A Multiplex Pharmacogenetics Assay using the MinION Nanopore Sequencing Device

Yusmiati Liau¹, Simone L. Cree¹, Simran Maggo¹, Allison L. Miller¹, John F. Pearson¹, Patrick A. Gladding², Martin A. Kennedy^{1*}

¹ Department of Pathology and Biomedical Science, University of Otago, Christchurch, New Zealand, ² Theranostics Laboratory, North Shore Hospital, Shakespeare Rd, Auckland, New Zealand

*Corresponding author:

Martin A. Kennedy

Email: martin.kennedy@otago.ac.nz

Abstract

Aim: The MinION nanopore sequencing device opens the opportunity to cost-effective and point-of-care DNA sequencing. We developed a multiplex assay targeting pharmacogenetic variants related to clopidogrel and warfarin, two commonly used drugs that show response variability due to genetic polymorphisms. **Materials & Methods:** Six reference and 78 clinical DNA samples were amplified by PCR to generate 15 amplicons targeting key variants. These products were then barcoded to enable sample multiplexing. Three variant calling tools were used to compare genotyping accuracy. **Results and Conclusions:** All but three samples were successfully sequenced and genotyped. Nanopolish software achieved accuracy > 90 % for all except one variant. While minor mis-genotyping issues exist, this work demonstrates that drug-specific or broad pharmacogenetic screening assays are possible on the MinION sequencing device.

Keywords: MinION, Nanopore sequencing, pharmacogenetics panel, clopidogrel, warfarin

Introduction

Clopidogrel and warfarin are two commonly used drugs for anticoagulation, with high variability in inter-individual responses. For both drugs, much of the variability is attributed to genetic factors, with pharmacokinetic variation in patients taking clopidogrel being associated with variation in the *CYP2C19* gene [1] and warfarin with *CYP2C9* and *VKORC1* [2]. Both *CYP2C19* and *CYP2C9* are polymorphic genes with more than 30 known alleles [1]. Moreover, recent publications suggest that other genes are also associated with variable drug response, including *B4GALT2* [3] and *ABCB1* [4, 5] for clopidogrel and *GGCX* [6] and *CYP4F2* [6] for warfarin.

Various assays have been developed to screen clopidogrel and warfarin related variants including point of care tests [7, 8], that usually detect a subset of common variants. Extensive panels of pharmacogenetic variants are available through technologies such as mass spectrometry or next generation sequencing. These approaches, however, involve high capital cost and are not always accessible by small-scale laboratories.

MinION (Oxford Nanopore Technologies (ONT), Oxford, UK) is a pocket-sized, portable sequencing device, using the disruptive nanopore technology [9]. It contains an array of nano-scale protein pores embedded in a membrane, across which a current is passed, and single stranded DNA or RNA can be streamed through the pores and sequenced in real-time [10, 11]. The device offers the advantage of portability and almost-zero capital cost, opening the opportunity for point-of-care sequencing, particularly given recent release of smaller-scale, cheaper flow cells (“flongles”, ONT). The main factor limiting its wide utility is the relatively high error rate compared to other sequencing technologies. However, the technology is evolving rapidly, with improvements in the device, chemistries and bioinformatic tools, which are contributing to improvements in accuracy [12].

As proof-of-principle for a potential point of care test, we developed a multiplex assay on this technology, focusing on pharmacogenetic variants related to variability in clopidogrel and warfarin response. The assay contained genetic variants reported as clinically relevant by the Clinical Pharmacogenomics Implementation Consortium (CPIC) guidelines [13], as well as covering new variants reported in recent publications [3-6, 15] (Table 1). Use of a polymerase chain reaction (PCR) barcoding approach enabled us to multiplex 84 samples in one sequencing run.

Materials and Methods

Samples

Five reference DNA samples (Coriell Institute, Camden, NJ, USA - see Supplementary Table 1) and one in-house DNA sample with known genotypes were used as controls. Seventy eight samples from a cohort of patients who had undergone percutaneous coronary intervention (PCI) [16], with available *CYP2C19**2, *CYP2C19**3, and *CYP2C19**17 genotype data, were included to validate the assay.

