

TMF-Workshop
Datenbank genomischer Varianten
für die klinische Anwendung und die medizinische Forschung
Satellitenworkshop im Rahmen der GfH-Tagung in Essen, 19. März 2014

Konzept einer transitorischen Genotyp-Phänotyp Datenbank

Thomas F. Wienker (wienker@molgen.mpg.de)
Max-Planck-Institut für Molekulare Genetik, Berlin-Dahlem



NGS

Next Generation Sequencing



Camille Flammarion, *L'Atmosphère. Météorologique populaire*; Paris 1888

Transitorische Genotyp–Phänotyp Datenbank

Problem:

- NGS (WGS, WES oder Panel-Diagnostik) erfolgreich durchgeführt
- alle seltenen/privaten Varianten sind annotiert und bewertet
- alle Computer-gestützten Analysen sprechen für Pathogenität
- kein Match in einer Datenbank für neutrale/polymorphe Varianten
- kein Match in einer Datenbank für krankheitsverursachende Varianten
- reguläre Segregation entspr. genetischem Modell in der Kernfamilie

Lösung (empirisch bzw. evidenzbasiert):

- “genotypisch und phänotypisch“ gleichgelagerte Fälle ($n \geq 2$)

Charles R. Darwin (1809 –1882):

«I have no faith in anything short of actual measurement and the Rule of Three.»

Transitorische Genotyp–Phänotyp Datenbank

Problem:

- NGS (WGS, WES oder Panel-Diagnostik) erfolgreich durchgeführt
- alle seltenen/privaten Varianten sind annotiert und bewertet
- alle Computer-gestützten Analysen sprechen für Pathogenität
- kein Match in einer Datenbank für neutrale/polymorphe Varianten
- kein Match in einer Datenbank für krankheitsverursachende Varianten
- reguläre Segregation entspr. genetischem Modell in der Kernfamilie

1. Eingrenzung & Beschreibung des Problems

2. Mittel & Wege zur Problemlösung

Transitorische Genotyp–Phänotyp Datenbank

Was ist eine “diagnostische Situation“ ?

gegeben: Patient/Familie mit einer klinisch-genetisch definierten Krankheit

gesucht: die pathogenetisch relevante Variante (Mutation)
im ursächlich relevanten Genlocus

D \Rightarrow **G**

nicht:

G \Rightarrow **D**

Untersuchung von Gesunden, z.B. Screening,
oder nebenläufige Befunde, “incidental/unsolicited“ findings

Clinical Application of Whole-Genome Sequencing Proceed With Care

William Gregory Feero, MD, PhD

Research

Original Investigation

Clinical Interpretation and Implications of Whole-Genome Sequencing

Frederick E. Dewey, MD; Megan E. Grove, MS; Cuiping Pan, PhD; Benjamin A. Goldstein, PhD;
Jonathan A. Bernstein, MD, PhD; Hassan Chaib, PhD; Jason D. Merker, MD, PhD; Rachel L. Goldfeder, BS;
Gregory M. Enns, MB, ChB; Sean P. David, MD, DPhil; Neda Pakdaman, MD; Kelly E. Ormond, MS;
Colleen Caleshu, MS; Kerry Kingham, MS; Teri E. Klein, PhD; Michelle Whirl-Carrillo, PhD; Kenneth Sakamoto, MD;
Matthew T. Wheeler, MD, PhD; Atul J. Butte, MD, PhD; James M. Ford, MD, PhD; Linda Boxer, MD;
John P. A. Ioannidis, MD, PhD; Alan C. Yeung, MD; Russ B. Altman, MD, PhD; Themistocles L. Assimes, MD, PhD;
Michael Snyder, PhD; Euan A. Ashley, MRCP, DPhil; Thomas Quertermous, MD

Transitorische Genotyp–Phänotyp Datenbank

Gene oder Varianten sind in ihrer diagnostischen Wertigkeit nicht abschliessend zu beurteilen, weil gleichgelagerte Vergleichsfälle fehlen; die Beurteilung ist bestenfalls “tentativ“, jedoch nicht definitiv.

