

Variant prioritization

Genetische Varianten

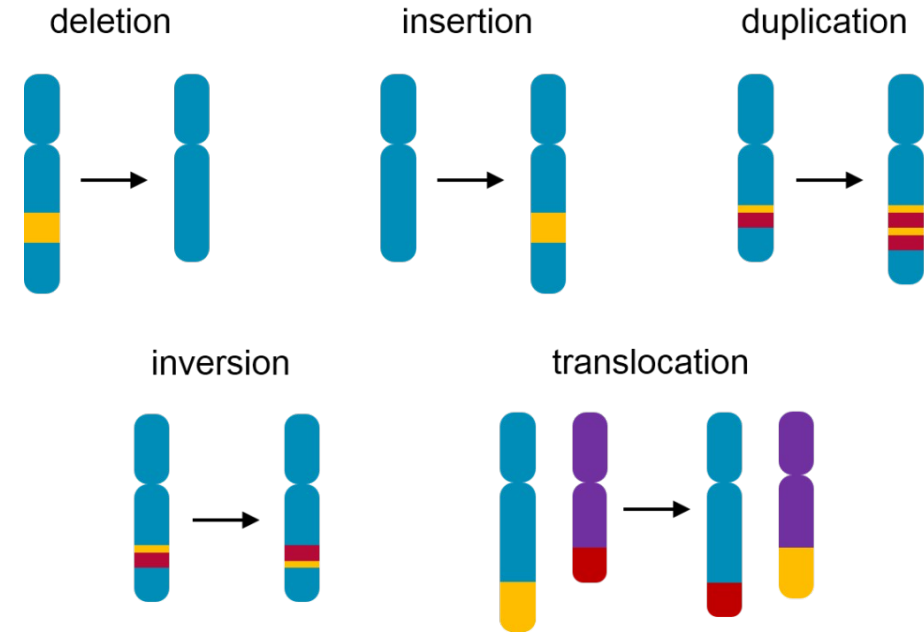
- Mutationen im Genom
- Typen:
 - Single nucleotide variants
 - Structural variants
- 8,9 m bp (13%) strukturell
- 3,6 m SNVs

Individual 1

...CGATATTCCTATCGAATGTC...
...GCTATAAGGATAGCTTACAG...

Individual 2

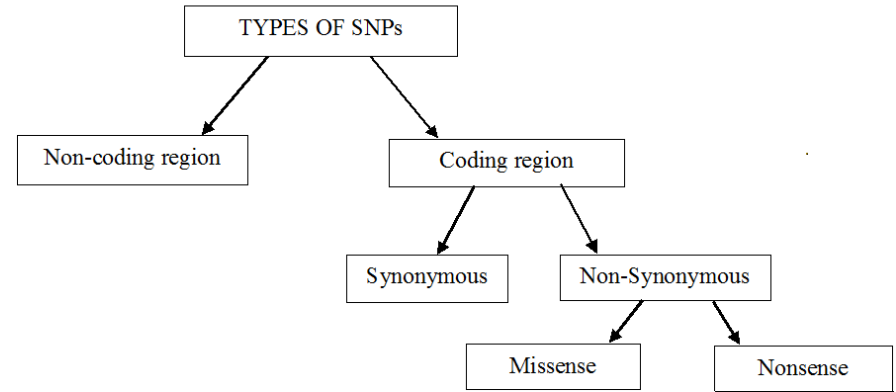
...CGATATTCCTATCGAATGTC...
...GCTATAAGGCTAGCTTACAG...



<https://www.pacb.com/wp-content/uploads/Structural-variant-types-1024x728.png> (21.05.2024)

Auswirkungen: Protein

- Keine Auswirkungen
 - Dominant-rezessiver Erbgang
 - Synonyme Aminosäure
- Dysfunktionales Protein
- Höhere / geringere Expression



By MansiG123 - Own work, CC BY-SA 4.0,
<https://commons.wikimedia.org/w/index.php?curid=48716253>

Auswirkungen: Phänotyp

Beschädigtes Protein \neq Krankheit

- Neutrale Mutation
- Stille Mutation
- Konditional-letal
- Letal

Priorisierung

- **Annotation:** Vorhersage der Auswirkung einer Variante
- **Priorisierung:** Sortierung der gefundenen Varianten nach bestimmten Kriterien:
 - Auswirkungen auf Gesundheit
 - Häufigkeit in der Bevölkerung
 - Vorhergesagte Auswirkung
- Ziel: Identifizierung von krankmachenden Varianten
- Entscheidungshilfe für klinische Experten

Methoden

Problem: sehr viele harmlose Varianten

- Möglichkeiten:
 - Vergleich mit Gesamtbevölkerung
 - Bekannte Effekte bestimmter Varianten
 - Patientenhistorie: Ethnie, Familiengeschichte, Phänotyp

