



**HAL**  
open science

## **Optimization of Next-Generation Sequencing Technologies for von Hippel Lindau (VHL) Mosaic Mutation Detection and Development of Confirmation Methods**

Lucie Coppin, Pascal Plouvier, Michel Crépin, Anne-Sophie Jourdain, Emilie Ait Yahya, Stéphane Richard, Brigitte Bressac-de Paillerets, Catherine Cardot-Bauters, Sophie Lejeune, Julie Leclerc, et al.

### ► To cite this version:

Lucie Coppin, Pascal Plouvier, Michel Crépin, Anne-Sophie Jourdain, Emilie Ait Yahya, et al.. Optimization of Next-Generation Sequencing Technologies for von Hippel Lindau (VHL) Mosaic Mutation Detection and Development of Confirmation Methods. *Journal of Molecular Diagnostics*, 2019, 21, pp.462 - 470. <10.1016/j.jmoldx.2019.01.005>. <hal-03486408>

**HAL Id: hal-03486408**

**<https://hal.science/hal-03486408v1>**

Submitted on 20 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

## **Optimization of next-generation sequencing technologies for von Hippel Lindau (VHL) mosaic mutation detection and development of confirmation methods**

Running title: *VHL* mosaic mutation screening

Lucie Coppin,<sup>\*†‡</sup> Pascal Plouvier,<sup>‡</sup> Michel Crépin,<sup>‡</sup> Anne-Sophie Jourdain,<sup>‡</sup> § Emilie Ait Yahya,<sup>‡</sup> Stéphane Richard,<sup>¶||</sup> Brigitte Bressac-de Paillerets,<sup>\*\*</sup> Catherine Cardot-Bauters,<sup>††</sup> Sophie Lejeune,<sup>‡‡</sup> Julie Leclerc,<sup>\*†‡</sup> and Pascal Pigny<sup>\*†‡</sup>

From UMR-S 1172 - JPARC - Jean-Pierre Aubert Research Center,<sup>\*</sup> and the Research Team on Rare and Developmental Metabolic Diseases (RADEME),<sup>§</sup> University Lille, Lille; Inserm,<sup>†</sup> UMR-S 1172, Team "Mucins, Epithelial Differentiation and Carcinogenesis", Lille; the Departments of Biochemistry and Molecular Biology,<sup>‡</sup> Endocrinology,<sup>††</sup> and Clinical Genetics,<sup>‡‡</sup> CHU Lille, Lille; Genetical Oncology,<sup>¶||</sup> Ecole Pratique des Hautes Etudes, PSL Research University, Inserm U1186, Gustave Roussy Cancer Campus, Villejuif; the National Expert Center for Rare Cancers PREDIR,<sup>||</sup> Bicêtre Hospital, Le Kremlin-Bicêtre; and the Biopathology Department,<sup>\*\*</sup> Inserm U1186, Gustave Roussy, Université Paris-Saclay, Villejuif, France

Correspondence to Pascal Pigny

E-Mail: [pascal.pigny@chru-lille.fr](mailto:pascal.pigny@chru-lille.fr)

**Funding:** Supported by CHRU de Lille.

**Disclosures:** None declared.

## **Abstract**

Von Hippel Lindau disease (VHL) is a monogenic disorder characterized by the development of tumors affecting the central nervous system, kidney, pancreas, or adrenal glands, and due to germline mutations in the *VHL* tumor suppressor gene. About 5% of patients harboring a typical VHL phenotype have no mutation detected by conventional techniques, so a postzygotic *VHL* mosaicism can be suspected. The aim of this study was therefore to implement a next-generation sequencing (NGS) strategy for *VHL* mosaic mutation detection, including an optimization of the original Personal Genome Machine design by enrichment with oligonucleotides corresponding to amplicons with insufficient depth of coverage. Two complementary strategies were developed for the confirmation of mosaic mutations identified by NGS, SNaPshot for variants present at an allelic ratio greater than 5%, and droplet digital PCR for allelic ratio above 1%. *VHL* mutant plasmids were generated to assess *VHL* mosaic mutation detection in different exons and to set up an internal quality control that could be included in each run or regularly to validate the assay. This strategy was applied to 47 patients with a suggestive or clinical VHL disease and mosaic mutations were identified in 8.5% of patients. In conclusion, NGS technologies combined with SNaPshot or droplet digital PCR allow the detection and confirmation of mosaic mutations in a clinical laboratory setting.

## Introduction

von Hippel Lindau (VHL) disease is an inherited neoplasia syndrome characterized by the development of several benign and malignant tumors affecting multiple organs, including cerebellar, spinal, and retinal haemangioblastoma (Hb), renal clear cell carcinoma (RCC), pheochromocytoma/paraganglioma, pancreatic cysts, and pancreatic neuroendocrine tumors. The clinical diagnosis is based on the occurrence of two characteristic VHL-related tumors, including at least one Hb, or the presence of a single characteristic VHL-related tumor with a family history of VHL (1,2). VHL disease is the result of a germline mutation inactivating the *VHL* tumor suppressor gene (NM\_000551.3, 3p25-26) and that is inherited in an autosomal dominant manner. In 20% of cases, VHL disease occurs without family history, in individuals with seemingly sporadic VHL-related tumors, and a *de novo* germline mutational event in the *VHL* gene is identified in these patients. The majority of *VHL* pathogenic variants are private and the spectrum is very large (about 200 distinct mutations reported to date), including missense and non-sense variants, frameshifts, splice site variants (60% to 80%), as well as small deletions or insertions and large or complete gene deletions (20% to 40%) (3). Conventional *VHL* genetic testing, based on DNA sequencing and PCR-based exonic rearrangement screening, allows the detection of a *VHL* pathogenic variant in approximately 95% of patients with a typical VHL disease (4). In the remaining 5% of patients harboring a typical VHL phenotype, a postzygotic *VHL* mosaicism can be suspected because of lack of sensitivity of conventional genetic testing (5). However, prior to the development of next-generation sequencing (NGS) technologies (6), mosaic mutation screening was cumbersome to perform routinely in a clinical laboratory setting. NGS provides higher sensitivity than conventional Sanger sequencing through a great depth of coverage that enables the detection of genetic variants in cell subpopulations (7-9). The identification of a mosaic mutation

present at a 1% allelic ratio with a good statistical significance by NGS requires a minimal depth of 700 reads with equilibrium between the forward and reverse strands in order to distinguish the mutation from a sequencing artefact (10). A few studies have reported levels of *VHL* mosaic mutation around 5% and have demonstrated that this kind of event was especially detected in patients with mild phenotype (5,11-13).