Dieser vorläufige Schwebезustand wird in Anlehnung an den in der Genkartierung üblichen Sprachgebrauch als ***in limbo*** bezeichnet.

Der “in limbo“ Beurteilung einer Beobachtung wird aufgelöst durch eine zweite Beobachtung :

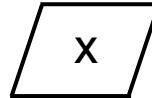
- neutrale Variante, keine pathogenetische Relevanz
- pathogenetisch relevante, kausal relevante Variante

(Evidenzbasierte Beurteilung)

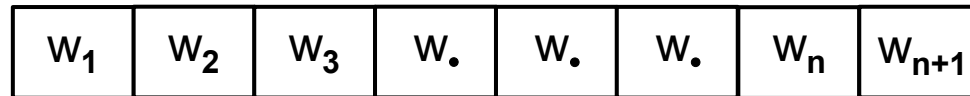
Die rasche Auflösung des “in limbo“ Zustandes ist Ziel dieser Datenbank.

Transitorische Genotyp-Phänotyp Datenbank

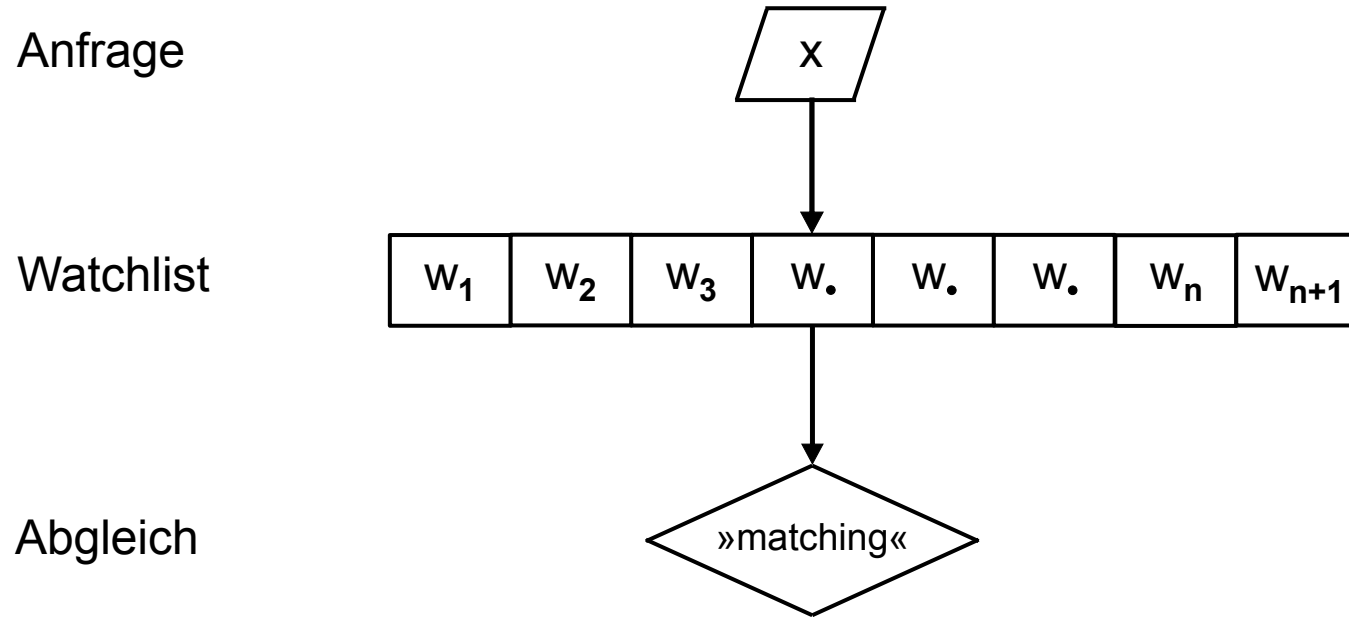
Anfrage



Watchlist

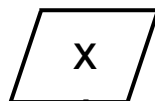


Transitorische Genotyp-Phänotyp Datenbank

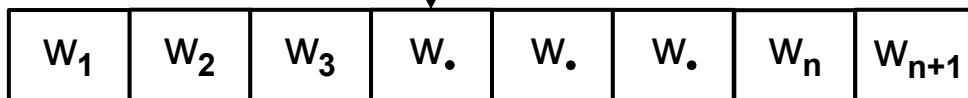


Transitorische Genotyp-Phänotyp Datenbank

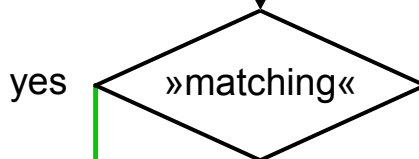
Anfrage



Watchlist



Abgleich



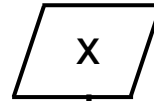
Aktion



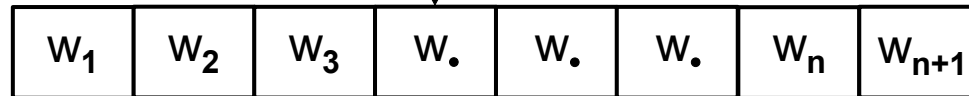
*automatische
Benachrichtigung*

Transitorische Genotyp-Phänotyp Datenbank

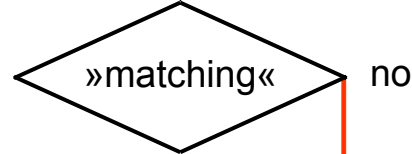
Anfrage



Watchlist



Abgleich



Rollenwechsel

Transitorische Genotyp–Phänotyp Datenbank

Rollenmodelle

Administrator / Curator

- Führung und Pflege der WATCH–LIST, Pflege der Benutzerkonten
- Überwachung und Eingangskontrolle der Anfragen, d.h. Abweisung von irrelevanten / trivialen /inkompetenten Anfragen
- Durchführung regelmässiger Prüfläufe zur Detektion geklärter Fälle