Multiplex PCR

A 15-plex PCR, covering a total of 27 variants, was designed using MFE-primer3.0 (<https://design.igenetech.com/>) (Table 1). Amplicon lengths varied between 163 – 240 bps. The forward and reverse primers were prepared separately to make 10 μ M of each pool. The multiplex PCR reaction consisted of 1 μ M each of the forward and reverse primer pools with 1x PCR reaction buffer, 2.5 mM of $MgCl_2$, 0.4 mM of dNTPs, 1 M of betaine and 2.5 units of Taq DNA polymerase in a 50 μ L reaction. The PCR was run for 35 cycles of 94 °C for 15 seconds, 59 °C for 15 seconds and 72 °C for 30 seconds followed by a final 72 °C for 1 minute. PCR product from each sample was purified using HighPrep™ PCR magnetic beads (Magbio, Gaithersburg, USA) and quantified using a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific).

Insert Table 1 here

Barcoding PCR

One nM of the purified product of each sample was assigned to a specific barcode, using the PCR Barcoding Expansion 1-96 kit (EXP-PBC096) (ONT) by a second-round PCR, according to the 1D PCR Barcoding (96) amplicon (SQK-LSK108) protocol (ONT). Reactions consisted of 1x PCR buffer, 1.5 mM of $MgCl_2$, 0.2 mM of dNTPs, 0.2 μ M of each barcode, and 5 units of Taq DNA polymerase in 100 μ L PCR reaction. The PCR consisted of 20 cycles of 94 °C for 15 seconds, 62 °C for 15 seconds, and 72 °C for 30 seconds with final extension at 72 °C for 1 minute. Taq Ti DNA polymerase (Fisher Biotec, Wembley, Australia) was used for both the initial multiplex PCR and the barcoding PCR.

MinION library preparation

The barcoded samples were purified using HighPrep™ PCR magnetic beads (Magbio, Gaithersburg, USA). Each barcoded sample was quantified with Qubit, normalized in concentration and pooled in equimolar amounts. Eighty ng of the pooled DNA was used in the MinION library preparation, according to the 1D PCR Barcoding (96) amplicons (SQK-LSK108) protocol (ONT).

Data analysis

The MinION sequencing was run for 48 hours as per protocol using an R9 flowcell. The fast5 files generated were basecalled and demultiplexed using Albacore version v2.3.0 (ONT), resulting in a fastq file for each barcode. The fastq files were aligned to a custom reference sequence which was built from human hg19 reference sequences of each amplicon in a fasta file (Supplementary file) using BWA MEM. Three different variant calling tools, marginCaller from marginAlign [17], nanopolish [18], and VarScan 2 [19] were used to call variants. NanoOK [20], Minion-QC [21] and Qualimap [22] were used to generate quality control parameters including read errors, read depth, and percentage of reads alignment. Integrative Genomics Viewer (IGV, Broad Institute, Boston, MA)[23] was used to visualise alignment files. Read depth plot was generated using ggplot2 in R (Vienna, Austria).

Sanger sequencing

A nested PCR using 0.5 μ M of sequence-specific primers was performed in a 20 μ L reaction containing 1x PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.5 units of Taq polymerase and 1 μ L of 10x diluted multiplex PCR product. PCR consisted of 20 cycles of 94 °C for 15 seconds, 59 °C for 15 seconds, and 72 °C for 30 seconds with final extension at 72 °C for 1 minute. Sanger sequencing was performed using 1 μ L product of nested PCR, 1 μ L of primer, 1x sequencing buffer, and 0.5 μ L of Big Dye Terminator in 10 μ L reaction. Dye incorporation was performed in a thermal cycler with 25 cycles of 96 °C for 10 seconds, 50 °C for 5 seconds, and 60 °C for 4 minutes, and subsequently purified using Sephadex G-50 before being sequenced using the AB3130xL genetic analyzer (Thermo Fisher Scientific).

