

## **Validation of a HIV Whole Genome Sequencing Method for HIV Drug Resistance Testing in an Australian Clinical Microbiology Laboratory**

Frances JENKINS<sup>1</sup>, Thomas LE<sup>1</sup>, Rima FARHAT<sup>1</sup>, Angie PINTO<sup>1,2</sup>, M Azim ANZARI<sup>3</sup>,  
David BONSALL<sup>3</sup>, Tanya GOLUBCHIK<sup>4</sup>, Rory BOWDEN<sup>5</sup>, Frederick J LEE<sup>6,7</sup>, Sebastiaan  
VAN HAL<sup>1,7</sup>

<sup>1</sup>Department of Infectious Diseases and Microbiology, Royal Prince Alfred Hospital, Sydney,  
Australia

<sup>2</sup> The Kirby Institute, UNSW Australia, Sydney, Australia

<sup>3</sup> Wellcome Centre for Human Genetics, University of Oxford, Oxford, United Kingdom

<sup>4</sup> Department of Medical Sciences, Faculty of Medicine and Health, University of Sydney,  
Sydney, Australia

<sup>5</sup> The Walter and Eliza Hall Institute of Medical Research, Advanced Genomics Facility,  
Melbourne, Australia

<sup>6</sup> Department of Clinical Immunology and Allergy, Royal Prince Alfred Hospital, Sydney,  
Australia<sup>7</sup> Sydney Medical School, University of Sydney, Sydney, Australia

**Corresponding author:** Sebastiaan van Hal, Department of Infectious Diseases and  
Microbiology, Royal Prince Alfred Hospital, Missenden Road, Camperdown NSW 2050,  
Sydney, Australia ([sebastiaan.vanhal@health.nsw.gov.au](mailto:sebastiaan.vanhal@health.nsw.gov.au))

**NOTE:** This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

**Abstract:** Detection of HIV drug resistance is vital to successful anti-retroviral therapy (ART). HIV drug resistance (HIVDR) testing to determine drug resistance mutations (DRMs) is routinely performed in Australia to guide ART choice in either newly diagnosed people living with HIV or in cases of treatment failure. In 2022, our Australian clinical microbiology laboratory sought to validate a Next-generation sequencing (NGS)-based HIVDR assay to replace the previous Sanger sequencing (SS)-based ViroSeq assay. NGS solutions for HIVDR offer higher throughput, lower costs and higher sensitivity for variant detection. We sought to validate the previously described low-cost probe-based NGS method (veSEQ-HIV) for HIV-1 recovery and HIVDR testing in a diagnostic setting. The implemented veSEQ-HIV assay displayed 100% and 98% accuracy in major and minor mutation detection respectively and 100% accuracy of subtyping (provided >1000 mapped reads were obtained). Pairwise comparison exhibited low inter-and intra-run variability across the whole genome (Jaccard similarity coefficient  $[J] = 0.993$ ;  $J = 0.972$ ) and limited to the Pol gene only ( $J = 0.999$ ;  $J = 0.999$ ) respectively. The veSEQ-HIV assay met all our pre-set criteria based on the WHO “Recommended methods for validating an in-house genotyping assay for surveillance of HIV drug resistance” and has successfully replaced the ViroSeq assay in our laboratory. Scaling-down the veSEQ-HIV assay to a limited batch size and sequencing on the Illumina iSeq100, allowed easy implementation of the assay into the workflow of a small sequencing laboratory with minimal staff and equipment and the ability to meet clinically relevant test turn-around times.

**Keywords:** HIV, HIVDR, NGS, whole genome sequencing, drug resistance, genotyping, viral genomics, Illumina, antiretroviral therapy

## 1 **Introduction**

2 HIV is a major public health concern worldwide. While Australia has a low prevalence of  
3 HIV in the general population (0.1% on 2017 estimate), HIV detection and management still  
4 play an important role in transmission prevention especially in at-risk populations (Guy,  
5 2007; Kirby Institute, 2018). HIV drug resistance (HIVDR) testing to determine drug  
6 resistance mutations (DRMs) is routinely performed in Australia to improve chances of  
7 treatment success by informing anti-retroviral therapy (ART) choice in newly-diagnosed  
8 people living with HIV or in cases of treatment failure.

9 Royal Prince Alfred Hospital in Sydney, Australia routinely performs HIVDR testing and  
10 from 2016 until 2021, utilised the Sanger sequencing (SS) based commercial assay, ViroSeq  
11 HIV-1 genotyping system (Celera Diagnostic, Abbott Laboratories, Alameda, CA, USA).  
12 The ViroSeq system includes sample purification and reverse -transcription polymerase chain  
13 reaction steps, to generate a 1.8 kb amplicon covering parts of the polymerase (*Pol*) region  
14 including the protease (PR) and partial reverse transcriptase (RT) genes. The generated  
15 amplicons act as the SS template to generate an approximate 1.2kb chromatogram which after  
16 manual curation acts as the sequence input to the HIVdb Program available on the Stanford  
17 University HIVDR database website (<https://hivdb.stanford.edu/hivdb/by-patterns/>). The  
18 HIVdb Program identifies known DRMs and inferred levels of drug resistance to commonly  
19 used anti-retroviral drugs (Liu TF, 2006; Rhee, 2003; Shafer, 2006).

