DNA-Analytik IV

Sequenzanalyse des Humangenoms

Human-Genom-Projekt

Ziele 1985-90

Hochauflösende Kartierung des Genoms

Vollständige Sequenzierung aller Chromosomen

Identifikation aller Gene

Technologie-Entwicklung

Analyse von Modellorganismen

Untersuchung der ethisch & sozialen Konsequenzen

Sequenzierung des Human-Genoms

Herausforderungen

Größe

25x größer als bisher größtes sequenziertes Genom 2000

8x größer als alles bisher Sequenzierte

Gehalt an sich wiederholenden Sequenzen

erstes wiederholungsreiches Genom, ca.40% erwartet

Ackerschmalwand 11%, Fadenwurm 7%, Fruchtfliege 3%

Medizinische Relevanz

gesellschaftliche und persönliche Konsequenzen, ethische Brisanz

Internationales Konsortium

20 Zentren aus 6 Ländern

privatwirtschaftliche Konkurrenz 1998-2001

von 15% auf 90% in 15 Monaten (Februar 99 - Juni 2000)

Sequenzierung des Human-Genoms

Meilensteine

Oktober 1990
15-Jahresplan NIH/DOE

Februar 1995
Sulston/Waterson Plan bis 2001

3 Zentren á 85,000 Sequenzen/Woche; 5 Jahre; 99,9% Genauigkeit

Februar 1996
Selbstverpflichtung zur unverzüglichen Datenfreigabe

"Bermuda" Regeln

Mai 1996
Venter/Smith/Hood BAC-End-Sequenzierung

300.000 BACs, 600.000 Sequenzen, ein STC/5kb bis 1999

April 1998
Venter/Hunkapiller genomweite "shotgun"-bis 2001

200 ABI3700 á 1000 Seq/Tag

1. Jahr: Drosophila, 2./3. Jahr: Mensch

Oktober 1998
NIH/DOE: 5-Jahresplan zur Fertigstellung bis 2003

Rohfassung zum Ende 2001, vollständig fertig bis 2003

April 1999
NIH/Welcome: Rohfassung bis Frühjahr 2000

G5-Initiative

Dezember 1999/ Mai 2000
Veröffentlichungen der ersten Chromosomen

Chr 22: GB/Japan/USA; Chr 21: Japan/ Deutschland

Juni 2000
IHGSC/Celera Abschluß der Rohdatensammlung

Pressekonferenz Clinton & Blair

Februar 2001
Publikationen der Rohfassung

Nature/IHGSC & Science/Celera

