Whole Genome Analysis

100,000 Genomes Project Cancer Programme

Preliminary analysis: somatic small non-synonymous variants and pertinent germline findings in cancer susceptibility genes v1.3

Participant information

Participant name	D.O.B	Gender	NHS number	Laboratory sample ID	GeL participant ID	GMC	Sample date	Date analysis issued
		F		XXX	XXX	XXX	27-10-2016	15-03-2017

Tumour information

Tumour type	Tumour subtype	ICD10 code	Sample type	Reported tumour content	Tumour sample cross-contamination
Ovarian	low grade serous adenocarcinoma	C56X	FF	N/A	Pass

Domain 1 somatic variants

Variants in a virtual panel of potentially actionable genes*. Actionable genes are defined as genes in which small variants (SNVs and indels <50bp) have reported therapeutic, prognostic or clinical trial associations**, as defined by the GenomOncology Knowledge Management System. Where known, the “variant-level actionability” category and applicable tumour type are indicated. For other variants in these genes, their impact on gene function has not yet been characterised and therefore their actionability status is unclear. This means:

(i) local evaluation will be required for listed variants which are not yet characterised (i.e. “variant-level actionability” is denoted N/A)

(ii) even if well characterised as actionable for some tumour types, the listed variants may not be actionable in the participant’s specific tumour type

*Current potentially actionable genes for solid tumours: 74 genes, listed at Actionable genes in solid tumour v1.3 document

**Links are provided to clinical trials within the United Kingdom which are both actively recruiting participants or closed to recruitment.

Gene	GRCh38 coordinates ref/alt allele	Transcript	cDNA and protein change	Predicted consequences	Population germline allele frequency (1KG)	VAF	Alt allele/ total read depth	COSMIC ID	Gene-level actionability	Variant-level actionability	Gene mode of action
TP53	17:7673717 T>TG	ENST00000269305	c.902dupC p.(Gly302fs)	frameshift_variant	N/A	0.21	15/71	COSM5132150 COSM5132151 COSM5132149 COSM5132148	Trial (breast ca); Trial (NSC lung ca); Trial (SCLC); Trial (solid neoplasm); Trial (ovarian ca)	Trial (breast ca); Trial (NSC lung ca); Trial (SCLC); Trial (solid neoplasm)	oncogene, tumour suppressor

Complex indels are only annotated at the cDNA level
(H) next to the consequence type denote indels intersecting with reference homopolymers of at least 8 nucleotides in length
(N) next to the consequence type denote indels in the regions with high levels of sequencing noise where at least 10% of the basecalls in a window extending 50 bases to either side of the indel's call have been filtered out due to the poor quality

Domain 2 somatic variants

Variants in a virtual panel of cancer-related genes***. Cancer-related genes are defined as genes in which any variants have been causally implicated in cancer, as defined by the Cancer Gene Census (Wellcome Trust Sanger Institute)
***Current cancer-related genes: 590 genes, listed at Cancer census genes v1.3 document

Gene	GRCh38 coordinates ref/alt allele	Transcript	cDNA and protein change	Predicted consequences	Population germline allele frequency (1KG)	VAF	Alt allele/total read depth	COSMIC ID	Gene mode of action
EIF4A2	3:186783629 G>A	ENST00000323963	c.19G>A p.(Asp7Asn)	missense_variant	N/A	0.09	12/133	N/A	N/A
HIST1H3B	6:26031913 G>T	ENST00000621411	c.148C>A p.(Arg50Ser)	missense_variant	N/A	0.08	10/128	N/A	N/A
HSP90AA1	14:102082452 C>G	ENST00000334701	c.2122-8G>C	splice_region_variant	N/A	0.09	9/105	N/A	N/A
MECOM	3:169122682 CTTGTA>C	ENST00000494292	c.871_875del5 p.(Tyr291fs)	frameshift_variant	N/A	0.07	11/155	N/A	oncogene
NIN	14:50770516 C>T	ENST00000245441	c.1306G>A p.(Glu436Lys)	missense_variant	N/A	0.13	14/112	N/A	N/A
NOTCH2	1:119925747 A>T	ENST00000256646	c.4069T>A p.(Cys1357Ser)	missense_variant	N/A	0.1	10/96	N/A	oncogene, tumour suppressor

Complex indels are only annotated at the cDNA level
(H) next to the consequence type denote indels intersecting with reference homopolymers of at least 8 nucleotides in length
(N) next to the consequence type denote indels in the regions with high levels of sequencing noise where at least 10% of the basecalls in a window extending 50 bases to either side of the indel's call have been filtered out due to the poor quality

Domain 3 somatic variants

Small variants in genes not included in domains 1 & 2. These are not included in this document but are accessible via the Supplementary Analysis.