20 In 2020, our laboratory was advised that the ViroSeq system would cease manufacture in  
21 mid-2021. We therefore sought to replace this assay with a suitable alternative that would be  
22 able to detect DRMs in samples with a plasma viral load around 2000 copies per mL (cp/mL)  
23 at similar costs and overcome several limitations of the ViroSeq system: (1) able to obtain  
24 results across all HIV-1 subtypes; (2) include the integrase (IN) region; 3) able to detect sub-

25 consensus variants as several studies have demonstrated that these subpopulations contribute  
26 to negative clinical outcomes (Metzner, 2009; Simen, 2009); and 4) was future-proof as novel  
27 ART agents become available. Next-generation sequencing (NGS) solutions for HIVDR have  
28 been proposed as superior to SS methods due to higher throughput, lower costs and improved  
29 sensitivity to detection minority variants at lower levels (Lee, 2020; Taylor, 2019). Several  
30 NGS-based HIVDR methods have been reported, however, the challenge remains finding  
31 protocols that work consistently without sampling bias across all HIV-1 subtypes and viral  
32 loads.

33 A robust low-cost probe-based high-throughput NGS method for HIVDR and viral load  
34 (veSEQ-HIV) was previously described by Bonsall et. al (2020) (Bonsall, 2020). Previous  
35 evaluation of its performance in comparison to ViroSeq, showed a high degree of  
36 concordance in DRMs detected, comparatively higher sensitivity in detection of sub-  
37 consensus variants and the ability to detect complete sequences in a majority of samples  
38 (70%), increasing with viral load (Fogel, 2020).

39 As such, the veSEQ-HIV assay best met our requirements. However, the veSEQ-HIV assay  
40 was designed to run in excess of 96 samples per batch (Bonsall, 2020). In our diagnostic  
41 setting, with a need to meet clinically relevant turn-around times, we reduced this batch size  
42 to 8 patient samples and performed sequencing on the Illumina iSeq100. These adjustments  
43 allowed us to implement the assay into the workflow of our small sequencing laboratory with  
44 minimal staff and equipment.

45 Here, we describe the validation of the veSEQ-HIV method in its utility for HIVDR testing as  
46 an in-house IVD in an Australian clinical microbiology laboratory.

## 47 **Materials and methods**

### 48 **Ethics**

49 Human Research Ethics Committee of the Sydney Local Health District approved the use of  
50 stored plasma samples in this study (2019/ETH01706).

### 51 **Study population/Sample selection**

52 Sixty-six stored plasma samples with known DRMs from both treatment naïve and -  
53 experienced people living with HIV-1 that had undergone testing on the ViroSeq system  
54 between 2016-2021 were included. Associated plasma viral loads were obtained from  
55 laboratory records and ranged from 950 to  $>10^6$  cp/mL. Subtypes included B (n=48), C (n=3),  
56 CRF01\_AE (n=13) and CRF02\_AG (n=2). Isolate baseline characteristics are summarized in  
57 **Table 1.**

58 As the ViroSeq assay does not cover the IN region, we included five samples that had had  
59 SS-based IN testing performed at an external accredited laboratory and had Stanford HIVdb  
60 reports available. We did not have access to nucleotide sequence for these samples and only  
61 included them in our comparisons of subtype and mutation detection.

62 For reference, a dilution series ( $10^5$  – 25 cp/mL) of the highest AcroMetrix HIV-1 Panel  
63 standard was prepared (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

64 A negative control (PCR confirmed HIV-negative plasma) was taken through the entire  
65 process from extraction to sequencing for each run.

### 66 **Nucleic acid extraction**

67 Methods were based on those described for veSEQ-HIV (Bonsall, 2020) with minor  
68 modifications. Briefly, automated total nucleic acid (NA) extraction was performed using the  
69 NucliSENS e-MAG (Biomérieux, Marcy-l'Étoile, France). Extraction was performed using

70 the manufacturers general extraction protocol with an input volume of 500  $\mu$ L and elution of  
71 25  $\mu$ L.

## 72 **cDNA generation and library preparation**

73 NA was concentrated using RNAClean XP beads (Beckman Coulter, Brea, California, USA)  
74 before quarter reactions for cDNA generation without RNA fragmentation were prepared  
75 using the SMARTer Stranded Total RNA-Seq Kit v2 – Pico Input Mammalian Kit (Takara  
76 Biosciences USA, San Jose, California, USA) with maximal NA input.

77 Libraries were prepared using the same kit. Indexed libraries were pooled equally by volume  
78 (8 patient and 1 negative control) before performing clean-up using a ratio of 0.68 Ampure  
79 XP beads (Beckman Coulter, Brea, California, USA). DNA quantification was performed  
80 using Qubit HS DNA kit and Qubit fluorometer (Thermo Fisher Scientific, Waltham,  
81 Massachusetts, USA) and fragment size determined using the HS D1000 assay on Tape  
82 Station 4200 (Agilent Technologies, Santa Clara, California, USA).

## 83 **Capture-based hybridization, post-capture PCR and clean-up**

84 Up to 500 ng of pooled library was taken through viral capture using the xGen hybridization  
85 and wash kit (Integrated DNA Technologies, Singapore). The probe pool contained 646  
86 custom 120-mer biotinylated probes targeting the HIV genome across multiple subtypes. The  
87 custom probe sequences are published here for the first time with permission  
88 (**Supplementary file 2**). Post-capture PCR was performed as per manufacturer's instructions.

89 Clean-up was performed with a ratio of 1.5x Ampure XP beads (Beckman Coulter, Brea,  
90 California, USA). DNA quantification was performed using Qubit HS DNA kit and  
91 fluorometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and fragment size  
92 determined using HS D1000 assay on Tape Station 4200 (Agilent Technologies, Santa Clara,  
93 California, USA). An additional post-capture PCR (6-cycles; total 16 cycles) was added and

94 clean-up repeated if minimum library molarity (1nM for iSeq100) was not meet at the final  
95 library stage.

