Sequence to Clinical Report: Enabling Rapid Variant Analysis and Clinical Interpretation for Somatic NGS Testing

Ritu Kamal[1], Cheryl Bellevia[1], Andrew Guo[1], Edward S. Kiruluta[1], Martin G. Reese[1] 1. rkamal@fabricgenomics.com, Fabric Genomics™ Inc., Oakland, CA

ABSTRACT

NGS testing of tumor variants is increasingly performed in many clinical settings, but is often inefficient, with pipelines cobbled together from several separate pieces of software. Here, we demonstrate an end-to-end bioinformatic pipeline, from FASTQ files through to a clinically actionable test report, optimized for the analysis and reporting of somatic variants.

NGS testing was performed on tumor-only samples using an Illumina sequencer and generated FASTQ files were uploaded to the cloud. A custom somatic secondary analysis pipeline aligned and called variants for each sample. Variant calling was performed using the same mathematics as the Mutect2 somatic variant caller for both SNP and InDel calling. The Opal™ tertiary analysis pipeline took variant calls as input in VCF format, and annotated all variants. Germline variants were removed using variant filtering based on population frequencies from the ExAc dataset. Selected variants were curated against publicly available literature and scored according to AMP somatic scoring guidelines.

The secondary analysis pipeline showed high sensitivity against known benchmarks. Tertiary analysis provided molecular and clinical interpretation for each sample, with at least 2 targeted therapy matches and multiple clinical trial matches reported for each tumor sample analysed. End-to-end data processing and analysis for this pre-curated panel was completed in under one hour. Case-specific interpretation using the same pipeline can be provided with next-day turnaround.

The Opal™ Clinical Somatic bioinformatics pipeline demonstrates a seamless and accurate method to process, annotate and clinically interpret somatic variants. Rapid turnaround time on the processing and analysis of NGS data enable faster reporting of clinically actionable results. We expect this to drive better cancer care.

Rapid and Accurate Analysis of Variants from Somatic Cancer Exomes

Cancer samples were collected from publicly available data sources, spiked and analyzed using the Fabric Enterprise for Oncology pipeline. The samples were all cancer exomes, which were analyzed using an in-silico panel, the Illumina TruSight Tumor 170 gene panel. This pipeline consists of 5 distinct steps:

1. FASTQ to VCF via alignment and variant calling using the Sentieon TNScope algorithm
2. VCF annotation, including variant and gene-level data from cancer-specific databases against the Illumina TruSight Tumor 170 gene panel
3. Detailed clinical curation, including drug and clinical trial matches for each variant
4. Scoring and classification of each variant, based on the 2017 AMP somatic variant scoring guidelines
5. Creation of final clinical report

AMPS

AKT2	CCNE1	ERCC2	FGF4	FGFR4	MYCL1	RAF1
ALK	CDK4	ESR1	FGF5	JAK2	MYCN	RET
AR	CDK6	FGF1	FGF6	KIT	NRAS	RICTOR
ATM	CHEK1	FGF10	FGF7	KRAS	NRG1	RPS6KB1
BRAF	CHEK2	FGF14	FGF8	LAMP1	PDGFRA	TFRC
BRCA1	EGFR	FGF19	FGF9	MDM2	PDGFRB	
BRCA2	ERBB2	FGF2	FGFR1	MDM4	PIK3CA	
CCND1	ERBB3	FGF23	FGFR2	MET	PIK3CB	
CCND3	ERCC1	FGF3	FGFR3	MYC	PTEN	