We previously described *VHL* mosaic mutation detection by 454 sequencing technology (Roche), and demonstrated that mosaicism in VHL disease is not restricted to patients with a mild phenotype (14). The aim of this study was i) to evaluate the efficiency of the Ion Torrent Personal Genome Machine (PGM) NGS-platform (based on semiconductor sequencing) for *VHL* mosaic mutation detection, ii) to develop confirmation methods for low allelic ratio mutations (droplet digital PCR (ddPCR) and SNaPshot), iii) to set up an internal quality control (iQc) adapted to mosaic mutation screening, and iv) to screen for *VHL* mosaic mutation the largest cohort (n=47) of VHL patients ever reported. So, we propose an overall technical protocol for *VHL* mosaic detection that can be implemented in a clinical laboratory setting.

## **Materials and Methods**

### **PATIENT SAMPLES**

Forty-seven patients with a phenotype evocative of a von Hippel Lindau disease and previously tested negative for heterozygous *VHL* mutation were screened for *VHL* mosaic mutation. Forty-six cases were sporadic (no familial history) and one case was familial (the patient had an affected daughter). Among these 47 patients, two DNA samples previously identified with *VHL* mosaic mutations were included as positive controls (P1: c.481C>T,

p.(Arg161\*); P2: c.500G>A, p.(Arg167Gln)) (14). DNA was extracted from peripheral blood mononuclear cells (PBMC) with the EZ1 DNA blood kit (Qiagen, Courtaboeuf, France). Tumor DNA was isolated from formalin-fixed, paraffin-embedded (FFPE) tumor tissues using QIAamp DNA FFPE Tissue kit (Qiagen) according to the manufacturer's instructions. All patients gave an informed consent allowing germline DNA analysis.

#### **NEXT-GENERATION SEQUENCING**

High-throughput sequencing was performed using PGM platform (Ion Torrent) and primers for amplification were designed using the "Ion AmpliSeq Designer" software v2.2.1 (coding regions plus 25 bp corresponding to intronic flanking regions) (Thermo Fisher Scientific Courtaboeuf, France). This design was part of a multigene panel previously designed to routinely perform the genetic testing of familial endocrine neoplasia syndromes. Primer pairs were split by the manufacturer into two mixes according to the compatibility between the different primers, so that primers for overlapping amplicons were in different tubes. Sequences of the primer pairs separated into two pools were the following : exon1 : amplicon#1710843567 (213pb) forward 5'-GCCTCCGGCCGGCTATT-3' and reverse 5'-CCCCGCCGTCTTCTTCA-3'; amplicon#1809669404 (198pb) forward 5'-AGGCAGGCGTCGAAGAGTA-3' and reverse 5'-GCTCGCCGTCGAAGTTG-3'; amplicon#1512083324 (114pb) forward 5'-GTGCTGCCCGTATGGCT-3'and reverse 5'-GGGCTTCAGACCGTGCTATC-3'; exon 2 : amplicon#5141954337 (250pb) forward 5'-CGGTGTGGCTCTTTAACAACC-3' and reverse 5'-GGGCTTAATTTTTCAAGTGGTCTATCCT-3'; exon 3 : amplicon#5364451121 (218pb) forward 5'-CCCTAGTCTGCCACTGAGGAT-3' and reverse 5'-ACCATCAAAGCTGAGATGAAACAGT-3'. The *in silico* coverage score predicted for this design was 100%. Libraries were constructed using the Ion AmpliSeq Panels, Ion AmpliSeq Library Kit, Ion Xpress Barcodes (for multiplexing), and 10 ng of patient DNA per

pool. Amount of DNA for *VHL* pCR2.1 constructions (wild-type and mutated) was adapted as described below. The libraries were transferred to the Ion OneTouch 2 System for emulsion PCR. Libraries were sequenced on the Ion PGM Sequencer (314 Chip for eight samples or 316 Chip for 16 samples) and automated data analysis was performed with SeqNext Module V4.1.2 of Sequencing Pilot Software (JSI Medical Systems GmbH, Ettenheim, Germany) with parameters allowing the detection of variants with a minimum of 1% of reads on both directions.

### **COUNT OF MOTIFS AND STATISTICS**

An in-house script was used to specifically analyze mutation loci. In this script, reference or alternate sequence plus anchor bases on both 5' and 3' of each mutation were used as motifs. Those motifs were searched and counted in raw read sequences from fastq files of all samples. It was assumed that no more than one patient sample per run could be positive for each variant. For each tested variant, candidate sample was compared to other samples—considered as controls—using pairwise Fisher tests on 2x2 contingency tables. This allowed systematic comparison in the worst scenario for background noise (cross-sample contamination, sequencing errors, etc.). A mutation was considered as detected when Fisher test P-value was inferior to 0.001 in all comparisons for a given candidate.

