Breast and ovarian cancer in Serbia: the importance of mutation detection in hereditary predisposition genes using NGS

dr sc. Ana Krivokuća
Laboratory for molecular genetics
Institute for Oncology and Radiology of Serbia
1977. Sanger = 1st GS
2003. Human genome sequenced

2005. 1st 454 Life Sciences NGS
2006. 1st Solexa NGS
2007. 1st Applied Biosystems NGS

2009. 1st Helicos single molecule sequencer
2011. 1st Pacific Biosciences single molecule sequencer
2012. 1st Oxford Nanopore Technologies ultra long single molecule reads

1st Gen Seq

2nd Gen Seq
Next Gen Seq

3rd Gen Seq
Next Next Gen Seq
<table>
<thead>
<tr>
<th>Coverage details</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA input</td>
<td>50ng</td>
</tr>
<tr>
<td>Number of probes</td>
<td>~ 4000</td>
</tr>
<tr>
<td>Probe size</td>
<td>80-mer/ NCBI37/hg19 reference genome</td>
</tr>
<tr>
<td>Number of target exons</td>
<td>&gt; 1700</td>
</tr>
<tr>
<td>Number of target genes</td>
<td>94 + 284SNPs</td>
</tr>
<tr>
<td>Cumulative target region size</td>
<td>255 kb</td>
</tr>
<tr>
<td>Recommended mean coverage</td>
<td>100x</td>
</tr>
<tr>
<td>Target minimum coverage</td>
<td>20x</td>
</tr>
<tr>
<td>Percent exons covered based on coverage metrics</td>
<td>≥ 95%</td>
</tr>
</tbody>
</table>
Sample collection

Cells and DNA isolation

DNA quality and quantity

NGS workflow

Results

Variant annotation

Gene lists

Primary analysis 24h

Sample Preparation

RBC Lysis

WBC Lysis

DNA Binding

Wash

Elution

Library prep (3 days)

Sequencing (2 days)
TruSight Cancer Panel – Library Prep

A. Sample Preparation

B. Denature double-stranded DNA library (for simplicity, adapters and indexes not shown)

C. Hybridize biotinylated probes to targeted regions

D. Enrichment using streptavidin beads

E. Elution from beads

Pooled Sample Library
- Adapters binding to the flow-cell surface
- Bridge amplification
- Cluster generation
- Sequencing by synthesis
DNA quality and quantity

NGS workflow

Sample collection

Cells and DNA isolation

DNA quality and quantity

Sample preparation

Library prep (3 days)

Primary analysis 24h

Gene lists

Variant annotation

Sequencing (2 days)

Results
NGS bioinformatics

RAW DATA → ALINGMENT → VARIANT CALLING → ANNOTATION

- Thousands of info
- Limited number of info

WES/WGS/RNASEq
TruSight Cancer Panel - Primary Analysis

- Demultiplexing and **FASTQ file** generation: Local Run Manager software/ BWA Enrichment BaseSpace App/ Galaxy/ MiSeq Reporter/ many other softwares
- Alignement and **BAM file** generation: Burrows-Wheeler Aligner (BWA)
- **Variant analysis**: The Genome Analysis Toolkit (GATK)/ VarScan/ VarDict/ SAMtools/ Beagle/ NGSEP/ Reveel ...
- **VCF file** generation
- **Summary files**: report on enrichment, variant calling, coverage, insert fragment length, and duplicates
FASTQ File Format Analysis

- Unique instrument name
- Flow cell ID
- Run ID
- Flowcell Lane
- Tile number within the flowcell lane
- X and Y coordinate within the tile
- The member of a pair, 1 or 2 (paired-end or mate-pair)
- Y if the read is filtered, N otherwise
- Index sequence

The "+" indicates the break between sequence and quality.

Encodes the quality values for the sequence letters

Raw sequence letters

0 when none of the control bits are on, otherwise it is an even number
Binary Alignment Map (BAM)
#fileFormat=VCFv4.0
#fileDate=2011-02-16
#source=GENSNP
#chromosome_id=131
#reference=GRC37.p1
#assocSNP_Typ=polypopulation in at least one population
#assocSNP_POP_ID=sex
#assocSNP_LOC_POP_ID=sex_SAMPLES

---

**Start of VCF Header**

---

**Start of Data table for this file**

---

**Reference allele**

---

**Format Statement for Population Frequency**

---

**Start of VCF File**

---

**POPFREQ for rs12121571**

---

### Before

<table>
<thead>
<tr>
<th>CHROM</th>
<th>POS</th>
<th>ID</th>
<th>REF</th>
<th>ALT</th>
<th>QUAL</th>
<th>FILTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>9999999</td>
<td>A</td>
<td>MLEAC=2; MLEAF=1.00; MQ=60.00; QD=25.82; SOR=5.136 GT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>72,39,0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10000017</td>
<td>C</td>
<td>T</td>
<td>262.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10000021</td>
<td>T</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10000043</td>
<td>T</td>
<td>G</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### After