Pathogenic germline cancer susceptibility variants

The following variants have been identified on analysis for pertinent germline findings (known pathogenic or likely pathogenic variants in cancer susceptibility genes relevant to the tumour type).

Gene	GRCh38 coordinates ref/alt allele	Transcript	cDNA and protein change	Predicted consequences	Population germline allele frequency (1KG)	Alt allele/ total read depth	Genotype	ClinVar ID	Gene mode of action
BRCA2	13:32338140 C>G	ENST00000544455	c.3785C>G p.(Ser1262*)	stop_gained	N/A	19/42	0/1	RCV000129108 RCV000113219 RCV000044266	tumour suppressor

For more details of analysis and the genes included for pertinent germline findings, please refer to Technical Information v1.3.main.

Sequencing quality information

See online Technical Information v1.3.main document and/or LabKey QC portal for details and expected ranges of QC metrics

Sample type	Mapped reads, %	Chimeric DNA fragments, %	Insert size median, bp	Genome-wide coverage mean, x	Unevenness of local genome coverage, x	COSMIC content with low coverage (<30x), %	Total somatic SNVs	Total somatic indels	Total somatic SVs
Germline	95.91	0.34	459.0	31.38	6.56	N/A	N/A	N/A	N/A
Tumour	95.94	0.29	483.9	98.64	15.32	0.89	19713	3576	473

Additional information

- **The pathways for sample processing and data analysis are not yet accredited end-to-end for diagnostic use. Accordingly, any result intended for use in informing clinical management should be confirmed using a test accredited for clinical use.**
- **Sensitivity: the depth of WGS used in this analysis will typically detect 99% of somatic SNVs with an allele frequency of ≥0.3, 95% of somatic SNVs with an allele frequency of ≥0.1 and 60% of somatic indels with an allele frequency of ≥0.2 (estimate is based upon admixtures analysis of a highly accurate catalog of variants produced in the "platinum genomes" project). Consequently, somatic variants with allelic frequencies below this level, or in areas of low coverage may not be detected. False negative results cannot be excluded.**
- Specificity: as yet the expected false positive rates across a range of somatic variant types and allele frequencies has not been determined. Therefore false positive results cannot be excluded.
- Variant calls are filtered according to the quality and quantity of reads. Full details of the filters used in this analysis can be found in the Technical Information v1.3.main.
- Variants present in the germline are subtracted to produce a list of somatic variants. Accordingly, variants detected in both the germline and the tumour will not be listed in this analysis with the exception of known pathogenic or likely pathogenic variants in cancer susceptibility genes relevant to the tumour type.
- In this analysis MNVs (multiple nucleotide variants) are reported as multiple consecutive SNVs and therefore the protein change may require correction.
- Complex indels are only annotated at the cDNA level owing to problems accurately annotating the protein change with the current pipeline.
- Only variants with specific consequences (transcript ablation, splice acceptor variant, splice donor variant, stop gained, frameshift variant, stop lost, start lost, transcript amplification, inframe insertion, inframe deletion, missense variant, splice region variant) in canonical transcripts are reported. The complete list of canonical transcripts can be accessed at List of canonical transcripts v1.3.
- A somatic variant may have multiple entries in COSMIC database due to the use of different reference sequences. In these cases links to all COSMIC entries are provided.
- Structural variants (SVs) and copy number variants (CNVs) are only included in this Supplementary Analysis. Please note that somatic SVs and CNVs that occur within domain 1 & 2 genes are not included in this Preliminary Analysis while the performance (recal and precision) of the calling algorithm for CNVs and SVs is under evaluation.
- The germline analysis undertaken may not be fully sensitive on account of coverage and the reference data used for assessment of non-truncating variants. If the patient has been evaluated as clinically eligible for germline genetic testing on account of their personal and/or family history of cancer, this testing should be performed as per standard local practise.
- •If a pathogenic or likely pathogenic germline susceptibility variant is detected, it is recommended that the variant is reviewed by a local clinical laboratory service with expertise in germline cancer genetics. Referral to a clinical cancer genetics unit and technical confirmation of the variant in a new blood sample may be recommended following local variant review.
- For a full description of the methods used to produce these results and for further information regarding QC metrics please refer to the Technical Information v1.3.main. All related documentation is available at Genomics England Website.
- 'N/A' indicates that information is not available or not applicable.

Genomics England
Queen Mary University of London
Dawson Hall
Charterhouse Square
London
EC1M 6BQ

Sequencing Laboratory
Illumina Laboratory Services United Kingdom - Hinxton
The Ogilvie Building, Wellcome Trust Genome Campus
Hinxton Nr Saffron Walden
Essex
CB10 1DR