### **PLASMIDS**

#### *Preparation of plasmids*

Several plasmid harboring different *VHL* mutations were generated. The first step consisted of the creation of *VHL* plasmids for each *VHL* exon. Amplicons corresponding to exon 1, 2, and 3 of *VHL* were generated from genomic DNA amplified with the following primers : Exon 1 forward 5'-TGGTCTGGATCGCGGAGG-3' and reverse 5'-

CCCGTCTGCAAAATGGACCC-3' (450pb length); Exon 2 forward 5'-ACCGGTGTGGCTCTTTAACA-3' and reverse 5'-GCCCAAAGTGCTTTTGAGAC-3' (376pb); Exon 3 forward 5'-TGGCAAAGCCTCTTGTTTCG-3' and reverse 5'-GCCCTAAACATCACAATGCC-3'(433pb). Proportions of the insert cloned into pCR2.1 were respectively 10.28%, 8.73%, and 9.92% for exon 1, 2, and 3. Primers were designed so that AmpliSeq primers for NGS sequencing can hybridize onto these amplicons to enable their use in subsequent library preparation. Then, each amplicon was cloned into a pCR2.1 vector (3.929 kb) with the original TA cloning kit (Thermo Fisher Scientific) and plasmids were prepared using the NucleoBond Xtra Maxi Plus kit (Macherey Nagel, Hoerd, France).

Then, starting from these wild-type plasmids, *VHL* mutant plasmids were generated. They corresponded to eight *VHL* pathogenic variants distributed across the whole *VHL* gene: three variants in exon 1: c.194C>G, p.(Ser65Trp); c.226\_228del, p.(Phe76del); c.232A>C, p.(Asn78His); one in intron 1-exon 2: c.341-1G>C, p.?; and four in exon 3: c.467A>G, p.(Tyr156Cys); c.481C>T, p.(Arg161\*); c.482G>A, p.(Arg161Gln); c.500G>A, p.(Arg167Gln). Site-directed mutagenesis was performed with QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara CA). Primers were as follows: VHL1-p.(Ser6Trp) forward 5'-CCGTGCTGCGCTGGGTGAACTCGCG-3' and reverse 5'-CGCGAGTTCACCCAGCGCAGCACGG-3' ; VHL1-p.(Phe76del) forward 5'-CCCAGGTCATCTGCAATCGCAGTCC-3' and reverse 5'-GGACTGCGATTGCAGATGACCTGGG-3' ; VHL1-p.(Asn78His) forward 5'-GTCATCTTCTGCCATCGCAGTCCGC-3' and reverse 5'-GCGGACTGCGATGGCAGAAGATGAC-3' ; VHL2-c.341-1G>C forward 5'-GCTTGTCCCGATACGTCACCTTTGGC-3' and reverse 5'-GCCAAAGGTGACGTATCGGGACAAGC-3' ; VHL3-p.(Tyr156Cys) forward 5'-

GCCCTTCCAGTGTGTACTCTGAAAGAGC-3' and reverse 5'-  
GCTCTTTCAGAGTACACACTGGAAGGGC-3' ; VHL3-p.(Arg161\*) forward 5'-  
GTATACTCTGAAAGAGTGATGCCTCCAGGTTG-3' and reverse 5'-  
CAACCTGGAGGCATCACTCTTTCAGAGTATAC-3' ; VHL3-p.(Arg161Gln) forward 5'-  
GTATACTCTGAAAGAGCAATGCCTCCAGGTTGTC-3' and reverse 5'-  
GACAACCTGGAGGCATTGCTCTTTCAGAGTATAC-3' ; VHL3-p.(Arg167Gln) forward  
5'-CTCCAGGTTGTCCAGAGCCTAGTCAAG-3' and reverse 5'-  
CTTGACTAGGCTCTGGACAACCTGGAG-3' The presence of each mutation was checked  
by Sanger sequencing.

#### *Library construction for plasmids*

Each mutated plasmid was mixed with the wild-type one of the corresponding exon to obtain an allelic ratio between 1% and 2% and then the wild-type plasmids of the two other exons were added. Starting amount of DNA had to be modified for library preparation with *VHL* amplicons cloned into plasmids. A total of 2 pg of insert was used for the three constructions (which corresponds to 0.66 pg per insert). Taking into account the ratio insert size/plasmid vector size, the insert amount (pg) was converted into plasmid amount (pg). 6.42 pg of pCR2.1-VHL exon 1 construct, 7.52 pg of pCR2.1-VHL exon 2 construct, and 6.64 pg of pCR2.1-VHL exon 3 construct were pooled. According to the mutation individually studied (unitary plasmid), the mutated plasmid were diluted and pooled with the corresponding wild-type construction and the wild-type constructions for the two other exons were added. Then, an internal control (iQc) was generated which contains all the *VHL* mutations generated at the level of 1% to 2%. Each mutated *VHL* construct was mixed with the wild-type corresponding construct to get a final level of mosaicism around 1% to 2% for each studied position.

#### **CONFIRMATION METHODS**

### *SNaPshot*

SNaPshot Multiplex System technology (Thermo Fisher Scientific) was used as previously described (14). Primers used for *VHL* mutation detection are the following : VHL-1-65 forward 5'-CGCGGCCCGTGCTGCGCT-3' and reverse 5'-AGGGCTCGCGCGAGTTCACC-3' ; VHL-1-76 forward WT 5'-\*N\*N\*AGCCCTCCCAGGTCATCTT-3', forward MUT 5'-\*N\*N\*NCCCTCCCAGGTCATCTGC-3' and VHL-1-76 reverse 5'-\*N\*N\*NGCGGACTGCGATTGCAGA-3'; VHL-1-78 forward 5'-\*N\*N\*N\*CCTCCCAGGTCATCTTCTGC-3' and reverse 5'-\*N\*N\*N\*N\*CACGACGCGCGGACTGCGAT-3'; VHL-3-156 forward 5'-\*N\*N\*N\*N\*TTGGTTTTTGCCCTTCCAGTGT-3' and reverse 5'-\*N\*N\*N\*N\*GAGGCATCGCTCTTTCAGAGTA-3'; VHL-1-c.341 forward 5'-AACCTTTGCTTGTCCCGATA-3' and reverse 5'-TCTGAAGAGCCAAAGGTGAC-3'; VHL-1-161Gln forward 5'-\*NAGTGTATACTCTGAAAGAGC-3' and reverse 5'-\*NTCCGGACAACCTGGAGGCAT-3'; VHL-1-167Gln forward 5'-\*N\*N\*NGATGCCTCCAGGTTGTCC-3' and reverse 5'-\*N\*N\*NTCTCAGGCTTGACTIONAGGCTC-3'; VHL-3-161\* forward 5'-\*N\*N\*N\*NCCAGTGTATACTCTGAAAGAG-3' and reverse 5'-\*N\*N\*N\*NCCGGACAACCTGGAGGCATC-3'. For *VHL* c.226\_228del, p.(Phe76del) mutation, specific forward primers targeting the wild-type and the mutated event were designed.