## 96 **Sequencing**

97 Final libraries were prepared for sequencing as per the Illumina iSeq 100 Sequencing System  
98 Product Documentation (# 200015511 v00). A final loading concentration of 75 pM was used  
99 with a 2% PhiX control spike. Libraries were run on the Illumina iSeq100 system using iSeq  
100 100 i1 Reagent v2 (300-cycles) with read lengths of 2 x 150 bp.

## 101 **Bioinformatics**

102 Raw short-read data was processed by an in-house automated pipeline utilising open-source  
103 software as described previously with minor modifications (Bonsall, 2020). Briefly, Kraken  
104 was first used to filter out human reads followed by trimming of adapters using fastp. Host-  
105 depleted, filtered reads were then taken forward including generating assemblies using and  
106 SPAdes which formed the inputs to assemble the contigs suitable as input for SHIVER (v3)  
107 (Wood DE, 2019; Wymant, 2018).

108 An updated Python 3 compliant SHIVER was used for analysis setting a previously validated  
109 minimum depth of 5-15 against the reference genome (Human immunodeficiency virus type  
110 1 - HBX2) (Ji, 2018). The resulting output is then run the Stanford University HIV Drug  
111 Resistance database to identify drug resistance mutations using the tool sierrapy (Liu TF,  
112 2006; Shafer, 2006; Rhee, 2003). Subsequent reports are generated using an in-house script  
113 based on the Stanford json file output.

## 114 **Assessment of assay performance**

115 Validation criteria were pre-defined and based on WHO “Recommended methods for  
116 validating an in-house genotyping assay for surveillance of HIV drug resistance” with minor

117 modifications (WHO, 2020). In brief, these criteria included that 100% of known mutations  
118 must be detected and  $\geq 90\%$  of pairwise comparisons must be at least 98% identical. The  
119 WHO guidance also recommended meeting sensitivity for amplification  $\geq 95\%$  of samples  
120 with viral loads between 500 and 1000 cp/mL amplified and successfully genotyped ( $n \geq 10$ ).  
121 This criteria had to be modified as the ViroSeq (our “gold-standard”) is not suitable for viral  
122 loads  $< 2000$  cp/mL to sensitivity for amplification  $\geq 95\%$  of samples with viral loads  $> 2000$   
123 cp/ml must be amplified and successfully genotyped ( $n \geq 10$ ).

124 Performance metrics used to assess the new method were as follows: (1) accuracy; (2)  
125 precision/reproducibility; (3) sensitivity.

## 126 **Accuracy**

127 Using sensitivity or specificity calculations in the context of sequencing comparisons are  
128 impractical and misleading as there are no “real” true negatives especially when including  
129 minor mutations in an RNA virus that has a low barrier for genomic change.

130 Therefore, the Jaccard similarity coefficient (J) was employed which gauges the diversity and  
131 similarity of samples. The Jaccard coefficient measures similarity between finite sample sets,  
132 and is defined as the size of the intersection divided by the size of the union of the sample  
133 sets.

$$134 \quad J(\text{ViroSeq}, \text{WGS}) = \frac{(| \text{ViroSeq} \cap \text{WGS} |)}{(| \text{ViroSeq} \cup \text{WGS} |)}$$

135 By design values are between 0 and 1 with higher values indicating greater similarity.

136 The level of agreement between the new test and the reference test was assessed by  
137 comparing consensus nucleotide sequences, amino acid sequences, subtype, major and minor  
138 mutations detected. To compare nucleotide sequences, both WGS data and the Viroseq  
139 FASTA files were mapped to HIV-1 reference HBX2 using clustalW (Chenna, 2003).



140 Relevant regions of the *Pol* region (1.3kbp) were extracted from the aligned files. Bases  
141 common to both datasets including ambiguous assigned bases in Viroseq data were extracted  
142 as well as mutations through pairwise comparisons.

### 143 **Linearity, Limit of Detection and Limit of Quantification**

144 Both the limit of detection (LoD) and limit of quantification (LoQ) were assessed using a  
145 panel of standards comparing viral load to the number of reads mapping to the reference  
146 (HIV-1- HBX2). Although, the veSEQ-HIV assay is capable of estimating viral load based  
147 on reads mapping to the reference provided a panel of quantification standards is included in  
148 each run (Bonsall, 2020), this was not formally assessed given the predicted number of  
149 samples in our setting.

### 150 **Precision/Reproducibility**

151 Inter- and intra-run agreement was assessed by running replicates on different runs and  
152 within the same run. For intra-run comparison, a set of five samples were included on the  
153 same run across three different runs; viral loads ranged between 3475 and 100,000 cp/mL and  
154 included subtypes B, C and CRF01\_AE. For inter-run comparison, a set of eight samples  
155 were included on two different runs; viral loads ranged between 3435 and 100, 000 cp/mL  
156 and included subtypes B and CRF01\_AE. A summary of the replicates can be found in **Table**  
157 **2** below.

158 The WHO recommends (3 x 5 replicates) to determine inter-run precision, due to limited  
159 remaining plasma from previously tested samples, inter-run precision was determined using a  
160 (2 x 8 replicates) approach.

161 **Specificity**

162 Specificity was assessed by a negative control (PCR HIV negative plasma) included in all  
163 runs. Subsequent routine clinical runs have included a phosphate-buffered saline negative  
164 control. Possible interfering substances were not assessed as sequencing would only be  
165 performed on confirmed HIV positive samples.