- Veranlassung des Datenabflusses, ggf. an Referenzdatenbanken

Nutzer / User

- Anfrager: ein legitimer Nutzer stellt eine Anfrage entsprechend einem der Standard-Szenarien (A oder B oder C)
- Beiträger / Contributor
bei negativer Antwort auf eine Anfrage wird diese augmentiert und als neuer Eintrag in die WATCH–LIST gestellt

Transitorische Genotyp–Phänotyp Datenbank

Rollenmodelle

Administrator / Curator

- Führung und Pflege der WATCH–LIST, Pflege der Benutzerkonten
- Überwachung und Eingangskontrolle der Anfragen, d.h. Abweisung von irrelevanten / trivialen /inkompetenten Anfragen
- Durchführung regelmässiger Prüfläufe zur Detektion geklärter Fälle
- Veranlassung des Datenabflusses, ggf. an Referenzdatenbanken

Nutzer / User

bei negativer Antwort auf eine Anfrage erfolgt ein Rollenwechsel
Anfrager → Beiträger

Transitorische Genotyp–Phänotyp Datenbank

Typische Szenarien für eine Anfrage

A) “novel gene“

Ein Genlocus, der bisher nicht als Krankheitsgen identifiziert wurde (OMIM, OrphaNet/OrphaData, HGMD, ClinVar, DmuDB, LOVD u.a.m.) mit einer seltenen/privaten, molekular-pathophysiologisch pathogenen Variante (by computational methods, z.B. Mutation Taster)

B) “novel variant“

Eine molekular-pathophysiologisch überzeugende oder verdächtige Variante in einem bekannten Krankheitsgen, bei kompatibelem Phänotyp (variant of unknown significance, VUS)

C) “novel phenotype“

Bekanntes Krankheitsgen, aber Phänotyp ausserhalb des bekannten Spektrums (OMIM: phenotypic series)