### *ddCPR*

Droplet digital PCR (ddPCR) was performed based on water-oil emulsion droplet technology. This technology requires a preliminary optimization step to determine the best annealing temperature. Temperatures ranging from 52°C to 60°C were tested and it was found that 56

°C can be used for all our assays. Sample partitioning was performed using the QX200 Droplet Generator (Bio-Rad, Marnes-la-Coquette, France), PCR amplification with the C1000 thermal Cycler with the following conditions : (95 °C: 10min; (94 °C: 30 s; 56 °C: 60 s) for 40 cycles; 98 °C: 10 min) and droplet reading with the Droplet Reader (QX 200), which provides absolute quantification in digital form. TaqMan probe-based assays were used (Bio-rad). Each mutated event was detected with a primer labeled with FAM fluorophore. The corresponding wild-type allele was detected with a primer labeled with HEX fluorophore. Reference of ddPCR Mutation Assay (BioRad) are the following:

p.(Ser65Trp):dHsaMDS389884862;	p.(Phe76del):dHsaMDS544200506;
p.(Asn78His):dHsaMDS8356884801;	c.341-1G>C: dHsaMDS618563711;
p.(Tyr156Cys):dHsaMDS456108094;	p.(Arg161*):dHsaLS2503094;
p.(Arg161Gln):dHsaMDS649820635;	p.(Arg167Gln):dHsaLS2504838.

## Results

### ***VHL* MOSAIC MUTATION DETECTION USING THE PGM PLATFORM**

#### *Coverage*

*In silico* primers were designed to perform sequencing of the three exons of *VHL* gene with the PGM platform using the “Ion AmpliSeq Designer” software v 2.2.1. The three exons of *VHL* were covered by a total of five amplicons (three in exon 1 (Supplemental Figure S1), one in exon 2, and one in exon 3).

A preliminary NGS run allowed identification of amplicons with a depth of coverage below 700 reads that would not allow mosaic detection at the level of 1% (1X concentration; Supplemental Table S1). This was the case for two of the three amplicons covering *VHL* exon 1 (#1710843567 and #1809669404). These two amplicons are longer (213 bp and 198 bp,

respectively) than amplicon #1512083324 (114 bp), which was also located in exon 1 and well covered. Amplicons in *VHL* exon 2 and 3 were well covered. To obtain a higher depth of coverage for these two *VHL*-exon 1 amplicons, the primer mixes provided by the manufacturer were enriched with individual primers available separately in a 384-well plate (+1.25X and +2.5X enrichment). The same DNA samples were then reanalyzed in a new library prepared with these enriched mixes. According to the results (Supplemental Table S1), +1.25X and +2.5X enrichments successfully enabled to reach a minimal coverage of 700X for amplicons #1809669404 and #1710843567, respectively.

### *Sensitivity*

The results generated by the 454 and PGM platforms for two positive control samples were compared using the same settings in the SeqPilot software with the SeqNext module. The mosaic mutations were properly identified in both samples regardless of the platform used, with a concordant allelic ratio of mutated alleles: for patient 1, the ratio of mutated alleles was 2.1% with the PGM platform versus 1.7% with 454 (patient P1; Supplemental Table S2) and for patient 2 the ratio was 7.4% with PGM versus 5.7% with 454 (patient P2; Supplemental Table S2). Mosaic events detected by PGM were equilibrated on both strands. These results demonstrate that PGM is able to detect *VHL* mosaic mutations as well as 454 platform.

Mosaic mutation at other positions on *VHL* gene was next detected. Eight plasmids, each harboring a specific *VHL* mosaic mutation at level comprised between 1% and 2%, were generated and analyzed on PGM platform. These eight mutations were distributed on the whole *VHL* gene. Individual mosaic mutation in one exon was systematically analyzed in the presence of the two other plasmids corresponding to the wild-type form of the remaining exons in order to be closer to real condition of *VHL* patient analyses. Results presented in

table 1 (unitary plasmid) and table 2 showed that mosaic mutations are found at levels corresponding to the theoretical dilution.

#### *Construction of an internal quality control*

Next, by mixing the eight mutated plasmids with the three wild-type plasmids, a 1% mosaic mutation positive sample was generated, which simultaneously harbored the eight *VHL* mutations previously studied, each at a level comprised between 1% and 2%. Results obtained on PGM platform are shown in table 1 (mixed plasmids). The P-values reflect the probability that the *VHL* event detected is significantly different from background noise. All mutations were detected in the sample at the expected level with a significant *P*-value. This positive sample, after evaluation, could be used as an iQc during NGS run to validate the assay, allowing assessing both the detection and the detection level of each *VHL* mosaic mutation.

#### *Clinical testing*

Twelve patients were analyzed with the PGM platform. Taking account the 35 patients previously analyzed with the 454 platform, it represents a total of 47 patients tested negative for heterozygous *VHL* germline mutations who were analyzed for *VHL* mosaic mutation (Table 3 and Supplemental Table S3). Twenty-three patients had a suggestive VHL disease (one tumor) and 24 presented a typical VHL disease (>2 tumors or one tumor with a familial context). Their clinical characteristics are described in Table 3 and Supplemental Table S3. Forty-six patients had a sporadic VHL disease and one patient had an affected offspring. A *VHL* mosaic mutation was identified in four patients (4/47 : 8.5%). Focusing on patients with a typical VHL disease, which corresponds to a *VHL* mosaicism frequency of 16.6 % (4/24). The identified events are c.500G>A p.(Arg167Gln) (allelic ratio: 5.7%), c.481C>T

p.(Arg161\*) in two unrelated patients (same allelic ratio: 1.7%) and c.490C>T p.(Gln164\*) (3.3%). One of them corresponds to the familial case of the cohort and his daughter was heterozygous for the same mutation. No *VHL* mosaic mutation was identified in the other patients at the detection level of 1%. The four patients harboring a *VHL* mosaic mutation had a clinical *VHL* disease with a median number of lesion of 2.5 which is significantly higher than patients without a mosaic mutation (median 1.0,  $P = 0.0058$ , Mann and Whitney test). Interestingly, all patients with a unique Hb were negative for *VHL* mosaic mutation.