<table>
<thead>
<tr>
<th>M</th>
<th>POS</th>
<th>QUAL</th>
<th>DP</th>
<th>NA12878.GQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>8</td>
<td>364.77</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>1222.73</td>
<td>29</td>
<td></td>
</tr>
</tbody>
</table>

---
Local database

Public mutation databases

Databases of unclassified variants

In silico predictions

Published data (in vivo, in vitro, in silico)

Variant annotation (1-5)
March 2016
Institute for Oncology Vojvodina
48 years, Dg. adenocarcinoma papillare serosum ovarii bilateralis, G2
Family history- unknown
NGS analysis

Info about variations

Functional prediction
- **BRCA1:c.5266dupC (p.Gln1756Profs)**
- **Frameshift**
- **Founder**
Confirmation

Sanger sequencing

Allele-specific PCR

Class 5- Pathogenic mutation
- BRCA1:c.5266dupC
  (p.Gln1756Profs)
- Frameshift
- Founder
Result: BRCA1:c.5266dupC (p.Gln1756Profs)
BRCA1, c.4356delA, p. (Ala1453Glnfs*3)

- December 2015
- Institute for Oncology and Radiology of Serbia
- 56 years, Dg. Adenocarcinoma serosum ovari
- Family history: yes
NGS analysis

- Info about variants
- Functional prediction
- Allele frequency
<table>
<thead>
<tr>
<th>Exon Type</th>
<th>Total Number of Entries</th>
<th>Distinct Mutations, Polymorphisms and Variants</th>
<th>Alterations reported only once</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2197</td>
<td>53</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>187</td>
<td>47</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>445</td>
<td>49</td>
<td>27</td>
</tr>
<tr>
<td>6</td>
<td>198</td>
<td>30</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>164</td>
<td>36</td>
<td>19</td>
</tr>
<tr>
<td>8</td>
<td>360</td>
<td>36</td>
<td>19</td>
</tr>
<tr>
<td>9</td>
<td>349</td>
<td>19</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>54</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11A</td>
<td>1251</td>
<td>224</td>
<td>115</td>
</tr>
<tr>
<td>11B</td>
<td>1613</td>
<td>223</td>
<td>123</td>
</tr>
<tr>
<td>11C</td>
<td>1835</td>
<td>235</td>
<td>124</td>
</tr>
<tr>
<td>11D</td>
<td>1673</td>
<td>221</td>
<td>116</td>
</tr>
<tr>
<td>12</td>
<td>118</td>
<td>47</td>
<td>28</td>
</tr>
<tr>
<td>13</td>
<td>655</td>
<td>51</td>
<td>28</td>
</tr>
<tr>
<td>14</td>
<td>111</td>
<td>28</td>
<td>19</td>
</tr>
<tr>
<td>15</td>
<td>241</td>
<td>47</td>
<td>24</td>
</tr>
<tr>
<td>16</td>
<td>790</td>
<td>88</td>
<td>52</td>
</tr>
<tr>
<td>17</td>
<td>578</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>18</td>
<td>466</td>
<td>55</td>
<td>26</td>
</tr>
<tr>
<td>19</td>
<td>123</td>
<td>32</td>
<td>18</td>
</tr>
<tr>
<td>20</td>
<td>1334</td>
<td>48</td>
<td>27</td>
</tr>
<tr>
<td>21</td>
<td>89</td>
<td>29</td>
<td>19</td>
</tr>
<tr>
<td>22</td>
<td>173</td>
<td>31</td>
<td>13</td>
</tr>
<tr>
<td>23</td>
<td>74</td>
<td>30</td>
<td>17</td>
</tr>
<tr>
<td>24</td>
<td>213</td>
<td>51</td>
<td>26</td>
</tr>
</tbody>
</table>

- Breast Cancer Core Database (BIC)
- Human Gene Mutation Database (HGMD-Professional)
- Universal Mutation Database (UMD)
- Leiden Open Variation Database (LOVD)
- NCBI ClinVar
- Not published before!
Nomenclature

**BRCA1:** U14680, NG_005905.2, NM_007294.3, LRG_292

\[ \text{c.4356delA, p. (Ala1453Glnfs*3)} \]
Confirmed with Sanger sequencing, c.4356delA, p. (Ala1453Glnfs*3)

- Frameshift, exon 12
- Splice site
- Likely pathogenic, class 4?
Conclusion: BRCA1, c.4356delA, p. (Ala1453Glnfs*3)