April 2003
50 Jahre Doppel-Helix

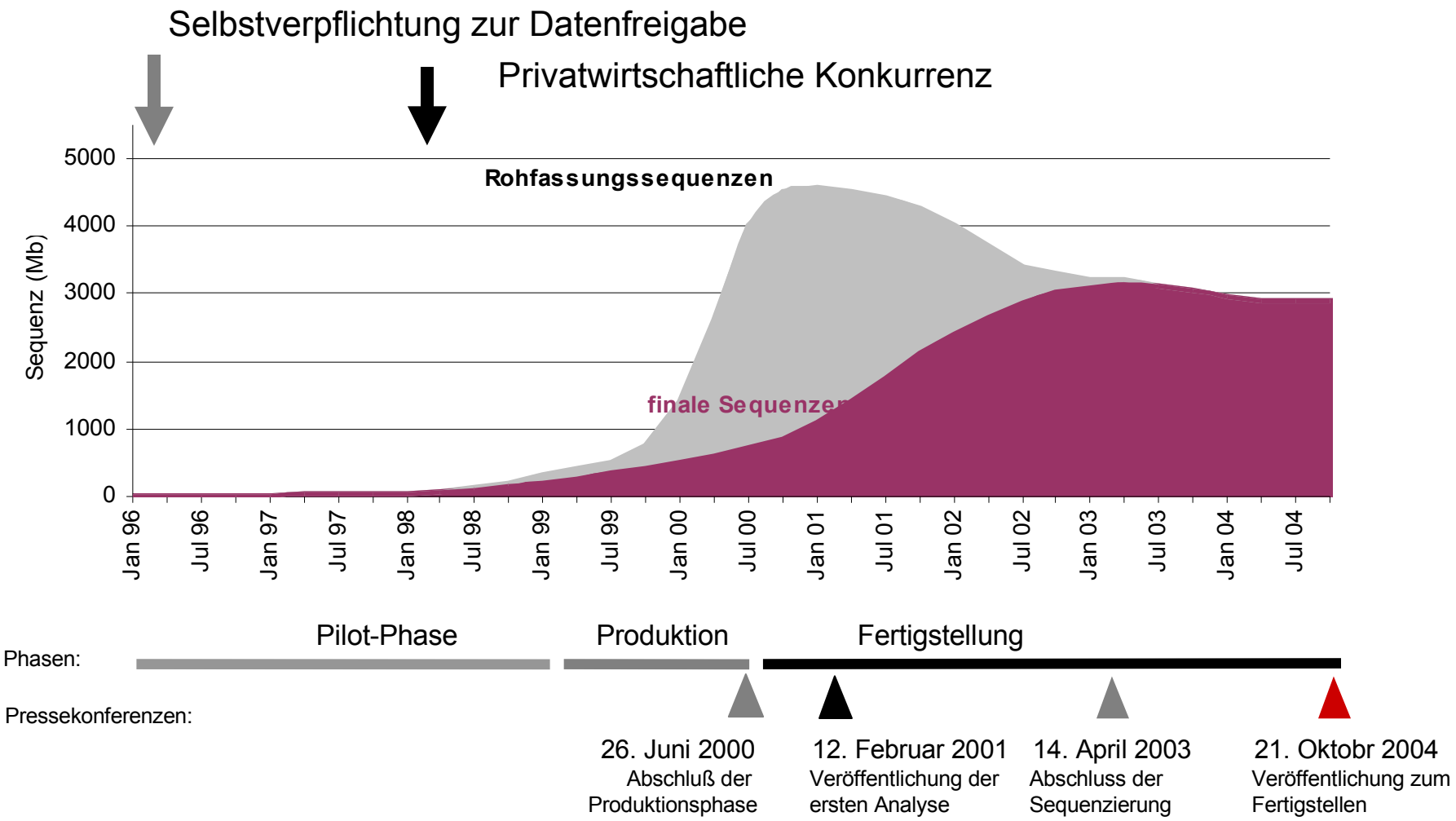
Watson & Crick

Oktober 2004
abschließende Genom-Publikation

Nature

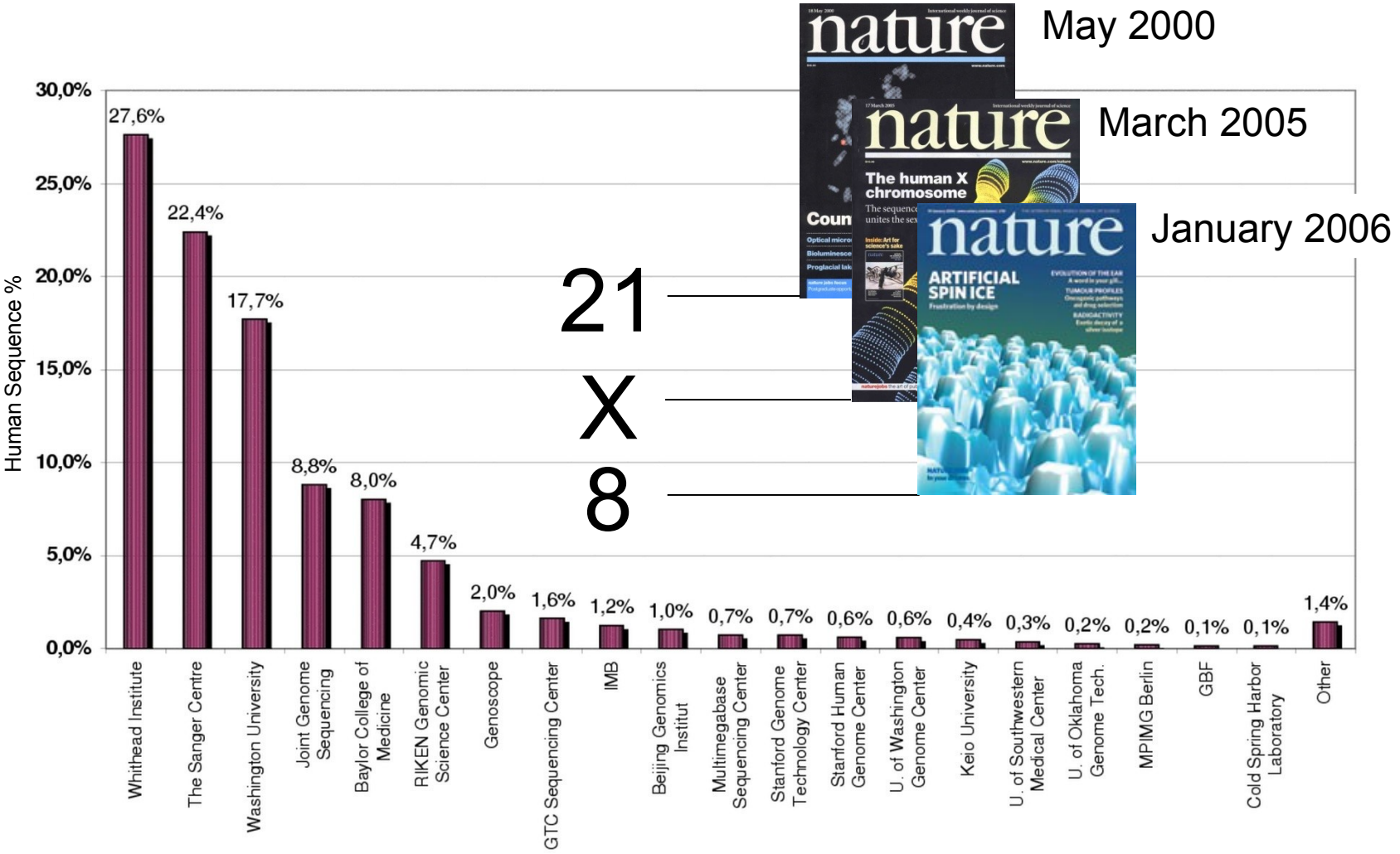
Sequenzierung des Human-Genoms

Phasen



Sequenzierung des Human-Genoms

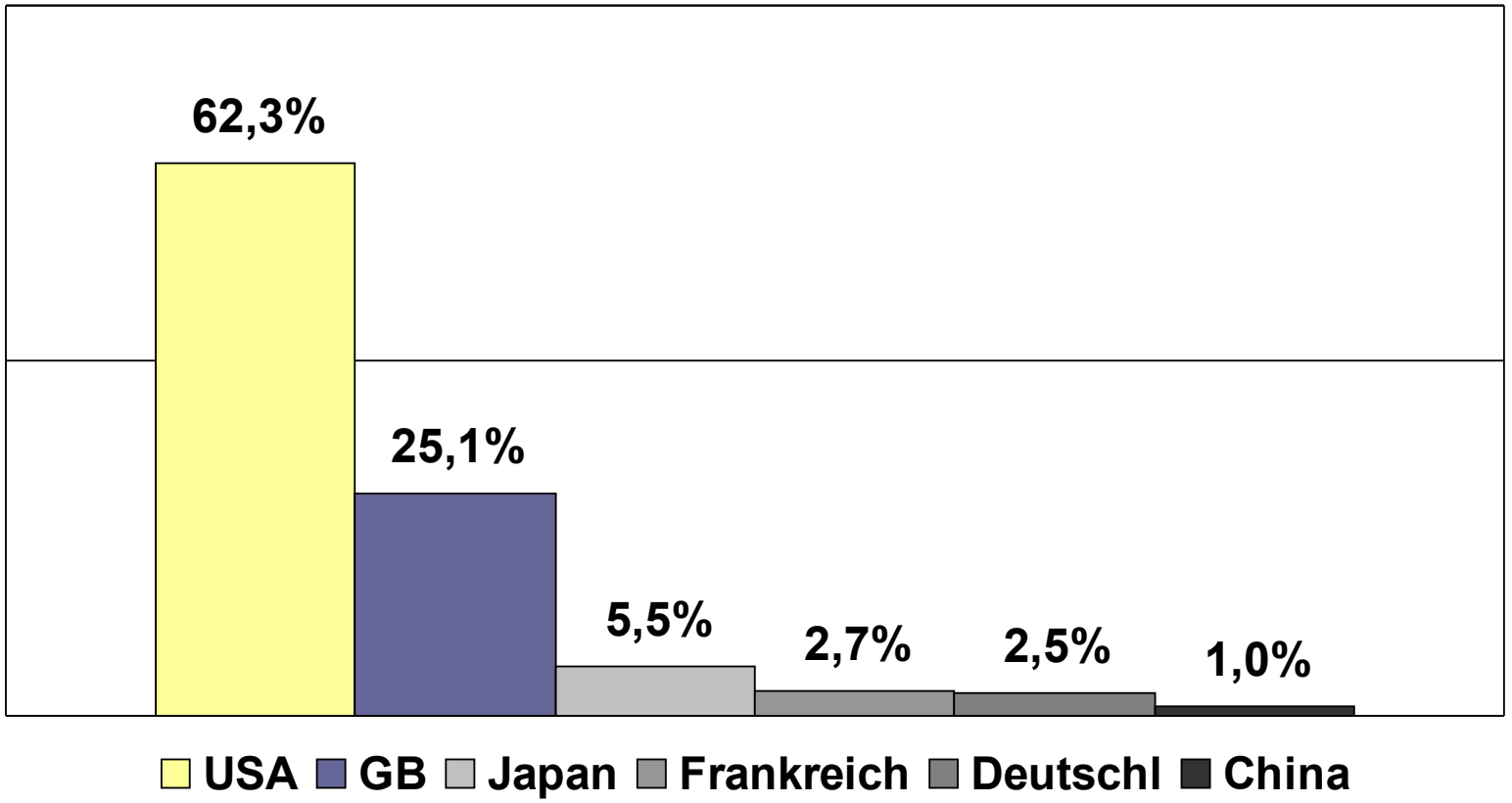
Beiträge der Teilnehmer



IHGSC (2001), Nature 409: 860

Fertigstellen des Human-Genoms

Beiträge der nationalen Genomprojekte



Sequenzierung des Human-Genoms

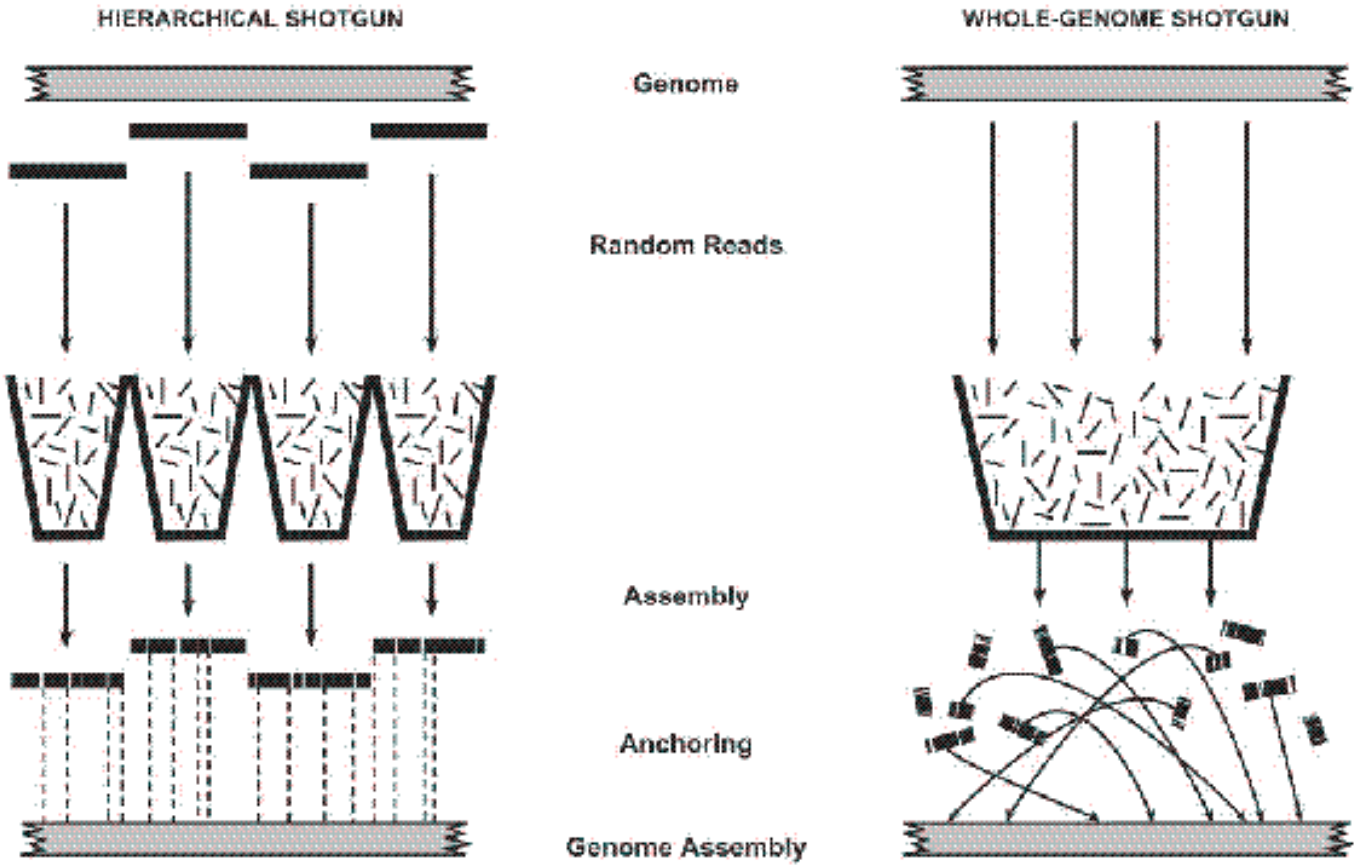
Publikationen der Rohfassung 2001

Internationales akademisches Konsortium

Privates Unternehmen Celera

Schrotschuß-Strategien

hierarchisch vs. gesamt-genomisch



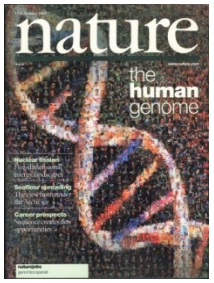
2.500 Teile

Puzzle

60.000.000 Teile

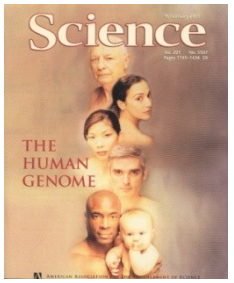
Human Genome

Working Draft versions February 2001



Academic

Private



Initial Sequencing & Analysis…		The Sequence of …
2.72 Gb	Sequenced Bases	2.65 Gb
1,000	Clone gaps	54,000
146,000	Sequence gaps	116,000
147.000	Gaps	170,000

overall coverage:	94%
quality of unfinished data :	< 1 error/10kb in 91%

human population heterogeneity:	1 SNP/kb
variation between two individuals:	1 SNP/10kb

Analyse der Human-Genom-Rohfassung

Besonderheiten

30.000 -35.000 Gene
erwartet 28.000 - 140.000
Ackerschmalwand 25.700, Fadenwurm 18.300,
Fruchtfliege 13.300

komplexere Transkription
60% aller Gene;
mRNAs/Gen: Chr22=2,6, Chr19=3,2
Fadenwurm: 22% aller Gene; 1,3 mRNAs/Gen

komplexere Proteinarchitektur
keine Neuen Domänen,
"Auflaufen" (accretion) zusätzlicher Domänen
an den Gen-Enden

Gentransfer von Prokaryonten
233 Gene mit Ähnlichkeit nur zu Prokaryonten
davon 133 bei Prokaryoten weit verbreitet

ungleichmäßige Genverteilung
25% des Genoms sind "genlose Wüste"
Gegensatz zu Ackerschmalwand, Fadenwurm, Fruchtfliege