### **CONFIRMATION METHODS**

In molecular diagnosis practice, the presence of a germline mutation should be validated before reporting the result to the physician who prescribed the molecular test. The confirmation of a mosaic event detected by NGS is possible either with another NGS technology or with another highly sensitive technique such as SNaPshot or droplet digital PCR (ddPCR).

To evaluate these two techniques for the detection of mosaic mutation, each mutant plasmid was used to determine the limit of detection. SNaPshot was able to detect mutation until a threshold set at 5% (Table 4 and Figure 1A). The ddPCR technique was able to detect an event with a threshold of 1% (Table 4 and Figure 1B). For c.194C>G, p.(Ser65Trp) *VHL* mutation, no signal was detected by ddPCR with the wild-type probe, maybe due to technical issues during probe production.

Then, these confirmation methods were evaluated on patient samples harboring mosaic mutations with a level of mutated alleles around the two defined thresholds. First, the sample was tested with the c.500G>A, p.(Arg167Gln) *VHL* mosaic mutation at 5.7% as assessed by NGS. This mutation was successfully detected both by SNaPshot (14) and ddPCR which yielded a level of mutated events of 5.5% (Table 3 and Supplemental Table S4). The

c.490C>T, p.(Gln164\*) mosaic mutation identified at the level of 3.3% by NGS was also confirmed by ddPCR. The proportion of mutated allele detected by ddPCR was 3.5 %. The sample was tested with the c.481C>T, p.(Arg161\*) *VHL* mosaic mutation at 1.7% as assessed by NGS. Although allelic frequency of this mutation in PBMC was too low to be detected by SNaPshot, the mutation was correctly detected by ddPCR and the level of mutated events was concordant with that determined by NGS (1.8%, Table 3 and Supplemental Table S4). The second sample with the c.481C>T, p.(Arg161\*) *VHL* mosaic mutation assessed by NGS at 1.7% was confirmed by ddPCR at the level of 1.9% (Table 3 and Supplemental Table S4). So there is a strong correlation between the percentages of mutated alleles detected by ddPCR and NGS which underline the accuracy of these two methods. These results demonstrate that both ddPCR and SNaPshot can be used to confirm the presence of mosaic mutations detected by NGS and that the choice between these two techniques depends on the level of mosaicism: SNaPshot is preferred when the level of mutated allele is higher than 5% (as it is less expensive) and ddPCR is preferred for mutated allele between 1% and 5 % (as it is more sensitive).

Tumoral tissue was available for two patients out of the four with a *VHL* mosaic mutation (a pheochromocytoma and a RCC tumor). These tumor DNAs were analyzed by ddPCR. Concerning the patient with the c.500G>A, p.(Arg167Gln) *VHL* mutation (5.5% of mutated alleles in blood, ddPCR), in his pheochromocytoma tumor tissue the same mutation was identified at the level of 34.6%. There were 81% of tumor cells in it. So the expected ratio of mutated alleles was 40.5% which is close to the mutated ratio obtained by ddPCR. Right RCC tumor was available for the patient with the c.481 C>T, p.(Arg161\*) mutation (1.8% of mutated alleles in blood, ddPCR). The same mutation was also identified at the level of 11.1% (Table 3). There were 54% of tumor cells in it. So the expected ratio of mutated alleles was 27%. The lower level of mutated alleles detected in tumoral tissue could be due to the

difficulty to assess the level of tumor cells in samples. So in the two tumoral tissues available, the presence of the mutation was validated at a higher allelic ratio than the one detected in blood.

## **Discussion**

Mosaicism is defined as the presence of at least two populations of cells with distinct genotypes within the same individual or tissue. Depending on the timing of occurrence during embryonic development, somatic or post-zygotic mosaic mutations will be present at various levels in blood and their detection can be challenging, thus explaining that the incidence of mosaicism in various inherited genetic diseases is still poorly documented. Most of described mosaic disorders concerns autosomal dominant diseases (15). The identification of mosaicism has important consequences in genetic counseling. In this context, we previously developed a method on the pyrosequencing-based 454 Roche Junior sequencing technology, which provided a high coverage of the three exons of *VHL* (14) for eight samples tested in a single run. *VHL* mosaic mutations are not always linked to a mild phenotype. However, 454 sequencing technology was withdrawn from the market at the end of 2016. So, another strategy was set up to detect *VHL* mosaic mutations. Here, we describe *i*) the analytical evaluation of the PGM platform, which detects pH variation following the incorporation of the nucleotide during sequencing (16), for the detection of *VHL* mosaic mutations, *ii*) the creation of mutant plasmids, harboring *VHL* mutations distributed on the whole *VHL* gene and of different types (non-sense, missense, small deletions, splice mutations), which could be mixed and used routinely as an iQc, *iii*) the results of the largest population of VHL patients (n=47) screened for *VHL* mosaic by NGS, and *iv*) the evaluation of confirmation methods for the detection of *VHL* mosaic mutations.

PGM specificity and sensitivity have been evaluated in the literature and PGM can be used as an alternative to Sanger sequencing to detect mutations (17). NGS is recognized to provide reliable results, the confirmation by Sanger sequencing of variants identified by NGS being controversial in the literature and mostly useful to detect potential sample mix-up during DNA extraction (18, 19). Moreover, several works demonstrated the usefulness of NGS to detect mosaic mutation which are not detectable by Sanger sequencing due to the lack of sensitivity of this latter technique (20).