Results

Read characteristics

One sequencing run containing 84 barcoded samples produced 0.9 million reads, of which one third passed the Albacore default quality threshold qscore of seven. These reads were then used for further downstream analysis. The majority of the passed reads (94.6 %) mapped to the reference sequence. The average error rate was 2.94%, 3.83%, and 3.82% for insertion, deletion, and substitution errors, respectively.

The average read depth of each amplicon varied significantly between 43 and 985x ($F(14,1200)=146.6, p<0.001$). Variability was also significantly different among amplicons (Levene's test $P < 0.001$) (Figure 1). Three amplicons, ABCB1, CYP2C96 and VKORC1rs7294 showed very high

read depth, while four amplicons, CYP2C192, CYP2C1917, CYP2C194, and CYP2C93 showed relatively low read depth of < 100 across all samples. Three out of the 78 clinical samples were excluded from analysis due to low yield in initial PCR and low read depth across all amplicons, with less than 50 % of reads being mapped to the reference sequence. Twelve samples for B4GALT2, two samples for CYP2C194, one sample for CYP2C93, and four samples for CYP4F2 amplicons had very low read depth (<10), and were also excluded from analysis (Supplementary Table 1).

Insert Figure 1 here

Genotyping accuracy across different variant calling tools

CYP2C19*2, CYP2C19*3, and CYP2C19*17

For three genotypes (*CYP2C19*2*, *CYP2C19*3*, and *CYP2C19*17*), data from two orthogonal assay platforms (Nanosphere Verigene® or Sequenom MassARRAY platform) were available [16], and initial comparative analysis was done on these three genotypes. Accuracy for *CYP2C19*3* genotyping was 100 % across all variant calling tools (Figure 2), with the three heterozygote samples for *CYP2C19*3* being accurately detected using the MinION assay.

Insert figure 2

Accuracy of *CYP2C19*2* and *CYP2C19*17* calls with the MinION varied depending on the bioinformatic tools used to base-call and generate the variant genotype calls. Nanopolish was the most accurate variant calling tool for this assay, with only four discordant genotypes for *CYP2C19*2* and two for *CYP2C19*17* across a total of 81 samples. All four discordant samples for *CYP2C19*2* and one discordant sample for *CYP2C19*17* were of low read depth (≤ 20). The other discordant sample for *CYP2C19*17* resulted from inaccurate calling of a homozygote variant as a heterozygote. The fraction of total reads that support the alternate allele for this sample was close to 1 (0.78), indicating a homozygote variant. Moreover, the read depth was > 100, however nanopolish incorrectly called this sample as heterozygote.

Other variants

Analysis of the other variants included in this assay was first done on control samples. Nanopolish resulted in 100% accuracy on control samples while marginCaller missed the CYP2C9*6 variant in one sample and VarScan 2 missed a total of 11 variants in six control samples (Supplementary Table 1). As nanopolish gave the best accuracy compared to the other two variant calling tools, it was used to analyze other variants across all samples. Many of the variants included in our custom designed assay are rare, and there were only eight variants which were detected in at least three of the 81

samples. We performed Sanger sequencing for the eight variants for randomly selected subsets of the samples. Accuracy of genotype data from MinION sequencing compared with Sanger sequencing is shown in Table 2.

Insert Table 2 here

Two heterozygote samples for *CYP2C9**3 were called as wild-type by nanopolish. These two samples were of relatively low read depth (< 30). The discordant samples for *B4GALT2*, *VKORC1* rs9923231, and *VKORC1* rs7294 variants consisted of homozygote samples which were mis-called as heterozygotes. This inaccuracy could not be explained by read depth, as some of these samples had a read depth > 100. Genotype data for all samples is available in Supplementary Table 1.

The discordant results for *VKORC1* rs9923231 could be resolved by analyzing the BAM files. As seen in Figure 3, mis-called homozygote samples (lines 1 and 2) had distinct alternate allele frequency compared to real heterozygote samples (lines 3 and 4) and could be easily differentiated.