FUSIONS

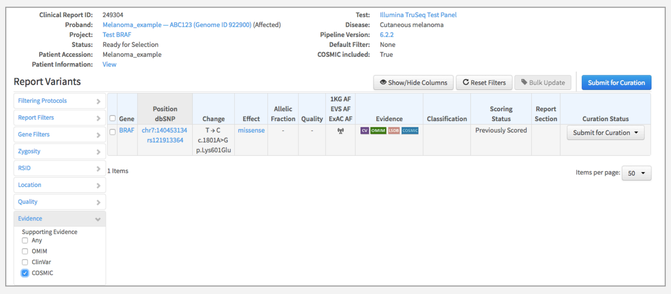
ABL1	BRCA2	ETS1	FGFR4	KMT2A (MLL)	NRG1	PIK3CA
AKT3	CDK4	ETV1	FLI1	MET	NTRK1	PPARG
ALK	CSF1R	ETV4	FLT1	MLLT3	NTRK2	RAF1
AR	EGFR	ETV5	FLT3	MSH2	NTRK3	RET
AXL	EML4	EWSR1	JAK2	MYC	PAX3	ROS1
BCL2	ERBB2	FGFR1	KDR	NOTCH1	PAX7	RPS6KB1
BRAF	ERG	FGFR2	KIF5B	NOTCH2	PDGFRA	TMPRSS2
BRCA1	ESR1	FGFR3	KIT	NOTCH3	PDGFRB	

SMALL VARIANTS

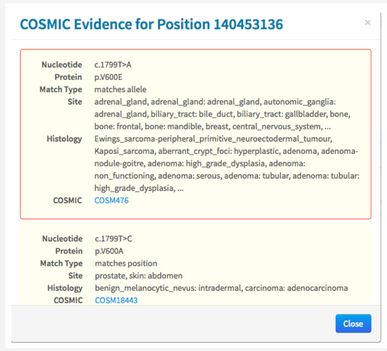
AKT1	CD79A	ERG	FGFR4	MDM2	NRAS	RICTOR
AKT2	CD79B	ESR1	FLT1	MDM4	NRG1	ROS1
AKT3	CDH1	EZH2	FLT3	MET	PALB2	RPS6KB1
ALK	CDK12	FAM175A	FOXL2	MLH1	PDGFRA	SLX4
APC	CDK4	FANCI	GEN1	MLLT3	PDGFRB	SMAD4
AR	CDK6	FANCL	GNA11	MPL	PIK3CA	SMARCB1
ARID1A	CDKN2A	FBXW7	GNAQ	MRE11A	PIK3CB	SMO
ATM	CEBPA	FGF1	GNAS	MSH2	PIK3CD	SRC
ATR	CHEK1	FGF10	HNF1A	MSH3	PIK3CG	STK11
BAP1	CHEK2	FGF14	HRAS	MSH6	PIK3R1	TERT
BARD1	CREBBP	FGF2	IDH1	MTOR	PMS2	TET2
BCL2	CSF1R	FGF23	IDH2	MUTYH	PPP2R2A	TP53
BCL6	CTNNB1	FGF3	INPP4B	MYC	PTCH1	TSC1
BRAF	DDR2	FGF4	JAK2	MYCL1	PTEN	TSC2
BRCA1	DNMT3A	FGF5	JAK3	MYCN	PTPN11	VHL
BRCA2	EGFR	FGF6	KDR	MYD88	RAD51	XRCC2
BRIP1	EP300	FGF7	KIT	NBN	RAD51B	
BTK	ERBB2	FGF8	KMT2A (MLL)	NF1	RAD51C	
CARD11	ERBB3	FGF9	KRAS	NOTCH1	RAD51D	
CCND1	ERBB4	FGFR1	MAP2K1	NOTCH2	RAD54L	
CCND2	ERCC1	FGFR2	MAP2K2	NOTCH3	RB1	
CCNE1	ERCC2	FGFR3	MCL1	NPM1	RET	

A Melanoma Case with Multiple Clinically and Biologically Relevant Mutations

In this sample set, we have a Stage III melanoma sample. Alignment, variant calling and in silico filtering with the TST170 gene panel yielded 4 variants as shown below. There are 2 somatic variants in the BRAF and NRG1 genes, one germline variant in the CDKN2A gene and a copy number variant in the MYC gene.



Over 90+ public and proprietary databases are queried to annotate variants in the Fabric pipeline. Annotation yields gene-level information from RefSeq, positional information from dbSNP, effects caused by the variant such as missense or frameshift, and quality information.