Here, PGM technology was tested to detect mosaic mutations. The results demonstrated that PGM platform is a valuable alternative to 454 platform to detect mosaic mutations. Initial PGM design did not provide a sufficient depth of coverage of all *VHL* exons. So, enrichment was performed that allowed achieving a 100% coverage with a minimal depth of coverage of 700 reads per exon, as this has been described by others to be the minimal depth required to detect mosaic mutation at a level of 1% (10). The comparison of 454 and PGM platforms for mosaic detection reveals that both technologies are highly sensitive. Eight patients could be analyzed per run on the 454 platform to achieve 700 reads per exon and detect mosaic mutation at a level of 1%. PGM platform has a greater capacity of sequencing and so is more adapted to the throughput of routine diagnosis (up to 16 patients analyzed in a run with a 316 Chip).

Plasmids were generated with *VHL* mutants and dilutions performed at a level comprised between 1% and 2% of mutated events. These mosaic mutations were all detected by PGM platform. Moreover, they were all detected at the expected levels, thus highlighting that this technique has the capacity to detect mosaic mutation at level comprised between 1% and 2%. PGM platform was able to detect mosaic mutations corresponding to different kinds of events: missense, non-sense, small deletions, and splice mutations.

Currently, no iQc dedicated to *VHL* gene is available to validate NGS for routine diagnosis, especially for mosaic mutation detection. So, created a positive control that could be regularly included in run to validate the assay both on the capacity of mosaic mutation detection and on the sensitivity of detection. All the mutant plasmids were mixed so that each mutation was present at a level of 1% to 2%. Results confirm that all mutations were detected in this same sample. So this mix constitutes a valuable iQc that should be included in each run and checked before patient sample analysis. Moreover, as reported recently, the use of plasmids alone or in combination, with known mutations (position and proportion) is recommended by the Association for Molecular Pathology and College of American Pathologists in the last “guidelines for validation of NGS” to validate bioinformatics pipelines, to demonstrate detection of the variant in each run and to evaluate the assay performance (21). Use of mutated plasmids is also recommended to evaluate NGS technologies through cross-platform comparison (22), and the performance of assay and reagents. So iQc based on a mix of *VHL* mutations may be used for all these purposes.

Sanger sequencing is the reference method to detect mutation and is usually used to confirm events identified by NGS sequencing. However, its sensitivity is limited to 15% to 20% of mutated allele and thus is not adapted to mosaic mutation detection and confirmation. New mutation quantification tools, such as ddPCR, have been developed. ddPCR is based on allelic discrimination with one probe specific to the mutation and a second probe specific to the wild-type sequence. These two probes are labeled with different fluorophores. As this technique is event-specific, it requires a different set of probes for each mutation and is quite expensive. However it enables quantification of the mutated allele. An alternative tool is SNaPshot, a semi-quantitative primer extension–based method, but its sensitivity is lower than that of ddPCR (23-24). This technique also requires different set of primers for each

mutation but is less expensive than ddPCR. ddPCR is very sensitive, but it requires many manual steps in comparison with other techniques. Moreover, ddPCR requires many controls, especially to detect the absence of mutation because the possibility of false-positive cases is high due to contamination or non-target amplification (25). Here, we tested confirmation techniques to validate the reliability of mosaic mutations detected by NGS. The results demonstrate that SNaPshot is a good strategy to confirm mosaic mutations detected at a level above 5%. ddPCR is able to detect mutations up to 1% but requires set up adaptation in each laboratory with positive controls to adapt experimental conditions and detect events at the expected threshold.

The largest cohort of *VHL* patients was screened for mosaic mutation detection (n=47) and 8.5% patients were found with a *VHL* mosaic mutation among these cases with no history of *VHL* in the ancestors. This frequency is in agreement with literature data which indicate a proportion around 5% for *VHL* mosaic mutations (5). Moreover, a frequency of *VHL* mosaicism of 16.6% was identified among patients with a typical *VHL* disease and negative for *VHL* heterozygous mutation. This high proportion justifies the investigation of *VHL* mosaicism in patients with a typical *VHL* disease when no heterozygous mutation could be identified. The patients with mosaic mutations had more lesions than those without. This result confirms that mosaic mutations are not associated with a mild phenotype (14). Since *VHL* mosaic mutation was not identified in patients with a unique Hb, these data also suggest that this analysis is not useful for patients harboring this unique lesion. The identification in the father's DNA of the *VHL* mutation present in his affected daughter at a heterozygous state further supports that mosaicism is a true finding. Moreover, the matching between tumor DNA and blood DNA analyses for the two cases with available tumor tissues is another element in favor of the reality of mosaicism.

In conclusion, *VHL* mosaic mutations can be detected using the PGM platform in a molecular genetics laboratory in a clinical setting. ddPCR seems to be the more suitable method to confirm *VHL* mosaic mutations identified by NGS, as the limit of detection of the two techniques are similar. However, SNaPshot is a good and a less expensive alternative to confirm mosaic mutation at a level above 5%. The identification and confirmation of *VHL* mosaic mutations is of great value in improving patient's care and providing genetic counseling to the family, as these mutations can be transmitted to the offspring.

### **Acknowledgments**

We thank our clinical collaborators who referred their patients to us for molecular evaluation, and Wassila Igoudjil and Jerome Gudewicz for their expert technical help.