Phylogenetische Konservierung

- Betrachtung einer Region bei unterschiedlichen Spezies
- Variante in stark konservierter Region = großes Schadenspotential
- Probleme:
 - stark und schwach konservierte Regionen
 - stop-codons/frameshifts
 - Viele falsch-positive Ergebnisse

Population allele frequency

- Datenbanken mit sequenzierten Genomen
- Katalog mit vorkommenden genetischen Varianten
- Allgemeiner Überblick über häufige genetische Varianten

Population stratification

- Manche Varianten sind sehr viel häufiger in bestimmten Ethnien
- Untersuchung der Häufigkeit unter Aspekt der Ethnie

Gene constraint

- Statistisches Modell
- Vorhersage: Toleranz für Varianten im Vergleich zum ganzen Genom
- Datenbank mit sequenzierten Genomen
- Vergleich: Anzahl funktioneller Varianten vs Gesamtzahl

Tool: Exomizer

- Identifikation von genetischen Auslösern von Krankheiten aus Genom oder Exom
- VCF + Phenotyp → Liste: mögliche verantwortliche Varianten
- Filter anhand von benutzerdefinierten Kriterien

Input: Variant Call Format

- UTF-8
- Nur Abweichungen von Referenzgenom
- Weitere Daten: Qualität, Beschreibung der Variante, ...

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	NA000001	NA000002	NA000003
20	14370	rs6054257	G	A	29	PASS	NS=3;DP=14;AF=0.5;DB;H2	GT:GQ:DP:HQ	0 0:48:1:51,51	1 0:48:8:51,51	1/1:43:5:.,.
20	17330	.	T	A	3	q10	NS=3;DP=11;AF=0.017	GT:GQ:DP:HQ	0 0:49:3:58,50	0 1:3:5:65,3	0/0:41:3
20	1110696	rs6040355	A	G,T	67	PASS	NS=2;DP=10;AF=0.333,0.667;AA=T;DB	GT:GQ:DP:HQ	1 2:21:6:23,27	2 1:2:0:18,2	2/2:35:4
20	1230237	.	T	.	47	PASS	NS=3;DP=13;AA=T	GT:GQ:DP:HQ	0 0:54:7:56,60	0 0:48:4:51,51	0/0:61:2
20	1234567	microsat1	GTC	G,GTCT	50	PASS	NS=3;DP=9;AA=G	GT:GQ:DP	0/1:35:4	0/2:17:2	1/1:40:3

https://en.wikipedia.org/wiki/Variant_Call_Format

Input: Human Phenotype Ontology

- Festgelegte Begriffe zur Beschreibung von Krankheitssymptomen
- Speziell für Erbkrankheiten
- Verwendbar für Algorithmen

Output

- HTML / JSON
- Priorisierte und gefilterte Gene mit Score

FGFR2

Exomiser Score: **0.998**
(p=2.0E-6)

Phenotype Score: **1.000**

Variant Score: **1.000**

Phenotype matches:

Phenotypic similarity 0.874 to [Jackson-Weiss syndrome](#) associated with FGFR2.

Best Phenotype Matches:

HP:0001156, Brachydactyly - HP:0010055, Broad hallux

HP:0001363, Craniosynostosis - HP:0001363, Craniosynostosis

HP:0011304, Broad thumb - HP:0010077, Broad distal phalanx of the hallux

HP:0010055, Broad hallux - HP:0010055, Broad hallux

Quellen

- Karen Eilbeck, Aaron Quinlan, Mark Yadaell (2017) , Nature Reviews (Volume 18), „Settling the score: variant prioritization and Mendelian disease”
- VFC specification: <https://samtools.github.io/hts-specs/VCFv4.3.pdf>
- Robinson PN, Köhler S, Oellrich A, Sanger Mouse Genetics Project, Wang K, Mungall CJ, Lewis SE, Washington N, Bauer S, Seelow D, Krawitz P, Gilissen C, Haendel M and Smedley D: „Improved exome prioritization of disease genes through cross-species phenotype comparison“ Genome research 2014;24;2;340-8
- Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003 Jul 1;31(13):3812-4. doi: 10.1093/nar/gkg509. PMID: 12824425; PMCID: PMC168916.
- Spencer, Paige S.; Barral, José M. (2012). "Genetic code redundancy and its influence on the encoded polypeptides". Computational and Structural Biotechnology Journal. 1: e201204006. doi:10.5936/
- <https://de.wikipedia.org/wiki/Mutation> (26.05.2024)
- https://en.wikipedia.org/wiki/Single-nucleotide_polymorphism#cite_note-15 (26.05.2024)