Class 4- Likely pathogenic

- Frameshift, splice-site
- Not published before
Result: BRCA1, c.4356delA, p. (Ala1453Glnfs*3)
Novel BRCA1 splice-site mutation in ovarian cancer patients of Slavic origin


Familial Cancer
ISSN 1389-9600
Familial Cancer
DOI 10.1007/s10689-017-0022-x
RAD51C, c.404+1G>C

- March 2016.
- Clinical Center Kragujevac
- 70 years, Dg. Adenocarcinoma serosum ovari
- Family history: unknown
NGS analysis

<table>
<thead>
<tr>
<th>Gene</th>
<th>Variant</th>
<th>Coordinate</th>
<th>Genotype</th>
<th>Exonic</th>
<th>Filters</th>
<th>Quality</th>
<th>Read Depth</th>
<th>Alt Read</th>
<th>Consequence</th>
<th>HGVS.c</th>
<th>HGVS.p</th>
<th>dbSNP ID</th>
<th>Allele Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRCA1</td>
<td>C&gt;C/T</td>
<td>41222975</td>
<td>het</td>
<td>yes</td>
<td>PASS</td>
<td>212117</td>
<td>179</td>
<td></td>
<td>missense</td>
<td>NP.00923</td>
<td>NM_007308.2</td>
<td>rs1799967</td>
<td>1/1 non-pathogenic/pathogenic</td>
</tr>
<tr>
<td>BRCA1</td>
<td>T&gt;C/C</td>
<td>41223096</td>
<td>hom</td>
<td>yes</td>
<td>PASS</td>
<td>640097</td>
<td>243</td>
<td></td>
<td>missense</td>
<td>NP.00923</td>
<td>NM_007308.2</td>
<td>rs1799967</td>
<td>33/40 unknown/non-pathogenic</td>
</tr>
<tr>
<td>BRCA1</td>
<td>T&gt;C/C</td>
<td>41244000</td>
<td>hom</td>
<td>yes</td>
<td>PASS</td>
<td>727117</td>
<td>246</td>
<td></td>
<td>missense</td>
<td>NP.00923</td>
<td>NM_007308.2</td>
<td>rs16942</td>
<td>32/40 unknown/non-pathogenic</td>
</tr>
<tr>
<td>BRCA1</td>
<td>T&gt;C/C</td>
<td>41244343</td>
<td>hom</td>
<td>yes</td>
<td>PASS</td>
<td>603439</td>
<td>184</td>
<td></td>
<td>missense</td>
<td>NP.00923</td>
<td>NM_007308.2</td>
<td>rs16941</td>
<td>30/38 non-pathogenic</td>
</tr>
<tr>
<td>RAD51C</td>
<td>G&gt;A/A</td>
<td>56772551</td>
<td>hom</td>
<td>yes</td>
<td>PASS</td>
<td>764532</td>
<td>241</td>
<td></td>
<td>missense</td>
<td>NP.00923</td>
<td>NM_007308.2</td>
<td>rs799917</td>
<td>48/45 unknown/non-pathogenic</td>
</tr>
<tr>
<td>RAD51C</td>
<td>G&gt;G/C</td>
<td>56772551</td>
<td>het</td>
<td>yes</td>
<td>PASS</td>
<td>92920</td>
<td>100</td>
<td></td>
<td>Intronic</td>
<td>UPSTREAM</td>
<td>NA</td>
<td>00118386.1</td>
<td>0/0 non-pathogenic</td>
</tr>
<tr>
<td>CHEK2</td>
<td>T&gt;T/C</td>
<td>29130012</td>
<td>het</td>
<td>no</td>
<td>PASS</td>
<td>111158</td>
<td>108</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>68/65 unknown/non-pathogenic</td>
</tr>
</tbody>
</table>
RAD51C, c.404+1G>C

- Human Gene Mutation Database (HGMD-Professional)
- Universal Mutation Database (UMD)
- Leiden Open Variation Database (LOVD)
- NCBI ClinVar
- Not published before

Class 4 - Likely pathogenic

- Frameshift, splice-site
These results confirm that RAD51C and RAD51D are moderate ovarian cancer susceptibility genes and suggest that they confer levels of risk of EOC that may warrant their use alongside BRCA1 and BRCA2 in routine clinical genetic testing.

The panel recommends that RRSO in RAD51C mutation carriers be considered beginning the ages 45 to 50!
**NGS limitations**

**Technical:**
- Bioinformatics
- Copy number variations (CNVs)
- Low depth of coverage
- Somatic mutations (FFPE sections)

**Clinical:**
- VUS interpretation
- Additional testing for family members
- Management of incidental findings

**Other:**
- Price!
- Repeating the run?
• Major challenge - complex computer infrastructure - expensive and unaffordable

• "Cloud computing" resources - storage, backup and computational resources
Genetic information is the key!