166 **Quality control Criteria**

167 A set of quality control (QC) criteria were established to assess both run and sample data  
168 quality prior to analysis. We included sequencing run metrics as specified by the  
169 manufacturer (Illumina) as well as criteria to review individual raw read data, inclusive of the  
170 negative control. For review of individual sample data, we referred to guidelines established  
171 by the Winnipeg Consensus for read quality control (QC), read alignment, reference  
172 mapping, HIVDR interpretation and reporting (Ji, 2018).

173 **Clinical Reporting**

174 Development of our clinical reporting strategy involved discussion with treating clinicians in  
175 our network and well as incorporating requirements set by our accreditation, WHO  
176 recommendations and guidelines established by the Winnipeg Consensus (Ji, 2018; WHO,  
177 2020).

178 **Results**

179 **Accuracy**

180 **Accuracy of nucleotide sequences**

181 The pairwise analysis of the *Pol* region resulted in an overall Jaccard coefficient of similarity  
182 of 0.998, with 98,459 bases shared and 145 SNPs (single nucleotide polymorphisms)  
183 differences between isolate pairs.

184 **Accuracy of subtyping**

185 Concordant subtypes were achieved for 98% (n=65/66) of patient samples. One sample was  
186 misidentified as subtype CRF\_14\_BG rather than subtype B. This sample had low viral load  
187 (2790 cp/ mL), resulting in less than 1000 mapped reads and suboptimal coverage across the  
188 PR gene. Provided samples obtained >1000 mapped reads 100% accuracy of subtyping could  
189 be obtained.

190 **Accuracy of Major Drug Mutations.** Due to the low prevalence and success of current  
191 ART, a limited repertoire of major drug mutations could be sourced and included 5 different  
192 PR, 19 NNRTI/NRTI and 5 IN major mutations. There were 19 discordant major drug  
193 mutation calls between the two methods. All discordances were able to be resolved. The  
194 majority (n=9) on the ViroSeq chromatogram were assigned as an ambiguous base with  
195 reporting calling both wild-type and a DRM. These were all present on WGS but at a sub-  
196 consensus level (defined when <50% but >5% of reads mapping to the alternative allele at a  
197 >20 fold read depth). The remaining discordance were as a result of: (1) incorrect manual  
198 base assignments on ViroSeq (n=5); and (2) no coverage across region of interest on WGS  
199 (n=5). No false positive DRMs were detected on WGS. Taking these additional results into  
200 account, the overall accuracy/concordance for detection of major mutations was 100% with a  
201 Jaccard similarity coefficient of 1.0. A summary of these results can be found in **Table 3**.

202 **Accuracy of Minor Drug Mutations**

203 Only accuracy for PR and RT were examined. Three (3/601; 0.4%) and 29 (29/1249; 2%)  
204 discordant minor PR and RT mutation calls were identified respectively while 1 (1/569;  
205 0.2%) and 13 (13/2499; 1%) potential WGS false positives were identified. The majority of  
206 the false positive (n=10) minor mutations occurred on the boundaries of the ViroSeq  
207 amplicons resulting in higher uncertainty of basecalls. Irrespective, the overall

208 accuracy/concordance for detection of Minor mutation was 98%. The overall Jaccard  
209 coefficient of similarity was 0.975. A summary of concordance of minor drug mutations is  
210 available in **Supplementary Table 4**.

## 211 **Limit of Detection and Limit of Quantification**

### 212 **Mapped reads**

213 Eight samples (8/82; 10%) run on WGS had suboptimal coverage (<1000 mapped reads).  
214 Five (5/8; 63%) of these samples had low viral loads (2000 - 5000 cp/mL). The depth of  
215 coverage affected all regions when <1000 mapped reads were obtained while >1000 mapped  
216 reads gave excellent coverage depth for all regions. Results are summarised in **Table 4**.

### 217 **By viral load**

218 A significant linear correlation ( $R=0.74$ ;  $p<0.01$ ) between viral load and number of mapped  
219 reads. (Figure 1A and 1B) was detected. Low viral loads (2000 - 5000 cp/mL) gave  
220 inconsistent results but were linked to the number of mapped reads with <1000 mapped reads  
221 in six of 13 isolates (46%). The remaining seven represented possible failed WGS runs.  
222 These were unable to be repeated as no further sample was available. Furthermore, as all  
223 samples had been stored with an unknown number of freeze-thaw cycles, re-quantification  
224 was not possible to exclude sample RNA degradation as the cause.

### 225 **Linearity**

226 A significant linear relationship ( $R=0.99$ ;  $p<0.001$ ) was detected when including a  
227 commercial standard AcroMetrix HIV-1 Panel (Thermo Fisher Scientific, Waltham,  
228 Massachusetts, USA). With decreasing viral loads a reduction in: (1) the percentage coverage  
229 across genes; and (2) the number of mutations was detected. At low viral loads <2000 cp/mL  
230 potential false positive (FP) mutations were detected (**Supplementary Table 1**). A limit of

231 detection for HIV-1 was determined to be 100 cp/mL. With respect to reporting the number  
232 of mapped reads was the best correlate with optimal reporting requiring a minimum of 1000  
233 mapped reads.

#### 234 **Inter-run (reproducibility) and intra-run (repeatability) of WGS method**

235 The inter-run variability (good Jaccard similarity index) both across the whole genome  
236 ( $J=0.993$ ) as well as across the *Pol* gene ( $J=0.999$ ) was within the expected performance with  
237 an average of 437 SNPs across the whole genome and 4 SNPs across the *Pol* gene between  
238 runs.

239 Similarly, the intra-run variability (good Jaccard similarity index) was adequate both across  
240 the whole genome ( $J=0.972$ ) as well as across the *Pol* gene ( $J=0.999$ ) with an average of 130  
241 SNPs across the whole genome and 4 SNPs across the *Pol* gene between runs. These results  
242 are summarised in **Table 2**.