Detailed variant annotation of these variants also yields information from cancer specific databases like COSMIC. The Fabric pipeline is capable of annotating somatic, germline and structural variants across all cancer types.

Detailed clinical curation and interpretation, including drug and clinical trial matches for each variant, is provided as part of the Fabric pipeline with the help of third-party curation services. Turnaround time is dependent on the number of variants curated, but is typically less than one day. Information returned includes, but is not limited to:

- Gene information
- Cancer pathway information
- Variant level information
- Targeted therapies approved in the tumor type
- Targeted therapies approved in other tumor types
- Therapy resistance information
- ASCO, NCCN guidelines
- Clinical trial matching, ranked by region or phase
- Interactive report with notes (molecular tumor board)
- Scored, PDF report with literature citations

In this melanoma sample, there are several clinically informative findings. A summary of the clinical curation is shown below:

Therapy	Relevant Marker	Approved indication	Likelihood of Response (if known)
Vemurafenib	BRAF V600E	Melanoma	Enhanced
Dabrafenib	BRAF V600E	Melanoma	Enhanced
Cobimetinib	BRAF V600E	Melanoma	Enhanced
Trametinib	BRAF V600E	Melanoma	Enhanced
Palbociclib	CDKN2NA (V126D)	Hormone receptor-positive HER2-negative Breast Cancer	
Ribociclib	CDKN2NA (V126D)	Hormone receptor-positive HER2-negative Breast Cancer	

Imatinib is recommended for tumors with activating c-KIT mutations. Hence, this therapy is not relevant for this mutation profile.

Clinically Actionable Reports Compliant with AMP Guidelines

The Fabric Enterprise for Oncology pipeline is compliant with the somatic variant scoring guidelines published as part of the Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists published in January 2017 in the Journal of Molecular Diagnostics. After gathering clinically actionable information about variants, the pipeline allows the user to review the evidence and score the variants as follows:

Variant Tier	Name	Level of Evidence
Tier I	Variants of Strong Clinical Significance	FDA approved drugs Drugs in NCCN guidelines Multiple, large scale studies completed
Tier II	Variants of Potential Clinical Significance	FDA approved in other cancers; "off label" Preclinical evidence
Tier III	Variants of Unknown Clinical Significance	Not seen in databases No convincing published evidence on cancer correlation
Tier IV	Benign or Likely Benign Variants	Seen in general population No published evidence on cancer correlation

In the melanoma case above, the variants were scored as follows:

Variant Name	Variant Score
BRAF V600E	Tier 1
CDKN2A p16	Tier 1
NRG1	Tier 2
MYC (cnv)	Tier 3

CONCLUSIONS

Here, we have demonstrated an end-to-end bioinformatic pipeline, from FASTQ files through to a clinically actionable test report, optimized for the analysis and reporting of somatic cancer variants.

1. "An open access pilot freely sharing cancer genomic data from participants in Texas," Becnel, L. et al, Scientific Data 3, Article number: 160010 (2016)
2. PrecisionFDA DREAM Challenge Results 2016: https://www.synapse.org/#!Synapse:syn312572/wiki/58893
3. TruSight Tumor 170 gene panel from Illumina https://www.illumina.com/products/by-type/clinical-research-products/trusight-tumor-170.html?scid=2016023VU5
4. "COSMIC: somatic cancer genetics at high-resolution," Forbes S.A., et al, Nucleic Acids Research, Volume 45, Issue D1, 4 January 2017, Pages D777–D783
5. "Tumor Genomic Profiling Reports from Different Vendors: A Comparison with Respect to Clinical Action Ability of the Provided Data," Mori Y, Levenson V, Otto J (2016). Adv Mol Diag 1:110.
6. "Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer," Li, Marilyn M. et al, The Journal of Molecular Diagnostics , Volume 19 , Issue 1 , 4 - 23

FABRIC GENOMICS™

FORMERLY Omicia

Learn more at

www.fabricgenomics.com ■ info@fabricgenomics.com ■ 510.595.0800