## References

1. Gossage L, Eisen T, Maher ER: VHL, the story of a tumour suppressor gene. *Nat Rev Cancer* 2015, 1:55-64
2. Haddad, NM, Cavallerano JD, Silva PS: Von hippel-lindau disease: a genetic and clinical review. *Semin Ophthalmol* 2013, 28:377-86
3. Decker J, Neuhaus C, Macdonald F, Brauch H, Maher ER: Clinical utility gene card for: von Hippel-Lindau (VHL). *Eur J Hum Genet* 2014, 22
4. Maher ER, Neumann HP, Richard S: Von Hippel-Lindau disease: a clinical and scientific review. *Eur J Hum Genet* 2011, 19:617-23
5. Sgambati MT, Stolle C, Choyke PL, Walther MM, Zbar B, Linehan WM, Glenn GM: Mosaicism in von Hippel-Lindau disease: lessons from kindreds with germline mutations identified in offspring with mosaic parents. *Am J Hum Genet* 2000, 66:84-91
6. Mardis ER: The impact of next-generation sequencing technology on genetics. *Trends Genet* 2008, 24:133-41
7. Pagnamenta AT, Lise S, Harrison V, Stewart H, Jayawant S, Quaghebeur G, Deng AT, Murphy VE, Sadighi Akha E, Rimmer A, Mathieson I, Knight SJ, Kini U, Taylor JC, Keays DA: Exome sequencing can detect pathogenic mosaic mutations present at low allele frequencies. *J Hum Genet* 2012, 57: 70-72
8. Qin L, Wang J, Tian X, Yu H, Truong C, Mitchell JJ, Wierenga KJ, Craigen WJ, Zhang VW, Wong LC: Detection and Quantification of Mosaic Mutations in Disease Genes by Next-Generation Sequencing. *J Mol Diagn* 2016, 18:446-53

9. Page K, Guttery DS, Fernandez-Garcia D, Hills A, Hastings RK, Luo J, Goddard K, Shahin V, Woodley-Barker L, Rosales BM, Coombes RC, Stebbing J, Shaw JA: Next Generation Sequencing of Circulating Cell-Free DNA for Evaluating Mutations and Gene Amplification in Metastatic Breast Cancer. *Clin Chem* 2017, 63:532-41
10. Izawa K, Hijikata A, Tanaka N, Kawai T, Saito MK, Goldbach-Mansky R, Aksentijevich I, Yasumi T, Nakahata T, Heike T, Nishikomori R, Ohara O: Detection of base substitution-type somatic mosaicism of the NLRP3 gene with >99.9% statistical confidence by massively parallel sequencing. *DNA Res* 2012, 19: 142-52
11. Murgia A, Martella M, Vinanzi C, Polli R, Perilongo G, Opocher G. Somatic mosaicism in von Hippel-Lindau Disease. *Hum Mutat.* 2000;15:114
12. Santarpia L, Sarlis NJ, Santarpia M, Sherman SI, Trimarchi F, Benvenga S. Mosaicism in von Hippel-Lindau disease: an event important to recognize. *J Cell Mol Med.* 2007;11:1408–15
13. Wu P, Zhang N, Wang X, Li T, Ning X, Bu D, Gong K. Mosaicism in von Hippel-Lindau disease with severe renal manifestations. *Clin Genet.* 2013;84:581–4
14. Coppin L, Grutzmacher C, Crépin M, Destailleur E, Giraud S, Cardot-Bauters C, Porchet N, Pigny P: VHL mosaicism can be detected by clinical next-generation sequencing and is not restricted to patients with a mild phenotype. *Eur J Hum Genet* 2014, 22:1149-52
15. Biesecker LG, Spinner NB: A genomic view of mosaicism and human disease. *Nat Rev Genet* 2013, 14:307-20
16. Knief C: Analysis of plant microbe interactions in the era of next generation sequencing technologies. *Front Plant Sci* 2014, 5:216

17. Diekstra A, Bosgoed E, Rikken A, van Lier B, Kamsteeg EJ, Tychon M, Derks RC, van Soest RA, Mensenkamp AR, Scheffer H, Neveling K, Nelen MR: Translating sanger-based routine DNA diagnostics into generic massive parallel ion semiconductor sequencing. *Clin Chem* 2015, 61:154-62
18. Beck TF, Mullikin JC, NISC Comparative Sequencing Program, Biesecker LG: Systematic Evaluation of Sanger Validation of Next-Generation Sequencing Variants. *Clin Chem* 2016, 62:647-54
19. Mu W, Lu HM, Chen J, Li S, Elliott AM: Sanger Confirmation Is Required to Achieve Optimal Sensitivity and Specificity in Next-Generation Sequencing Panel Testing. *J Mol Diagn* 2016, 18:923-32
20. Yamaguchi K, Komura M, Yamaguchi R, Imoto S, Shimizu E, Kasuya S, Shibuya T, Hatakeyama S, Takahashi N, Ikenoue T, Hata K, Tsurita G, Shinozaki M, Suzuki Y, Sugano S, Miyano S, Furukawa Y: Detection of APC mosaicism by next-generation sequencing in an FAP patient. *J Hum Genet* 2015, 60: 227-31
21. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, Temple-Smolkin RL, Voelkerding KV, Nikiforova MN: Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. *J Mol Diagn* 2017, 19:341-65
22. Sims DJ, Harrington RD, Polley EC, Forbes TD, Mehaffey MG, McGregor PM 3rd, Camalier CE, Harper KN, Bouk CH, Das B, Conley BA, Doroshov JH, Williams PM, Lih CJ: Plasmid-based materials as multiplex quality controls and calibrators for clinical next-generation sequencing assays. *J Mol Diagn* 2016, 18:336e349

23. Magnin S, Viel E, Baraquin A, Valmary-Degano S, Kantelip B, Pretet JL, Mouglin C, Bigand M, Girardo B, Borg C, Ferrand C: A multiplex SNaPshot assay as a rapid method for detecting KRAS and BRAF mutations in advanced colorectal cancers. *J Mol Diagn* 2011, 13: 485-92
24. Lamy PJ, Castan F, Lozano N, Monté lion C, Audran P, Bibeau F, Roques S, Montels F, Laberrenne AC: Next-Generation Genotyping by Digital PCR to Detect and Quantify the BRAF V600E Mutation in Melanoma Biopsies. *J Mol Diagn* 2015, 17:366-73
25. Huggett JF, Cowen S, Foy CA: Considerations for digital PCR as an accurate molecular diagnostic tool. *Clin Chem* 2015, 61:79-88

### **Figure legend**

**Figure 1.** Limit of detection tests for confirmation techniques. **A:** Detection of the c.467A>G p.(Tyr156Cys) event by SNaPshot with a proportion of mutated plasmid at 100%, 50%, 25%, 10%, 5%, and 1%. **B:** detection of the c.341-1G>C event by ddPCR with a proportion of mutated plasmid at 100 %, 50%, 25%, 10%, 5%, 1%, and 0.1%.