Insert figure 3 here

***CYP2C9**6, a single base deletion**

All the variants included in the assay are single base substitutions, except for *CYP2C9**6 which is a single base deletion (delA). This variant is very rare in populations other than African, where the allele frequency is approximately 1% [24]. We included a reference sample containing a heterozygote *CYP2C9**6 variant. Nanopolish and VarScan 2 successfully detected this variant, while marginCaller failed. This showed that when using an appropriate variant calling tool, a single base deletion can be accurately detected in MinION sequencing data.

Discussion

Drug responses are often affected by multiple variants in multiple genes, and detection of such variants using conventional approaches like Sanger sequencing or single PCR is both laborious and time-consuming. Extensive genotyping panels containing high number of variants can be analysed in advanced instruments like next generation sequencers, or multiplex PCR-based technologies with bead-based or mass spectrometry detection. However, these approaches require high-capital instrument costs which are not always accessible by diagnostic laboratories. Moreover, modification of such genotyping panels might be difficult and expensive, as special primers or probes are needed. The MinION on the other hand is a portable sequencing tool, available to any laboratory scale and has been use in field settings as far as outer space [18, 25, 26].

We have described an assay detecting multiple genetic variants known to affect response to clopidogrel and/or warfarin therapy. This assay can easily be modified by designing new sets of primers. The capability of multiplexing up to 96 samples in one sequencing run resulted in a reasonably cost-effective approach to genotyping. Cost-effectiveness might further be improved with the smaller-scale, cheaper “Flongle” flow cells recently released by ONT.

The simple workflow of only one PCR reaction per sample was achieved by designing short amplicons, which were previously shown to be compatible with the MinION [27]. All contigs were amplified successfully, and despite the uneven yield among amplicons, the majority of genotypes could be accurately determined using appropriate bioinformatics tools. Genotyping accuracy was in part dependent on the variant calling tool. We showed that nanopolish achieved the best accuracy compared with marginCaller and VarScan 2. Instead of calling variants as single bases, nanopolish investigates a block of 10 bases at a time, and calculates the best likely haplotype using a hidden Markov model [18]. Nanopolish could detect both single base substitutions and a single base deletion.

Although generally accurate, a small proportion of incorrect genotype calls were observed. The majority of these were homozygote variants miscalled as heterozygotes. As previously reported, this seems to be a MinION and/or nanopolish specific problem [10]. Using nanopolish to genotype a fraction of variants on chromosome 20, the authors reported that 3,217 out of 4,781 mislabelled variants were due to this kind of error [10]. Improvements to variant calling tools for nanopore sequencing data may lead to more accurate calling.

Accuracy in MinION sequencing is also sequence-dependent [28]. This was well reflected in this assay with some genotypes showing better accuracy than the others. In particular, *CYP2C19**3 showed perfect accuracy regardless of the bioinformatic tools used to genotype. *VKORC1* rs9923231, on the contrary, showed the least accurate genotyping despite the relatively high read depth (> 100). However, these errors could be resolved by visual inspection of the alignment (BAM) file.

Conclusion

Drug-specific or broad pharmacogenetic assays are possible on the MinION sequencer, with high numbers of amplicons and sample multiplexing. High accuracy was achieved for most of the variants in our custom genotyping panel by applying suitable bioinformatic tools. The choice of variant calling tool is important, with nanopolish performing best for the assay described here. However, some mis-genotyping issues remain, which need to be resolved before such assays could be applied in a clinical or diagnostic setting. Continual improvements in chemistry and flowcells by ONT, and continuing

182 development of bioinformatic tools by the research community is likely to resolve these issues in the
183 future.