#### 243 **Specificity**

244 No evidence of contamination across all validation runs or subsequent clinical runs was  
245 detected with an average of 20 (range 0-548) reads for the negative control. Although reads  
246 represented HIV reads in a small subset, all negative control assemblies failed and mapping  
247 to the reference (HIV-1- HXB-2) was unsuccessful with 0% coverage and 0% depth across  
248 the three regions (PR, RT, IN) of the HIV genome.

#### 249 **Quality Control Criteria**

250 QC metrics and their corresponding limits/criteria for sequencing runs are presented in

251 **Supplementary Table 2**.

## 252 **Clinical reporting**

253 In the transition to NGS-based methods, WHO has recommended presenting resistance  
254 results in the same way as previous SS-based methods (WHO, 2020). Thus, reports include  
255 the same data output generated by the ViroSeq assay- subtype identification and lists  
256 identified DRMs mutations and their associated susceptibility interpretations.

257 With interest expressed from the treating clinicians in our network, the presence of sub-  
258 consensus variants containing major drug resistance mutations are also included (reporting  
259 when frequency greater than 5% and with appropriate caveats) (Ji, 2018). These findings are  
260 presented in an iterative way with additional mutation score and interpretation change  
261 reported alongside consensus interpretations. With the clinical relevance of sub-consensus  
262 variants of <20% frequency still under debate, this leaves the choice to consider any changes  
263 in interpretation to the treating clinician. An example report can be found in **Supplementary**  
264 **file 1**.

## 265 **Discussion**

266 In this study, we provide our in-house validation of the veSEQ-HIV assay in its utility for  
267 HIVDR in an Australian clinical microbiology laboratory. The veSEQ-HIV assay with minor  
268 modifications, met our pre-specified validation criteria based on the WHO “Recommended  
269 methods for validating an in-house genotyping assay for surveillance of HIV drug resistance”  
270 (WHO, 2020). We therefore deemed the assay fit for purpose and able to replace the ViroSeq  
271 assay in our laboratory.

272 We achieved a Jaccard similarity coefficient (J) of 0.998 for pairwise nucleotide comparisons  
273 of the *Pol* gene and detected 100% of major mutations (Criteria 1: Accuracy 100% of known  
274 mutations must be detected).

275 A limitation of our validation includes the small number of mutations encountered. We only  
276 encountered (11/56; 19%) NRTI; (8/49; 16%) NNRTI, 5/57; 9% PR and 3/43; 7% IN DRMs  
277 (All DRMs encountered are listed in **Supplementary Table 3**). Ideally, our validation set  
278 would have included all known mutations that affect clinical interpretation i.e. all known  
279 DRMs, however, practically, this is infeasible as hundreds of mutations would need to be  
280 covered (WHO, 2020). We attempted to cover as many mutations as possible but were  
281 limited to samples that had previously been referred to our laboratory for testing that had  
282 sufficient remaining plasma. Ongoing clinical testing should add to the DRMs encountered.

283 Pairwise comparison exhibited low inter-and intra-run variability across the whole genome  
284 ( $J=0.993$ ;  $J=0.972$  respectively). Agreement was highest when limited only to the *Pol* gene  
285 (both  $J=0.999$ ). (Criteria 2:  $\geq 90\%$  of pairwise comparisons must be at least 98% identical).  
286 This supports previous finding of relative conservation across the *Pol* gene relative to the rest  
287 of the genome (Troyano-Hernández, 2022).

288 We identified 100% of subtypes correctly (Criteria 3: Sensitivity for amplification:  $\geq 95\%$  of  
289 samples with viral loads  $>2000$  cp/mL must be amplified and successfully genotyped ( $n$   
290  $\geq 10$ )).

291 There were 4 subtypes (B, C, CRF01\_AE and CRF02\_AG) represented in our sample set.  
292 The majority of samples were subtype B ( $n=48/66$ ; 72%) followed by subtype CRF01\_AE  
293 ( $n=20/66$ ; 18%). This is in keeping with previous Australian epidemiological data identifying  
294 subtype B as the most prevalent and CRF01\_AE as one of the dominant non-B subtypes of  
295 increasing prevalence (Castley, 2017). This indicates that while we could not equally  
296 represent all HIV genotypes due to lack of sample availability the subtypes in our sample are  
297 representative of circulating HIV in Australia. Irrespective, we do not expect there to be any  
298 negative bias in “uncommon” genotypes as veSEQ-HIV probes have been designed and

299 previously shown to consistently recover genomes from these “uncommon” subtypes  
300 including C, A (A1 &A2), D, G and J (Bonsall, 2020).

301 The veSEQ-HIV assay is now in routine use in our laboratory and has been used to  
302 successfully sequence and generate over 100 HIVDR and genotyping reports to clinicians  
303 across the state of New South Wales, Australia.

304 Running limited batches (eight patient samples) has allowed easy implementation of the  
305 assay into the workflow of our small sequencing laboratory with minimal staff and equipment  
306 and has been an ideal use of a small desktop sequencer such as the Illumina iSeq100. Despite  
307 small batches, reagent cost per assay is still significantly reduced compared to the previous  
308 commercial assay (from \$360 to \$200). Assay batching flexibility also allows for future  
309 throughput increases utilising automated liquid-handing for library preparation and larger  
310 sequencing platforms (e.g. Illumina MiSeq).