50% repetitive Sequenzen
Ackerschmalwand 11%, Fadenwurm 7%, Fruchtfliege 3%
"fossiles Archiv" der genomischen Evolution
dramatische Reduktion der Akkumulation
in den letzten 50 MioJ
SINEs als Symbionten von Genen (Chr19=57%)

Mutationsraten Frau:Mann 1:2
Vergleich X und Y Chr

ungleichmäßige SNP-Verteilung
63% aller 5kb-Fragmente enthalten SNP(s)
interindividueller Unterschied 0.01%

IHSC (2001), Nature 409: 860

Fertigstellen des Human-Genoms

Nature, 21. Oktober 2004

Finishing the euchromatic sequence of the human genome

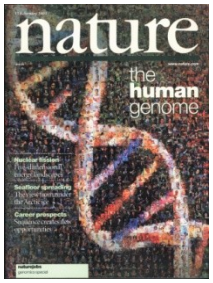
International Human Genome Sequencing Consortium*

* *A list of authors and their affiliations appears in the* Supplementary Information

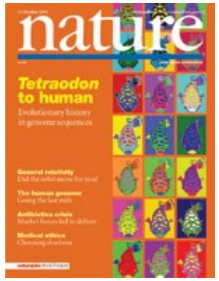
The sequence of the human genome encodes the genetic instructions for human physiology, as well as rich information about human evolution. In 2001, the International Human Genome Sequencing Consortium reported a draft sequence of the euchromatic portion of the human genome. Since then, the international collaboration has worked to convert this draft into a genome sequence with high accuracy and nearly complete coverage. Here, we report the result of this finishing process. The current genome sequence (Build 35) contains 2.85 billion nucleotides interrupted by only 341 gaps. It covers ~99% of the euchromatic genome and is accurate to an error rate of ~1 event per 100,000 bases. Many of the remaining euchromatic gaps are associated with segmental duplications and will require focused work with new methods. The near-complete sequence, the first for a vertebrate, greatly improves the precision of biological analyses of the human genome including studies of gene number, birth and death. Notably, the human enome seems to encode only 20,000–25,000 protein-coding genes. The genome sequence reported here should serve as a firm foundation for biomedical research in the decades ahead.

Human Genome

Final version October 2004



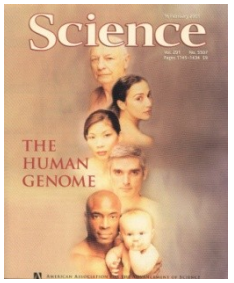
Initial Seq…



Academic

Finishing the euchromatic sequence…

Private



The Sequence of …

2.72 Gb	2.85 Gb	Sequenced Bases	2.65 Gb
1,000	283	Clone gaps	54,000
146,000	58	Sequence gaps	116,000
147.000	341	Gaps	170,000

near-complete sequence:	99% of euchromatin
extremely high quality:	< 1 error/100kb

Analyse des fast vollständigen Human-Genoms

Besonderheiten

fast vollständige Sequenz

2,85 Mb, 99% des Euchromatins

genreich, fast alle Lebensprozesse

nur noch 341 Lücken

33 heterochromatisch (Centromere)