References

1. Gaedigk A, Ingelman-Sundberg M, Miller NA, Leeder JS, Whirl-Carrillo M, Klein TE. The Pharmacogene Variation (PharmVar) Consortium: Incorporation of the Human Cytochrome P450 (CYP) Allele Nomenclature Database. *Clinical Pharmacology and Therapeutics* 103(3), 399-401 (2018).
2. Johnson J, Cavallari L. Warfarin pharmacogenetics. *Trends Cardiovasc Med*. 25(1), 33-41 (2015).
3. Scott SA, Collet JP, Baber U *et al*. Exome sequencing of extreme clopidogrel response phenotypes identifies B4GALT2 as a determinant of on-treatment platelet reactivity. *Clinical Pharmacology and Therapeutics* 100(3), 287-294 (2016).
4. Zhang JH, Tang XF, Zhang Y *et al*. Relationship between ABCB1 polymorphisms, thromboelastography and risk of bleeding events in clopidogrel-treated patients with ST-elevation myocardial infarction. *Thrombosis Research* 134(5), 970-975 (2014).
5. Su J, Yu Q, Zhu H *et al*. The risk of clopidogrel resistance is associated with ABCB1 polymorphisms but not promoter methylation in a Chinese Han population. *PLoS ONE* 12(3), 1-16 (2017).
6. Sun Y, Wu Z, Li S *et al*. Impact of gamma-glutamyl carboxylase gene polymorphisms on warfarin dose requirement: A systematic review and meta-analysis. *Thrombosis Research* 135(4), 739-747 (2015).
7. Maurice CB, Barua PK, Simses D, Smith P, Howe JG, Stack G. Comparison of assay systems for warfarin-related CYP2C9 and VKORC1 genotyping. *Clinica Chimica Acta* 411(13-14), 947-954 (2010).
8. Zhou Y, Armstead AR, Coshatt GM, Limdi NA, Harada S. Comparison of Two Point-of-Care CYP2C19 Genotyping Assays for Genotype-Guided Antiplatelet Therapy. 47(6), 738-743 (2017).
9. Jain M, Olsen HE, Paten B *et al*. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biology* 17(1), 239-239 (2016).
10. Jain M, Koren S, Miga KH *et al*. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nature Biotechnology* 36(4), 338-345 (2018).
11. Garalde DR, Snell EA, Jachimowicz D *et al*. Highly parallel direct RN A sequencing on an array of nanopores. *Nature Methods* 15(3), 201-206 (2018).
12. Leggett RM, Clark MD. A world of opportunities with nanopore sequencing. *Journal of Experimental Botany* 68(20), 5419-5429 (2017).
13. Relling MV, Klein TE. CPIC: Clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. *Clin Pharmacol Ther* 89(3), 464-467 (2011).
14. Evans WE, Relling MV. Moving towards individualized medicine with pharmacogenomics. *Nature* 429(6990), 464-468 (2004).
15. Liang R, Wang C, Zhao H, Huang J, Hu D, Sun Y. Influence of CYP4F2 genotype on warfarin dose requirement-a systematic review and meta-analysis. *Thrombosis Research* 130(1), 38-44 (2012).
16. Sathananthan J, Seif El-Jack AK, Armstrong G *et al*. TCT-503 - Prevalence of CYP2C19 Variants and Associated Stent Thrombosis in Patients Undergoing Percutaneous Coronary Intervention. *Journal of the American College of Cardiology* 64(11), B148 (2014).
17. Jain M, Fiddes IT, Miga KH, Olsen HE, Paten B, Akeson M. Improved data analysis for the MinION nanopore sequencer. *Nature Methods* 12(4), 351-356 (2015).
18. Quick J, Loman NJ, Duraffour S *et al*. Real-time, portable genome sequencing for Ebola surveillance. *Nature* 530(7589), 228-232 (2016).
19. Koboldt DC, Zhang Q, Larson DE *et al*. VarScan 2 : Somatic mutation and copy number alteration discovery in cancer by exome sequencing VarScan 2 : Somatic mutation and copy

