



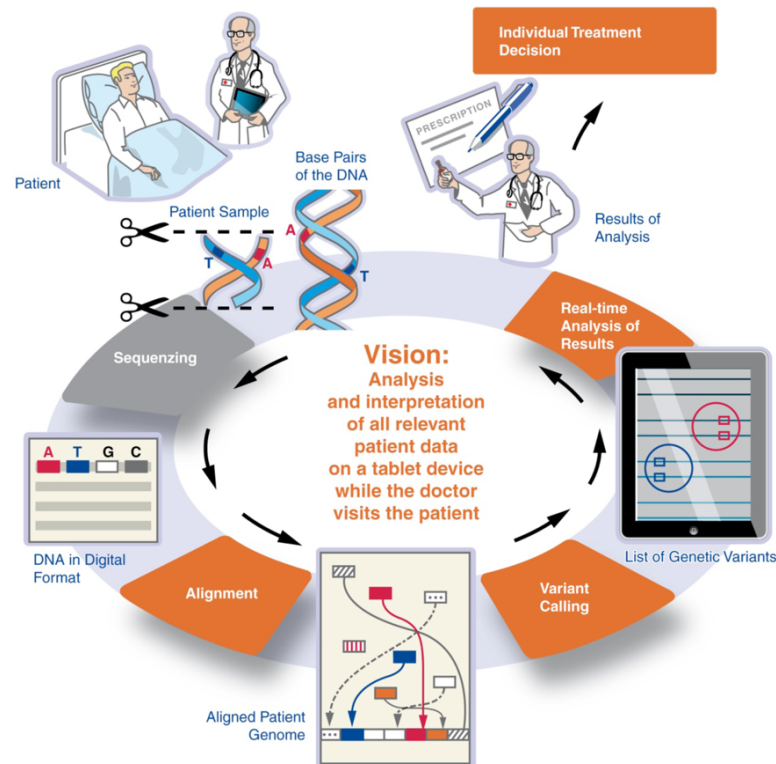
Genetic Annotations

Dr. Matthieu-P. Schapranow
Code of Life

November 2016

From Raw Genome Data to Analysis

- **DNA Sequencing:** Transformation of analogues DNA into digital format
- **Alignment:** Reconstruction of complete genome with snippets
- **Variant Calling:** Identification of genetic variants
- **Data Annotation:** Linking genetic variants with research findings



Genetic Annotations
Schapranow, Nov 2016

Genetic Annotations

- Purpose:
 - Assess impact of genetic changes
 - Understand gene function and possible medical therapy options

- Input: List of genetic variants

CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	SAMPLE
chr7	140753336	rs113488022	T	A	61	PASS	NS=1	GT	0/1

- Output: Details about certain genetic locus

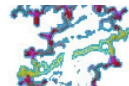
Genetic Annotations
Schapranow, Nov 2016

Challenges Today

- Manual, time-consuming Internet search, e.g. publications, annotations, guidelines
- International consortiums provide fragmented information
- Missing linkage across individual data sources
- Annotations vary in completeness and correctness

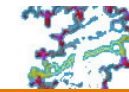


Genetic Annotations
Schapranow, Nov 2016



Interpretation of Annotations: BRAF Gene dbSNP

dbSNP
Short Genetic Variations



■ https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=rs113488022

Reference SNP (refSNP) Cluster Report: rs113488022

**** With Pathogenic allele ****

Clinical Significance: With Pathogenic allele

Primary Assembly Mapping

Assembly	SNP to Chr	Chr	Chr position	Contig	Contig position	Allele
GRCh38.p2	Rev	7	140753336	NT_007933.16	78246557	A

RefSeqGene Mapping

RefSeqGene	Gene (ID)	SNP to RefSeqGene	Position	Allele
NG_007873.3	BRAF (673)	Fwd	176429	T

Gene Model(s)