311 Processing of NGS data has been identified as a major hurdle in implementation of NGS-  
312 based HIVDR. The Winnipeg Consensus has established preliminary guidelines for read  
313 quality control (QC), read alignment and reference mapping, variant calling and QC, HIVDR  
314 interpretation and reporting as well as general analysis of data management (Ji, 2018). These  
315 guidelines were invaluable in developing our bioinformatics, QC and clinical reporting  
316 strategy.

317 Further hurdles for NGS-based HIVDR include lack of specifically developed external  
318 quality assessment (EQA) strategies and programs (Lee, 2020; Ji, 2018) as current HIVDR  
319 EQA are still designed for SS-based methods. In the transition to NGS-based methods, WHO  
320 has recommended configuring NGS methods to present results in the same way as SS-based  
321 methods (WHO, 2020). Thus, while not ideal, NGS-based HIVDR can still be used to  
322 produce results formatted as SS-based HIVDR outputs, permitting participation in current



323 EQA programs such as the HIVDR programs from Quality Control for Molecular  
324 Diagnostics (QCMD) that our laboratory participates in.

325 Informal sample exchange between laboratories offers an alternative transitional EQA  
326 strategy to meet accreditation needs, however, very few laboratories are currently performing  
327 NGS-based HIVDR in Australia. We look forward to more laboratories taking up NGS-based  
328 methods to make this another transitional EQA option.

329 Despite this, there remains an urgent need for appropriate EQA programs and materials to be  
330 made available as these approaches overlook the abundance and available complexity in NGS  
331 HIVDR data, particularly its ability to detect sub-consensus variants at low levels (Lee,  
332 2020).

333 In addition, appropriate reference materials for validation of sub-consensus variant detection  
334 do not yet exist. Our reporting strategy for sub-consensus variants followed the  
335 recommendations from the Winipeg consensus with a 5% frequency threshold suggested to  
336 account for errors/bias introduced during assay steps such as reverse-transcription, PCR,  
337 sequencing steps etc. (Ji, 2018; Casadellà, 2017). While we did not quantify the sensitivity of  
338 variant detection in this validation, this threshold is likely conservative with estimates of  
339 minority variant frequency likely to be more robust using the probe-based veSEQ-HIV  
340 method as biases introduced by PCR are minimised and contamination is computationally  
341 controlled (Bonsall, 2020). Our reporting includes necessary caveats to sub-consensus variant  
342 reporting and leaves the choice act on subsequent drug susceptibility interpretation changes to  
343 the treating clinician.

344 Well-characterised sample panels including both clinical and synthetically constructed  
345 samples containing sub-consensus variants are needed to assess each laboratory's ability to  
346 detect sub-consensus variants accurately (Lee, 2020).

347 **Conclusion**

348 The veSEQ-HIV assay provides a robust and cost-effective NGS method for HIVDR in a  
349 diagnostic setting. veSEQ-HIV demonstrated satisfactory and consistent performance  
350 including low inter-and intra-run variability and good accuracy in detection of DRMs in  
351 samples with >1000 mapped reads (achievable in samples with VL >1000 cp/mL) when  
352 compared to the “gold-standard” SS-based ViroSeq assay. The veSEQ-HIV assay met all our  
353 pre-set criteria based on the WHO “Recommended methods for validating an in-house  
354 genotyping assay for surveillance of HIV drug resistance” and has successfully replaced the  
355 previous SS-based ViroSeq assay in our laboratory.

356 The assay meets current needs for clinical reporting (as SS-like results) to meet requirements  
357 of our diagnostic laboratory accreditation and to participate in current EQA programs as well  
358 as the ability to report on the presence of sub-consensus variants with appropriate caveats. As  
359 HIVDR transitions from SS-based methods to NGS and the clinical relevance of sub-  
360 consensus variants is further clarified, the veSEQ-HIV meets both current and future needs  
361 for HIVDR testing in a diagnostic setting.

362 **Funding**

363 None

364 **Declaration of Competing Interest**

365 All authors have completed the ICMJE uniform disclosure form at  
366 [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) and declare: no support from any organisation for the  
367 submitted work; FJL is on the Gilead Sciences Advisory board and has received consulting  
368 fees and payments for educational events.

369

370 **Acknowledgements**

371 Thanks to Alicia Beukers for her preliminary work in this validation. Thanks also go to the  
372 Clinical Immunology Laboratory at RPAH for providing the study samples and sharing the  
373 ViroSeq comparison data. Thanks to the Oxford group for providing the custom probe  
374 sequences to allow us to perform the veSEQ-HIV in our laboratory.