- number alteration discovery in cancer by exome sequencing. *Genome Research* 22(3), 568-576 (2012).
20. Leggett RM, Heavens D, Caccamo M, Clark MD, Davey RP. NanoOK: Multi-reference alignment analysis of nanopore sequencing data, quality and error profiles. *Bioinformatics* 32(1), 142-144 (2015).
21. Lanfear R, Schalamun M, Kainer D, Wang W, Schwessinger B. MinIONQC: fast and simple quality control for MinION sequencing data. *Bioinformatics* doi:10.1093/bioinformatics/bty654(July), 1-3 (2018).
22. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* 32(2), 292-294 (2015).
23. Robinson JT, Thorvaldsdóttir H, Winckler W *et al.* Integrative Genomics Viewer. *Nat Biotechnol* 29(1), 24-26 (2011).
24. Lek M, Karczewski KJ, Minikel EV *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536 285-291 (2016).
25. Quick J, Grubaugh ND, Pullan ST *et al.* Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nature Protocols* 12 1261-1276 (2017).
26. Castro-Wallace SL, Chiu CY, John KK *et al.* Nanopore DNA Sequencing and Genome Assembly on the International Space Station. *Scientific Reports* 7(1), 1-12 (2017).
27. Wei S, Williams Z. Rapid short-read sequencing and aneuploidy detection using minION nanopore technology. *Genetics* 202(1), 37-44 (2016).
28. Krishnakumar R, Sinha A, Bird SW *et al.* Systematic and stochastic influences on the performance of the MinION nanopore sequencer across a range of nucleotide bias. *Scientific Reports* 8(1), 1-13 (2018).

Table 1. Primers and variants detected in the pharmacogenetics assay panel

Amplicons	Primer pairs	Primer sequences (5'-3')	Variants detected
CYP2C192	CYP2C192_F CYP2C192_R	tttctgttggtgctgatattgcATGCAATAATTTTCCCACTATCATTGA acttgctgtcgctctatcttcAAGCAGTCACATAACTAAGCTTTTG	CYP2C19*2 (rs4244285), CYP2C19*10 (rs6413438), CYP2C19*7 (rs72558186)
CYP2C193	CYP2C193_F CYP2C193_R	tttctgttggtgctgatattgcTCCCTGCAATGTGATCTGCTC acttgctgtcgctctatcttcTACTTCAGGGCTTGGTCAATATAGAATT	CYP2C19*3 (rs4986893)
CYP2C194	CYP2C194_F CYP2C194_R	tttctgttggtgctgatattgcAGCATGGAGTGTTATAAAAAGCTTGG acttgctgtcgctctatcttcACATTGGTTAAGGATTTGCTGACATC	CYP2C19*4 (rs28399504), CYP2C19*15 (rs17882687)
CYP2C195	CYP2C195_F CYP2C195_R	tttctgttggtgctgatattgcATTACCGAACAGTTCTTGCATATTC acttgctgtcgctctatcttcGGTCAGGAATAAAAACAGCTCCATG	CYP2C19*5 (rs56337013), CYP2C19*16 (rs192154563)
CYP2C196	CYP2C196_F CYP2C196_R	tttctgttggtgctgatattgcGGATCTCCCTCCTAGTTTCGTTTC acttgctgtcgctctatcttcCAGTCCAGAAAGGTCAGTGATAGAG	CYP2C19*6 (rs72552267), CYP2C19*8 (rs41291556), CYP2C19*9 (rs17884712), CYP2C19*11 (rs58973490)
CYP2C1917	CYP2C1917_F CYP2C1917_R	tttctgttggtgctgatattgcGGGCTGTTTTCTTAGATAAATAAGTGG acttgctgtcgctctatcttcTTCTGATGCCCATCGTGGC	CYP2C19*17 (rs12248560)
B4GALT2	B4GALT2_F B4GALT2_R	tttctgttggtgctgatattgcTCTGTCCGTCCCCATCCTC acttgctgtcgctctatcttcCCACTGGGGGATCAGGGAT	B4GALT2 rs1061781
ABCB1	ABCB1_F ABCB1_R	tttctgttggtgctgatattgcCCACATGCTCCCAGGCTG acttgctgtcgctctatcttcCACCTGGGCATCGTGTCC	ABCB1 rs1045642
CYP2C92	CYP2C92_F CYP2C92_R	tttctgttggtgctgatattgcGGATCTCCCTCCTAGTTTCGT acttgctgtcgctctatcttcAGAGAAGATAGTAGTCCAGTAAGGTC	CYP2C9*2 (rs1799853), CYP2C9*8 (rs7900194), CYP2C9*14 (rs72558189)
CYP2C93	CYP2C93_F CYP2C93_R	tttctgttggtgctgatattgcCATGCCCTACACAGATGCTGT acttgctgtcgctctatcttcCCGGTGATGGTAGAGGTTTAAAAATG	CYP2C9*3 (rs1057910), CYP2C9*5 (rs28371686)
CYP2C96	CYP2C96_F CYP2C96_R	TttctgttggtgctgatattgcCCCGGGAACACAAACAAATTAC acttgctgtcgctctatcttcAAGCAGTCACATAACTAAGCTTTTGTT	CYP2C9*6 (rs9332131), CYP2C9*9 (rs2256871), CYP2C9*10 (rs9332130)
GGCX	GGCX_F GGCX_R	tttctgttggtgctgatattgcATAGGCCACTTATTCTTCCAGTAAATCA acttgctgtcgctctatcttcCCGGCGAAATACTCCTTTCCA	GGCX rs11676382
CYP4F2	CYP4F2_F CYP4F2_R	tttctgttggtgctgatattgcTTGTGTGTGTCTTTGAGGGAGG acttgctgtcgctctatcttcTGCTCATCAGTGTTTTCGGAA	CYP4F2 rs2108622
VKORC1rs7294	VKORC1rs7294_F VKORC1rs7294_R	tttctgttggtgctgatattgcGCAATGGAAAGAGCTTTGGAGAC acttgctgtcgctctatcttcCTGACCTCATCTGCTTTGCTTTG	VKORC1 rs7294
VKORC1rs9923231	VKORC1rs9923231_F VKORC1rs9923231_R	tttctgttggtgctgatattgcTTGGCCAGGCTTGTCTTAACT acttgctgtcgctctatcttcGAAACAGCATCTGGAGAGGGAG	VKORC1 rs9923231