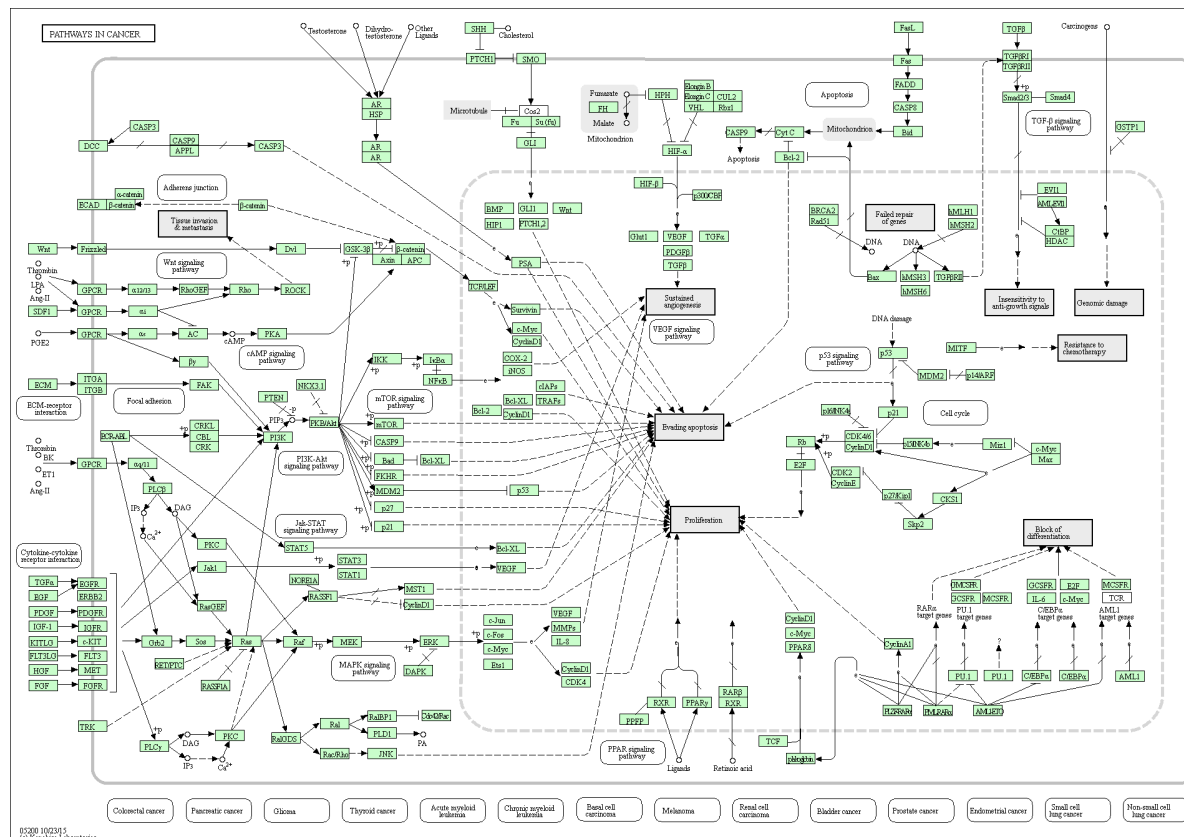
Function	mRNA				Protein		
	SNP to mRNA	Accession	Position	Allele change	Accession	Position	Residue change
missense	Fwd	NM_004333.4	1860	GTG ⇒ GAG	NP_004324.2	600	V [Val] ⇒ E [Glu]
missense	Fwd	XM_005250045.1	1807	GTG ⇒ GAG	XP_005250102.1	600	V [Val] ⇒ E [Glu]
missense	Fwd	XM_005250046.1	1807	GTG ⇒ GAG	XP_005250103.1	600	V [Val] ⇒ E [Glu]
missense	Fwd	XM_011516529.1	1807	GTG ⇒ GAG	XP_011514831.1	600	V [Val] ⇒ E [Glu]
ncRNA	Fwd	YP_042400.1	1807	N/A ⇒ N/A	N/A	N/A	N/A

Sample Ascertainment					Genotypes		Alleles	
ss#	Population	Individual Group	Chrom. Sample Cnt.	Source	HWP	A	T	
ss1688979043	ExAc Aggregated Populations		121410	AF		0.00001647	0.99998355	

Genetic Annotations

Schapranow, Nov 2016

Interpretation of Annotations: BRAF Gene Kegg



Genetic Annotations
Schapranow, Nov 2016

GeneCards
HUMAN GENE DATABASE



Aliases for BRAF Gene

(Serine/Threonine Kinase) is a Protein Coding gene. Diseases associated with BRAF include [Cardiofaciocutaneous Syndrome](#) and [Lung Cancer](#). Among its [parent Slit-Robo signaling](#) and [MAPK Signaling: Oxidative Stress Pathway](#). GO annotations related to this gene include *calcium ion binding* and *transferase is-containing groups*. An important paralog of this gene is [RAF1](#).