**Table 1 : Proportion of mutated *VHL* event detected by NGS on the PGM platform for each individual mutant plasmid and for the mixed iQc (dilution 1% to 2%).**

Events	<b>c.194C&gt;G</b> <b>p.(ser65Trp)</b>	<b>c.226_228del</b> <b>p.(Phe76del)</b>	<b>c.232A&gt;C</b> <b>p.(Asn78His)</b>	<b>c.341-1G&gt;C</b>	<b>c.467A&gt;G</b> <b>p.(Tyr156Cys)</b>	<b>c.481C&gt;T</b> <b>p.(Arg161*)</b>	<b>c.482G&gt;A</b> <b>p.(Arg161Gln)</b>	<b>c.500G&gt;A</b> <b>p.(Arg167Gln)</b>
Proportion (%) of mutated event (unitary plasmid)	1.34	1.1	1.35	1.03	1.31	1.72	1.68	2.39
<i>P</i> -value	5.69x10 <sup>-18</sup>	1.20 x10 <sup>-17</sup>	1.09 x10 <sup>-20</sup>	6.06 x10 <sup>-15</sup>	4.73 x10 <sup>-23</sup>	4.11 x10 <sup>-23</sup>	1.05 x10 <sup>-23</sup>	2.81 x10 <sup>-48</sup>
Proportion(%) of mutated event (mixed plasmids (iQc))	1.41	2.3	1.54	1.41	1.85	1.73	1.12	1.02
<i>P</i> -value	2.79 x10 <sup>-17</sup>	8.63 x10 <sup>-17</sup>	1.99 x10 <sup>-10</sup>	9.21 x10 <sup>-18</sup>	6.07 x10 <sup>-33</sup>	2.30 x10 <sup>-23</sup>	4.21 x10 <sup>-16</sup>	4.10 x10 <sup>-15</sup>

“Unitary plasmid” corresponds to plasmid individually analyzed (each had a different barcode), “mixed plasmids” corresponds to all plasmid mixed all together in the same sample (all had the same barcode).

**Table 2: SeqNext analyses of mutant unitary plasmids.**

Event	Total read number (forward + reverse) for the position of the event	Number of read for each nucleotide (forward +reverse) at the position of the event				Proportio n of the event (%)
		A	T	C	G	
c.194C>G p.(Ser65Trp)	10848	58	14	10631	145	1.34%
c.226_228del p.(Phe76del)	15744	180 reads with the deletion				1.1%
c.232A>C p.(Asn78His)	17286	17021	3	233	29	1.35%
c.341-1G>C	6101	3	6	63	6029	1.03%
c.467A>G p.(Tyr156Cys)	18348	18107	0	1	240	1.31%
c.481C>T p.(Arg161*)	7213	1	124	7087	1	1.72%
c.482G>A p.(Arg161Gln)	6682	112	1	6569	0	1.68%

c.500G>A p.(Arg167Gln)	21848	523	3	19	21303	2.39%
------------------------	-------	-----	---	----	-------	-------

---

At each mutated position, the number of reads for each nucleotide is shown in the table.

**Table 3: Clinical characteristics of patients harboring a *VHL* mosaic mutation.**

Sex	Date of birth	Age of first clinical manifestation (years)	Familial history	Phenotype	Number of tumors	Clinical diagnosis of VHL	Event	% of mutated allele in blood identified by NGS	% of mutated allele in blood identified by ddPCR	% of mutated allele in tumor identified by ddPCR
M	1990	16	No	HB, pheochromocytoma, pancreatic endocrine tumor, left adrenal nodule	3	Yes	c.500G>A  p.(Arg167Gln)	5.7%	5.5%	34.6%  (pheochromocytoma)
M	1994	16	?	several retinal HB, renal cysts	2	Yes	c.481C>T  p.(Arg161*)	1.7%	1.9%	Tissue not available
M	1939	48	Yes†	bilateral RCC	2	Yes	c.481C>T  p.(Arg161*)	1.7%	1.8%	11.1%  (right RCC)
M	1984	34	No	central nervous system HB, two RCC	3	Yes	c.490C>T  p.(Gln164*)	3.3%	3.5%	Tissue not available 27

†An affected daughter.

HB: haemangioblastoma; RCC: renal clear cell carcinoma.

**Table 4: Proportion of *VHL* mutated event detected by SNaPshot and ddPCR for each individual plasmid mutant (ND: not detected).**

Theoretical proportion of mutated event	SNaPshot				ddPCR			
	10 %	5 %	1 %	Threshold	10 %	5 %	1 %	Threshold
Observed proportion of mutated event (Forward+reverse)	%	%	%	Threshold	%	%	%	Threshold
<b>c.194C&gt;G p.(Ser65Trp)</b>	17.7	8.43	ND	5%	ND	ND	ND	ND
<b>c.226_228del p.(Phe76del)</b>	7.8	3.37	ND	5%	6.38	2.95	0.69	1%
<b>c.232A&gt;C p.(Asn78His)</b>	20.0	12.9	5.5	<5%	4.71	2.04	0.053	1%
<b>c.341-1G&gt;C</b>	25.8	18.21	ND	5%	13.61	6.8	1.01	1%
<b>c.467A&gt;G p.(Tyr156Cys)</b>	12.1	7.02	ND	5%	21.29	7.99	1.19	1%
<b>c.481C&gt;T p.(Arg161*)</b>	13.2	ND	ND	>5%	24.65	9.98	2.89	1%
<b>c.482G&gt;A p.(Arg161Gln)</b>	13.3	5.1	ND	5%	15.19	8.73	2.32	1%
<b>c.500G&gt;A p.(Arg167Gln)</b>	7.9	ND	ND	>5%	21.03	7.48	4.36	1%

Figure 1

