase chain reaction: a practical approach. J Clin Lab Anal 16:47–51.

Meuzelaar LS, Lancaster O, Pasche JP, Kopal G, Brookes AJ. 2007. MegaPlex PCR: a strategy for multiplex amplification. Nat Methods 4:835–837.

Meyer M, Stenzel U, Myles S, Prufer K, Hofreiter M. 2007. Targeted high-throughput sequencing of tagged nucleic acid samples. Nucleic Acids Res 35:e97.

Okou DT, Steinberg KM, Middle C, Cutler DJ, Albert TJ, Zwick ME. 2007. Microarray-based genomic selection for high-throughput resequencing. Nat Methods 4:907–909.

Porreca GJ, Zhang K, Li JB, Xie B, Austin D, Vassallo SL, LeProust EM, Peck BJ, Emig CJ, Dahl F, Gao Y, Church GM, Shendure J. 2007. Multiplex amplification of large sets of human exons. Nat Methods 4:931–936.

Rachlin J, Ding C, Cantor C, Kasif S. 2005. MuPlex: multi-objective multiplex PCR assay design. Nucleic Acids Res 33:W544–W547.

Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, editors. Bioinformatics methods and protocols. Totowa, NJ: Humana Press. p 365–386.

SantaLucia JJr. 1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc Natl Acad Sci USA 95:1460–1465.

Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Wang MD, Zhang K, Mitra RD, Church GM. 2005. Accurate multiplex polony sequencing of an evolved bacterial genome. Science 309:1728–1732.

Sleegers K, Brouwers N, Gijselinck I, Theuns J, Goossens D, Wauters J, Del Favero J, Cruts M, van Duijn CM, Van Broeckhoven C. 2006. APP duplication is sufficient to cause early onset Alzheimer's dementia with cerebral amyloid angiopathy. Brain 129:2977–2983.

Timmerman V, Lupski JR, De Jonghe P. 2006. Molecular genetics, biology, and therapy for inherited peripheral neuropathies. Neuromolecular Med 8:1–2.

Wang HY, Luo M, Tereshchenko IV, Frikker DM, Cui X, Li JY, Hu G, Chu Y, Azaro MA, Lin Y, Shen L, Yang Q, Kambouris ME, Gao R, Shih W, Li H. 2005. A genotyping system capable of simultaneously analyzing >1000 single nucleotide polymorphisms in a haploid genome. Genome Res 15:276–283.

Weckx S, Del Favero J, Rademakers R, Claes L, Cruts M, De Jonghe P, Van Broeckhoven C, De Rijk P. 2005. novoSNP, a novel computational tool for sequence variation discovery. Genome Res 15:436–442.

Yamada T, Soma H, Morishita S. 2006. PrimerStation: a highly specific multiplex genomic PCR primer design server for the human genome. Nucleic Acids Res 34:W665–W669.