HGNC: [1097](#) Entrez Gene: [673](#) Ensembl: [ENSG00000157764](#) OMIM: [164757](#) UniProtKB: [P15056](#)

GC07M138705, GC07M139764, GC07M139841, GC07M139754, GC07M139887, GC07M140080, GC07M140424, GC07M134729

Export aliases for BRAF gene to outside databases

Summaries for BRAF Gene

Schapranow, Nov 2016

Interpretation of Annotations: BRAF Gene

PubMed



Neuropathology. 2016 Oct 28. doi: 10.1111/neup.12347. [Epub ahead of print]

A comprehensive analysis identifies **BRAF hotspot mutations** associated with gliomas with peculiar epithelial morphology.

Hatae R¹, Hata N², Suzuki SO³, Yoshimoto K², Kuga D², Murata H², Akagi Y², Sangatsuda Y², Iwaki T³, Mizoguchi M⁴, Iihara K².

Author information

Abstract

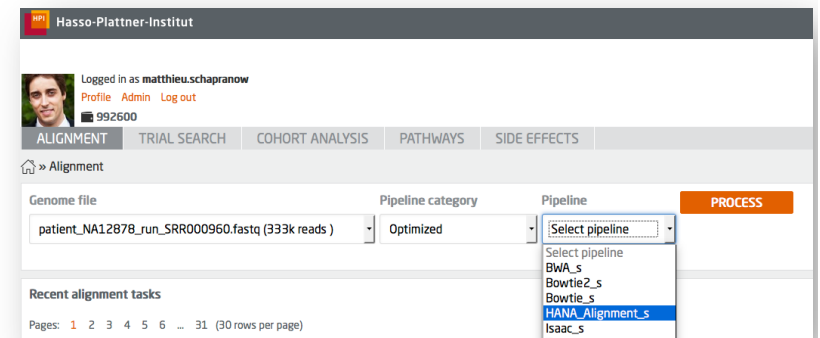
Brain tumors harbor various BRAF alterations, the vast majority of which are the BRAF kinase-activating V600E mutation. BRAF mutations are most frequently detected in certain subtypes of low-grade glioma, such as pilocytic astrocytoma (PA), pleomorphic xanthoastrocytoma (PXA), ganglioglioma (GG) and dysembryoplastic neuroepithelial tumor (DNT). However, it is unclear whether gliomas harboring BRAF mutations can be invariably regarded as these glioma subtypes or their derivatives. To address this question, we analyzed 274 gliomas in our institutional case series.

Genetic Annotations

Schapranow, Nov 2016

What is Next?

- How to guarantee reproducibility of findings?
- How to enable scalability for high-throughput sequencing?



Genetic Annotations

Schapranow, Nov 2016

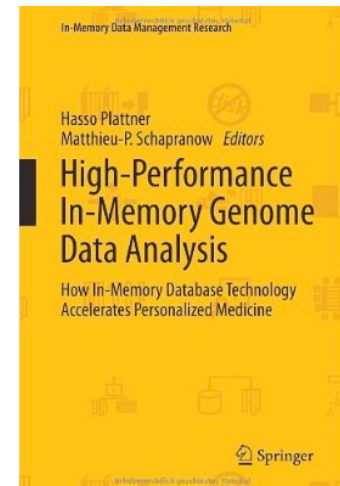
Keep in contact with us!



Dr. Matthieu-P. Schapranow
Program Manager E-Health & Life Sciences

Hasso Plattner Institute
August-Bebel-Str. 88
14482 Potsdam, Germany

schapranow@hpi.de
<http://we.analyzegenomes.com/>



Genetic Annotations
Schapranow, Nov 2016