*Lowercase indicates tails for MinION sample barcoding by PCR. Uppercase indicates sequence-specific primer regions.

Table 2. Genotyping accuracy of eight variants found in the samples

	Average read depth (x)	n samples	Discordant* sample(s)	Accuracy (%)
<i>B4GALT2 rs1061781</i>	176	21	1	95.2%
<i>ABCB1 rs1045642</i>	1020	29	0	100%
<i>CYP2C9*2</i>	204	26	0	100%
<i>CYP2C9*3</i>	45	25	2	92%
<i>GGCX rs11676382</i>	123	29	0	100%
<i>CYP4F2 rs2108622</i>	126	27	0	100%
<i>VKORC1 rs9923221</i>	168	30	7	76.7%
<i>VKORC1 rs7294</i>	814	32	1	96.9%

*Compared to Sanger Sequencing

Figure 1 Boxplot of read depth of amplicons

Figure 2 Genotyping accuracy across different bioinformatics workflow for three CYP2C19 SNPs.

Threshold for heterozygote variants was set at 0.2 and homozygote at 0.65, except for nanopolish, where there is no threshold for homozygote. Nanopolish uses its own algorithm to call a variant as homozygote based on the score of the variant compared to its reference sequence (<https://github.com/jts/nanopolish>). Homozygote variants mis-genotyped as heterozygote were considered false negative.

Figure 3. Bam files of homozygote (AA) samples (line 1 and 2), heterozygote (GA) samples (line 3 and 4) and wild-type (GG) samples (line 5 and 6) of VKORC1 rs9923231, as visualized in IGV.