375 **Contributions**

376 FJ drafted the manuscript.

377 FJ, TL, RF performed the validation testing

378 SVH, AP, FJL designed the validation

379 SVH developed the in-house pipeline, performed the bioinformatics and statistical analysis

380 MAA developed the custom probe sequences

381 TG assisted with pipeline development

382 DB, RB shared protocols and assisted with assay implementation

383 DB, RB, SVH, AP, FJL reviewed the manuscript

384 **Appendix. Supplementary materials**

385 Supplementary file 1: Example Clinical Report

386 Supplementary file 2: Probe sequences

387 Supplementary file 3: Supplementary data

388

389 **References**

- 390 Bonsall, D. G.-C.-D.-M. (2020). A Comprehensive Genomics Solution for HIV Surveillance  
391 and Clinical Monitoring in Low-Income Settings. *Journal of Clinical Microbiology*,  
392 58(10). doi:<https://doi.org/10.1128/JCM.00382-20>
- 393 Casadellà, M. &. (2017). Deep sequencing for HIV-1 clinical management. *Virus research*,  
394 239, 69–81. doi:<https://doi.org/10.1016/j.virusres.2016.10.019>
- 395 Castley, A. S. (2017, May 10). Australian Molecular Epidemiology Network-HIV (AMEN-  
396 HIV). A national study of the molecular epidemiology of HIV-1 in Australia 2005-  
397 2012. *PLoS One*, 12(5), e0170601. doi:doi: 10.1371/journal.pone.0170601
- 398 Chenna, R. S. (2003). Multiple sequence alignment with the Clustal series of programs.  
399 *Nucleic acids research*, 31(13), 3497–3500. doi:<https://doi.org/10.1093/nar/gkg500>
- 400 Chris Wymant, F. B. (2018). Easy and accurate reconstruction of whole HIV genomes from  
401 short-read sequence data with shiver. *Virus Evolution*, 4(1), vey007.
- 402 Fogel, J. M. (2020). Performance of a high-throughput next-generation sequencing method  
403 for analysis of HIV drug resistance and viral load. *The Journal of antimicrobial*  
404 *chemotherapy*, 75(12), 3510–3516. doi:<https://doi.org/10.1093/jac/dkaa352>
- 405 Guy, R. J. (2007). HIV diagnoses in Australia: diverging epidemics within a low-prevalence  
406 country. *The Medical journal of Australia*, 187(8), 437–440.  
407 doi:<https://doi.org/10.5694/j.1326-5377.2007.tb01353.x>
- 408 Ji, H. E.-R.-J. (2018). Bioinformatic data processing pipelines in support of next-generation  
409 sequencing-based HIV drug resistance testing: the Winnipeg Consensus. *Journal of*  
410 *the International AIDS Society*, 21(10), e25193.  
411 doi:<https://doi.org/10.1002/jia2.25193>

- 412 Kirby Institute. (2018). *HIV, viral hepatitis and sexually transmissible infections in Australia:*  
413 *annual surveillance report 2018*. Kirby Institute. Sydney: UNSW Sydney. Retrieved  
414 from <https://kirby.unsw.edu.au/report/asr2018>
- 415 Lee, E. R. (2020). External Quality Assessment for Next-Generation Sequencing-Based HIV  
416 Drug Resistance Testing: Unique Requirements and Challenges. *Viruses*, *12*(5), 550.  
417 doi:<https://doi.org/10.3390/v12050550>
- 418 Liu TF, S. R. (2006, April 28). Web Resources for HIV type 1 Genotypic-Resistance Test  
419 Interpretation. *Clin Infect Dis*, *42*(11), 1608-18.
- 420 Metzner, K. J. (2009). Minority quasispecies of drug-resistant HIV-1 that lead to early  
421 therapy failure in treatment-naive and -adherent patients. *Clinical infectious diseases :*  
422 *an official publication of the Infectious Diseases Society of America*, *48*(2), 239–247.  
423 doi:<https://doi.org/10.1086/595703>
- 424 Rhee, S. Y. (2003). Human immunodeficiency virus reverse transcriptase and protease  
425 sequence database. *Nucleic Acids Research*, *31*(1), 298-303.  
426 doi:<https://doi.org/10.1093/nar/gkg100>
- 427 Shafer, R. W. (2006). Rationale and Uses of a Public HIV Drug-Resistance Database.  
428 *Journal of Infectious Diseases*, *194*(Suppl 1), S51-8.
- 429 Simen, B. B. (2009). Low-abundance drug-resistant viral variants in chronically HIV-  
430 infected, antiretroviral treatment-naive patients significantly impact treatment  
431 outcomes. *The Journal of infectious diseases*, *199*(5), 693–701.  
432 doi:<https://doi.org/10.1086/596736>

- 433 Taylor, T. L. (2019). A MiSeq-HyDRA platform for enhanced HIV drug resistance  
434 genotyping and surveillance. *Scientific reports*, 9(1), 8970.  
435 doi:<https://doi.org/10.1038/s41598-019-45328-3>
- 436 Troyano-Hernández, P. R. (2022). Genetic Diversity and Low Therapeutic Impact of Variant-  
437 Specific Markers in HIV-1 Pol Proteins. *Frontiers in microbiology*, 13, 866705.  
438 doi:<https://doi.org/10.3389/fmicb.2022.866705>
- 439 WHO. (2020). *WHO HIVResNet HIV DRUG RESISTANCE; LABORATORY OPERATIONAL*  
440 *FRAMEWORK; ANNEX 3. Recommended methods for validating an in-house*  
441 *genotyping assay for surveillance of HIV drug resistance*. World Health Organization.  
442 Retrieved from <https://www.jstor.org/stable/pdf/resrep27859.13.pdf>
- 443 Wood DE, L. J. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology*,  
444 20(1), 257.
- 445 Wymant, C. B.-B. (2018). Easy and accurate reconstruction of whole HIV genomes from  
446 short-read sequence data with shiver. *Virus evolution*, 4(1), vey007.  
447 doi:<https://doi.org/10.1093/ve/vey007>
- 448

1 **Table 1. Isolate baseline characteristics**

Subtype	Number of samples	Mean viral load (cp/mL)	Viral load range	Number Surveillance Drug Resistance Mutations		
				PR**	RT**	IN
B	48	364275	2260-2914185	6	54	5*
C	3	41444	14309-81222	0	5	0*
CRF01_AE	13	167506	1700-847000	0	26	0*
CRF02_AG	2	53025	1050- 105000	0	4	0*

2 \* Only 5/66 (8%) samples had integrase testing at baseline. Integrase is not included in ViroSeq assay.

3 \*\* Surveillance Drug Resistance Mutations as determined by ViroSeq assay.