Transitorische Genotyp–Phänotyp Datenbank

Typische Szenarien für eine Anfrage

A) “novel gene“

Ein Genlocus, der bisher nicht als Krankheitsgen identifiziert wurde

Protein coding genes: ~ 22.000 (RefSeqGene, hg18)
Disease genes: ~ 3.000 (OMIM, OrphaData)

B) “novel variant“

Eine molekular-pathophysiologisch überzeugende oder verdächtige Variante in einem bekannten Krankheitsgen, bei kompatibelem Phänotyp (variant of unknown significance, VUS)

C) “novel phenotype“

Bekanntes Krankheitsgen, aber Phänotyp ausserhalb des bekannten Spektrums (OMIM: phenotypic series)

Transitorische Genotyp–Phänotyp Datenbank

Minimal-suffiziente Daten für die typischen Szenarien

A) “novel gene“

- gene symbol (HGNC)
- genetisches Modell (mode of inheritance, MOI)
- Phänotyp-Kategorie (top level HPO term, u.U. auch second level)

B) “novel variant“

- gene symbol (HGNC, OMIM genotype [*] and phenotype [#])
- genetisches Modell, MOI
- genomische Variante, HGVS ohne Mutationsbeschreibung
z.B. **NC_000015.9:g.66995761**

C) “novel phenotype“

Dieses Szenario wird überführt nach A oder B, oder A und B

Human Phenotype Ontology, HPO: <http://www.human-phenotype-ontology.org/>

HPO:0001871	Abnormality of blood and blood-forming tissues
HPO:0003549	Abnormality of connective tissue
HPO:0000152	Abnormality of head and neck
HPO:0001939	Abnormality of metabolism/homeostasis
HPO:0001197	Abnormality of prenatal development or birth
HPO:0001438	Abnormality of the abdomen
HPO:0000769	Abnormality of the breast
HPO:0001626	Abnormality of the cardiovascular system
HPO:0000598	Abnormality of the ear
HPO:0000818	Abnormality of the endocrine system
HPO:0000478	Abnormality of the eye
HPO:0000119	Abnormality of the genitourinary system
HPO:0002715	Abnormality of the immune system
HPO:0001574	Abnormality of the integument
HPO:0003011	Abnormality of the musculature
HPO:0000707	Abnormality of the nervous system
HPO:0002086	Abnormality of the respiratory system
HPO:0000924	Abnormality of the skeletal system
HPO:0001507	Growth abnormality
HPO:0002664	Neoplasm

REPORT

The Human Phenotype Ontology: A Tool for Annotating and Analyzing Human Hereditary Disease

Peter N. Robinson,^{1,2,*} Sebastian Köhler,^{1,2} Sebastian Bauer,¹ Dominik Seelow,^{1,3} Denise Horn,¹ and Stefan Mundlos^{1,2,4}

There are many thousands of hereditary diseases in humans, each of which has a specific combination of phenotypic features, but computational analysis of phenotypic data has been hampered by lack of adequate computational data structures. Therefore, we have developed a Human Phenotype Ontology (HPO) with over 8000 terms representing individual phenotypic anomalies and have annotated all clinical entries in Online Mendelian Inheritance in Man with the terms of the HPO. We show that the HPO is able to capture phenotypic similarities between diseases in a useful and highly significant fashion.

¹Institute for Medical Genetics, ²Berlin-Brandenburg Center for Regenerative Therapies, ³Department of Neuropaediatrics, Charité-Universitätsmedizin Berlin, 13353 Berlin, Germany; ⁴Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany

*Correspondence: peter.robinson@charite.de

DOI 10.1016/j.ajhg.2008.09.017. ©2008 by The American Society of Human Genetics. All rights reserved.