308 euchromatisch (28 Mb),
50% mit segmentalen Duplikationen

extrem hohe Genauigkeit

99.999%

ein Fehler auf 100.000 Bausteine

20.000 -25.000 Gene

erwartet 28.000 - 140.000

Pflanze 25.700, Wurm 18.300,Fliege 13.300

verbesserte Gen-Vorhersagen

58% aller Gene der Rohfassung waren fehlerbehaftet

neue & verlässliche Analysen

Pseudogene

1.183 neue entstandene Gene

Geruchsrezeptoren, Immunität, Schwangerschaft

37 ‘kürzlich gestorbene’ Gene

10 Geruchsrezeptoren

Segmental Duplications

Problems of the human reference sequence

~50%

of the

273 interior euchromatic gaps

located in

segmentally duplicated regions

Segmental Duplication

Definition

genomic regions >1kb

with nt identity >90%

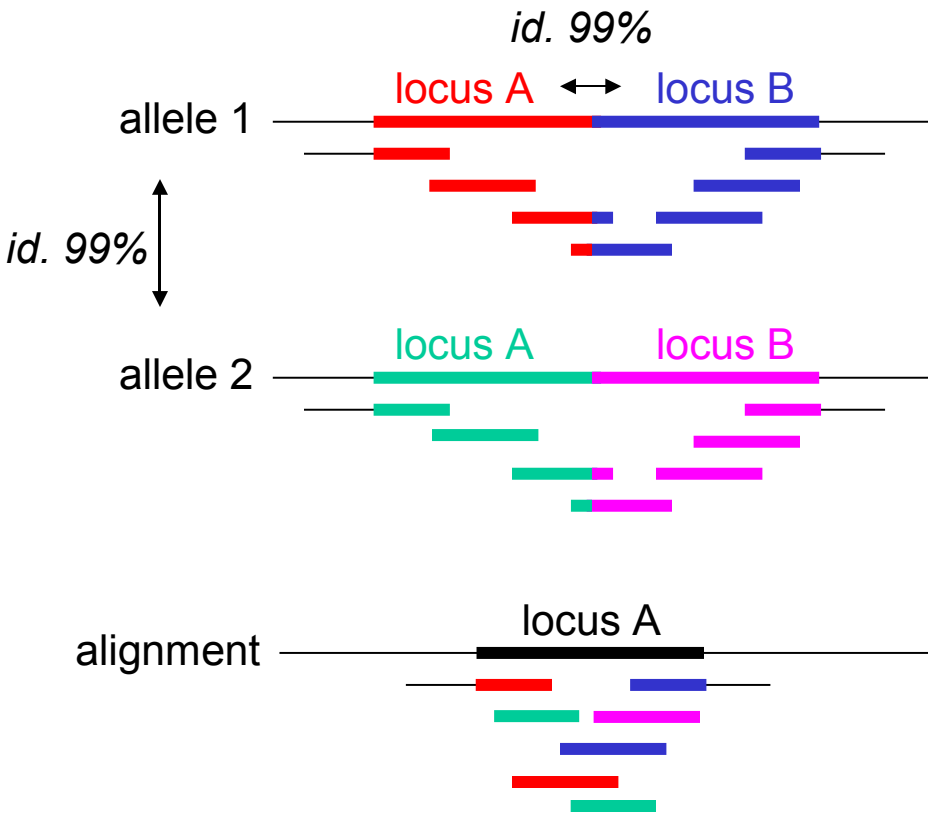
Human genome

5.3% segmentally duplicated

87% of all segmental duplications >50 kb