4 **Table 2. Intra- and inter-run comparison including Jaccard Similarity and SNP numbers across whole-genome and across *Pol* region**

	<b>Number of samples</b>	<b>Replicates per sample</b>	<b>Viral load (cp/mL)</b>	<b>Subtypes</b>	<b>Jaccard Similarity- whole genome (mean SNPs)</b>	<b>Jaccard Similarity- Pol gene (mean SNPs)</b>
<b>Intra-run</b>	5	3	3435 - 100,000	B, C, CRF01_AE	0.972 (130)	0.999 (4)
<b>Inter-run</b>	8	2	3435 - 100,000	B, CRF01_AE	0.993 (437)	0.999 (4)



5 **Table 3. Concordance of major drug mutations between methods**

<b>Major mutation location</b>	<b>Concordant mutation calls</b>	<b>Discordant mutation calls</b>	<b>Presence of sub-population</b>	<b>ViroSeq poor coverage/false calls</b>	<b>WGS poor coverage</b>	<b>Discordance resolved (%)</b>	<b>Overall accuracy</b>
<b>PR</b>	6	3	1	2	-	100 (n=3/3)	100% (n=6/6)
<b>RT</b>	84	16	8	3	5	100 (n=16/16)	100 (n=84/84)
<b>IN</b>	5*	-	-	-	-	-	100 (n=5/5)
<b>Total</b>	-	-	9	5	5	100 (n=19/19)	100 (n=95/95)

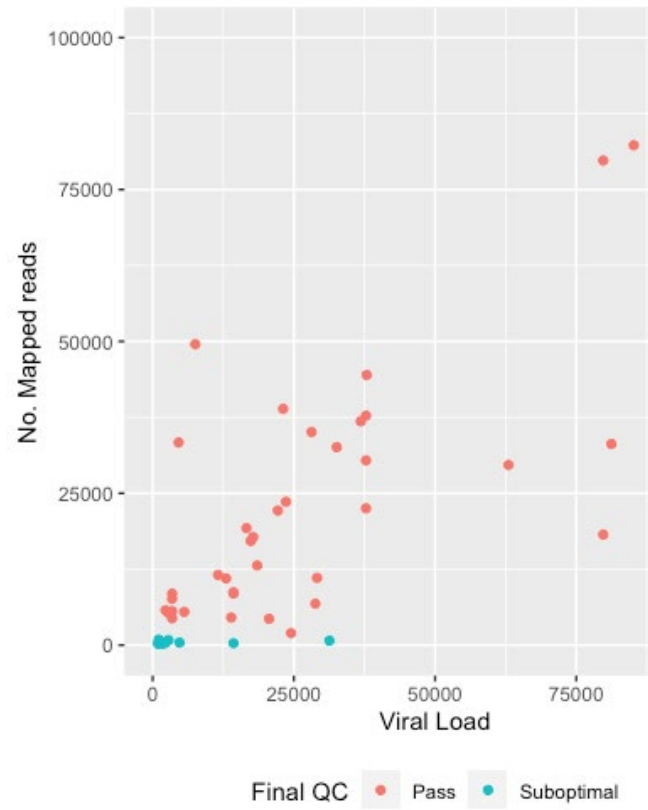
6 Please note: Replicates were included in mutation analysis

7 \* SS IN results only available for 5/66 samples

8 **Table 4. WGS QC metrics based on number of mapped reads**

No. mapped Reads	No. of samples	Average Depth Coverage		
		PR region	RT region	IN region
<1000	8	71% @ 9 fold	42% @ 8 fold	52% @ 7 fold
>1000	78	99% @ 2914 fold	99% @ 3465 fold	99% @ 3372 fold

1A



1B

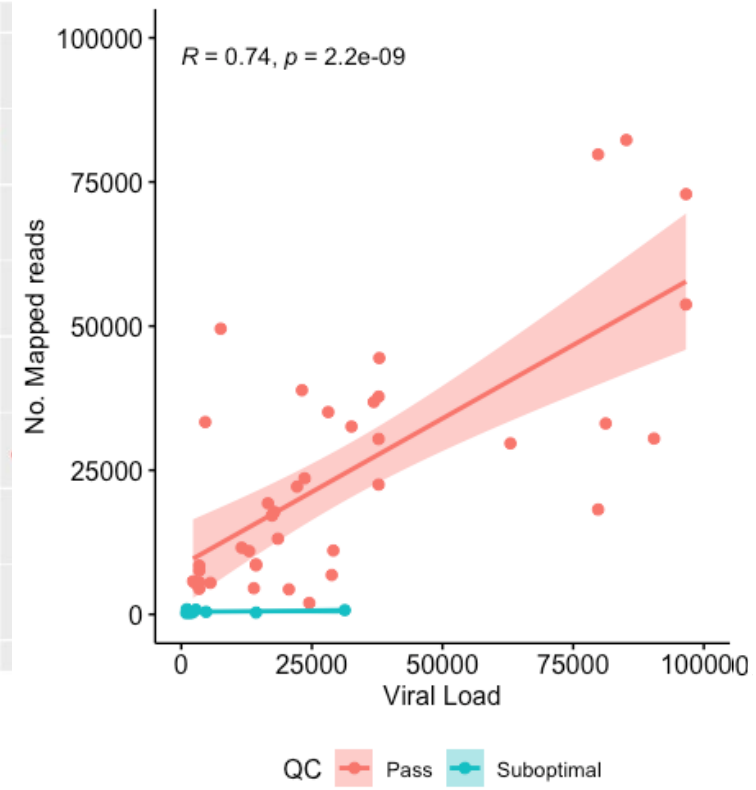


Figure 1A& 1B. Scatterplot of viral load and number of mapped reads for sequenced isolates (1A) and correlation and prediction curve (1B).