The American Journal of Human Genetics 83, 610–615, November 7, 2008

NGS

next generation sequencing

WGS

whole genome sequencing

WGA

whole genome analysis

WES

whole exome sequencing

WEA

whole exome analysis

TES

targeted exome sequencing

TEA

targeted exome analysis

all disease genes, ca. 3200 – 4800
selected disease genes, 2 – 525 – 1222

Transitorische Genotyp–Phänotyp Datenbank

3 leitende Prinzipien für die Datenerfassung und Datenhaltung

- beobachtungsnahe Daten, d.h. primäre molekulargenetische Befunde, weitestgehender Verzicht auf Daten der Interpretationsebenen
- Beachtung des Sparsamkeitsprinzips (*parsimony principle*), es werden **nur** die Daten erhoben, die für das Match-making **unbedingt** erforderlich sind
- Sonder- und Spezialfälle werden ggf. suboptimal oder gar nicht abgebildet

KISS

Keep it simple, stupid!

Pedigree: M 90 00 029 Gene: TLR10 HGNC:15634 OMIM: *606270 Location: 4p14

Mutation: Genomic: NC_000004.11:g.38,776,478G>T
 Transcript(s): NM_001017388.2:c.734C>A
 NM_001195106.1:c.734C>A
 NM_001195107.1:c.734C>A
 NM_001195108.1:c.692C>A
 NM_030956.3:c.734C>A
 NM_001017388.1:c.734C>A
 NM_030956.2:c.734C>A
 Protein: NP_001017388.1(TLR10):p.S245X



T/T

Pedigree: M 90 00 029 Gene: TLR10 HGNC:15634 OMIM: *606270 Location: 4p14

Mutation: Genomic: NC_000004.11:g.38,776,478G>T

Transcript(s): NM_001017388.2:c.734C>A

NM_001195106.1:c.734C>A

NM_001195107.1:c.734C>A

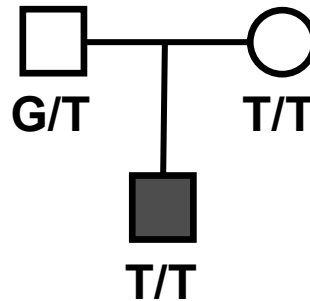
NM_001195108.1:c.692C>A

NM_030956.3:c.734C>A

NM_001017388.1:c.734C>A

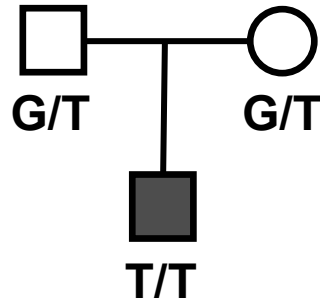
NM_030956.2:c.734C>A

Protein: NP_001017388.1(TLR10):p.S245X



Pedigree: M 90 00 029 Gene: TLR10 HGNC:15634 OMIM: *606270 Location: 4p14

Mutation: Genomic: NC_000004.11:g.38,776,478G>T
 Transcript(s): NM_001017388.2:c.734C>A
 NM_001195106.1:c.734C>A
 NM_001195107.1:c.734C>A
 NM_001195108.1:c.692C>A
 NM_030956.3:c.734C>A
 NM_001017388.1:c.734C>A
 NM_030956.2:c.734C>A
 Protein: NP_001017388.1(TLR10):p.S245X



... there is no better way to advance the proper practice of medicine than to give our minds to the discovery of the usual laws of nature by careful investigation of cases of rarer forms of diseases.

*William Harvey (1578–1657)
in a letter to the physician Jan Vlackfeld
Haarlem, Netherlands, April 24th, 1657*

cf. Groft SC and Posada M (2010)
in: Rare Disease Epidemiology, chpt. 1
Adv. Exp. Med. Biol. Vol. 686, pp. 3 –14



About TMF

Our Mission

We bring together researchers in different disciplines and develop concepts, infrastructures, and methods that promote medical research. We thus increase the safety, quality, and efficiency of medical research, ensure the necessary standardization, and offer researchers, irrespective of their scientific issues, assistance in dealing with the increasingly complex legal and regulatory infrastructure of medical research.

Vielen Dank fürs Zuhören!

Fragen werden gerne beantwortet!

Kommentare sind erwünscht!