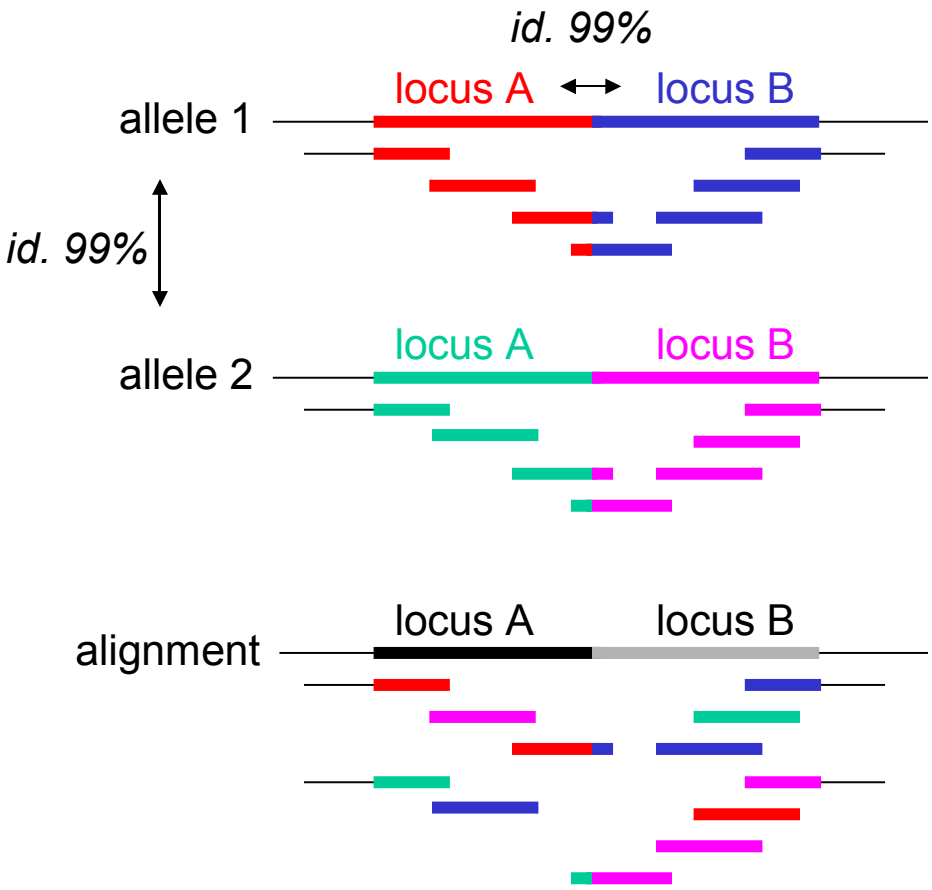
Segmental Tandem Duplications

Assembly problems



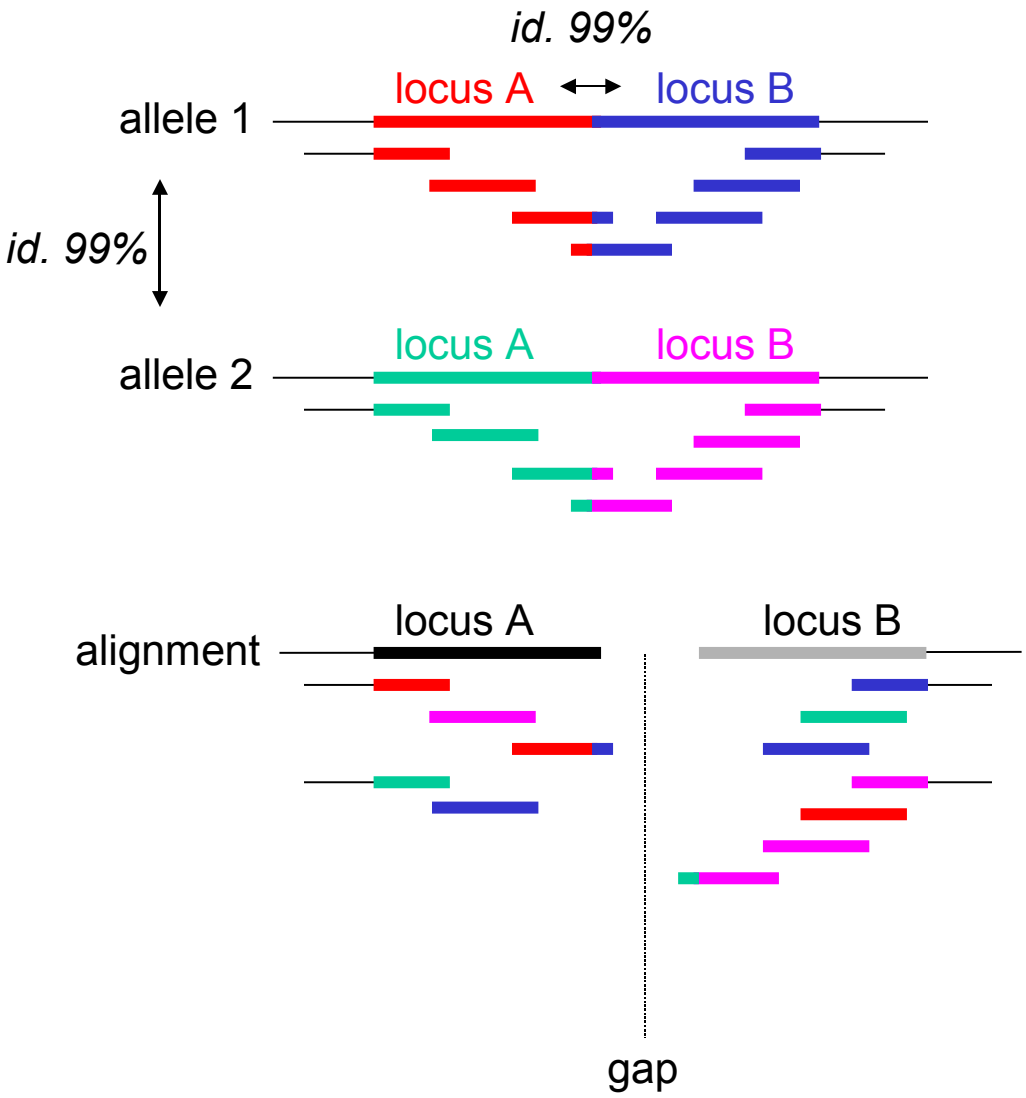
Segmental Tandem Duplications

Assembly problems



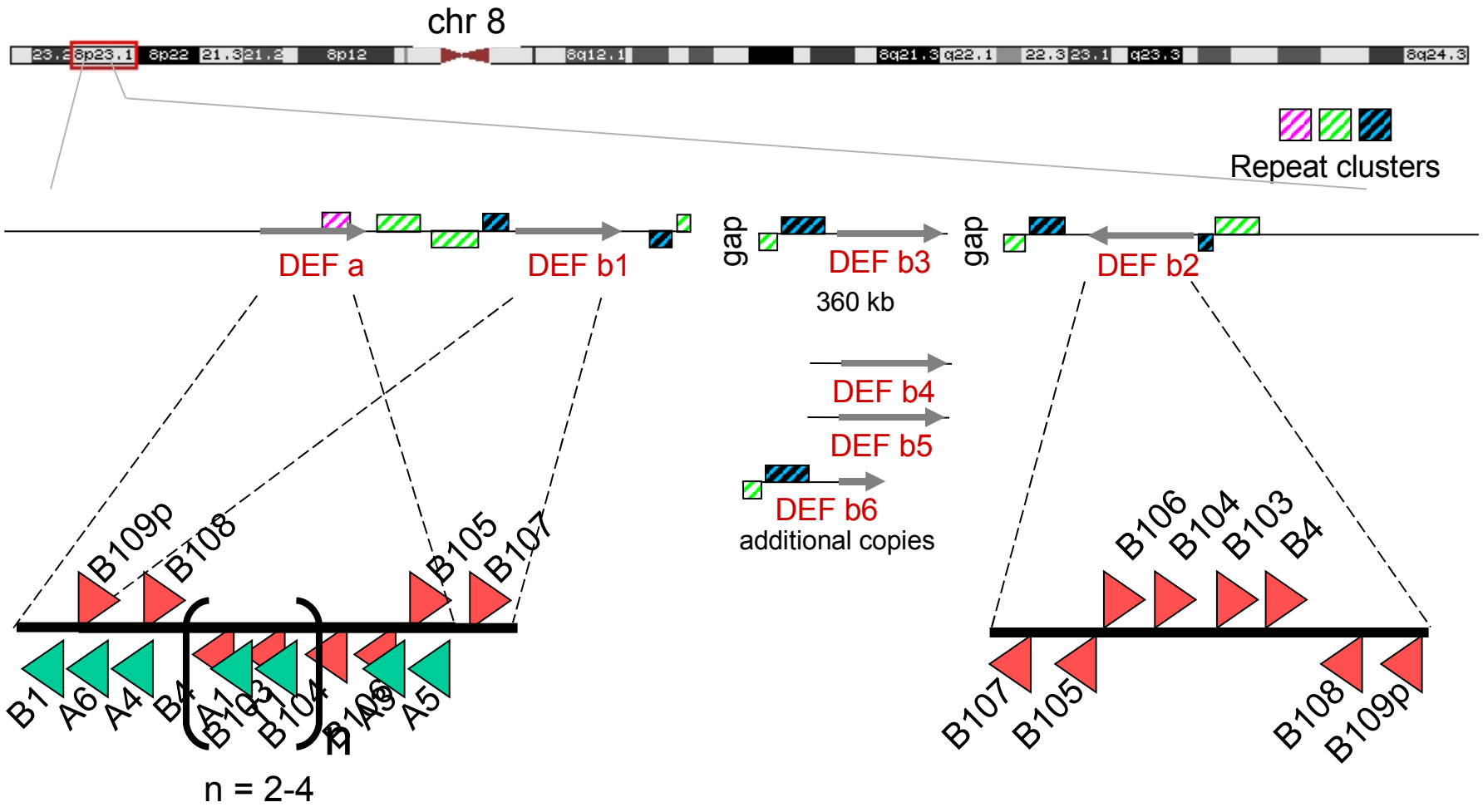
Segmental Tandem Duplications

Assembly problems



DEF cluster at 8p23.1

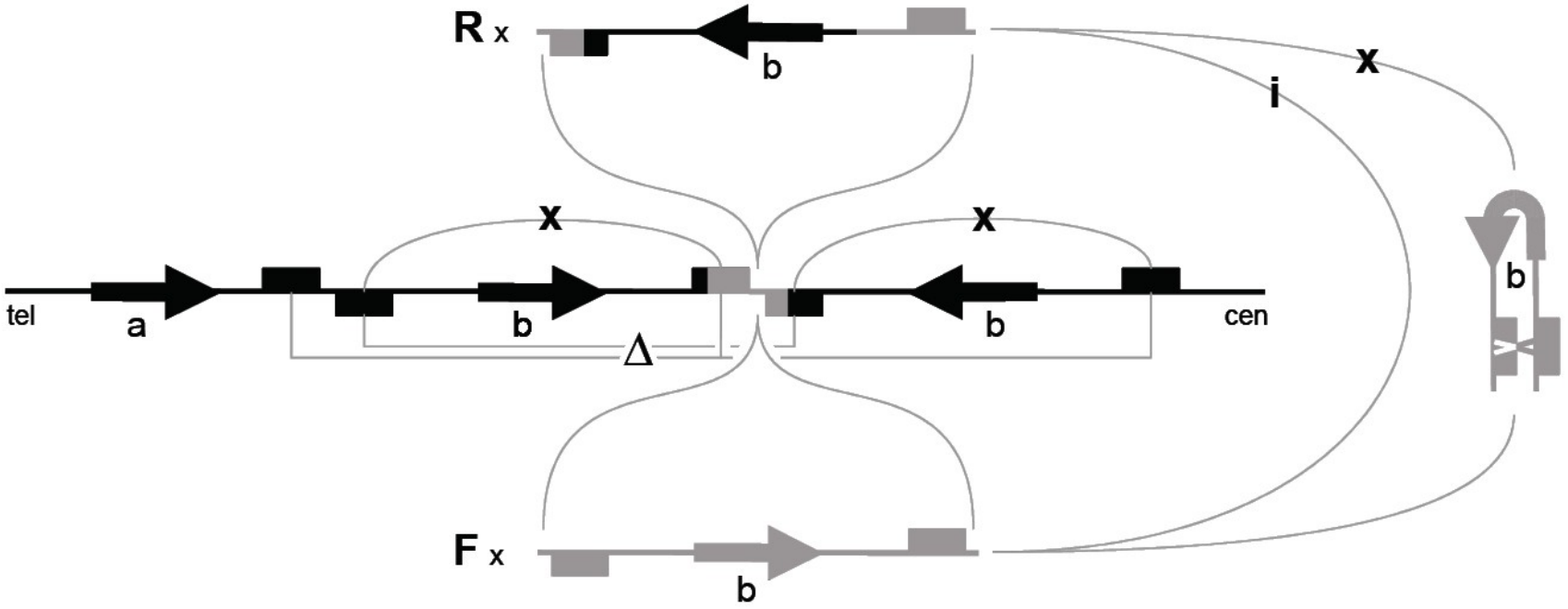
hg16: 6.3-8.3 Mb



Taudien et al., BMC Genomics 5:92 (2004)

Genomic variability of 8p23.1 DEF locus

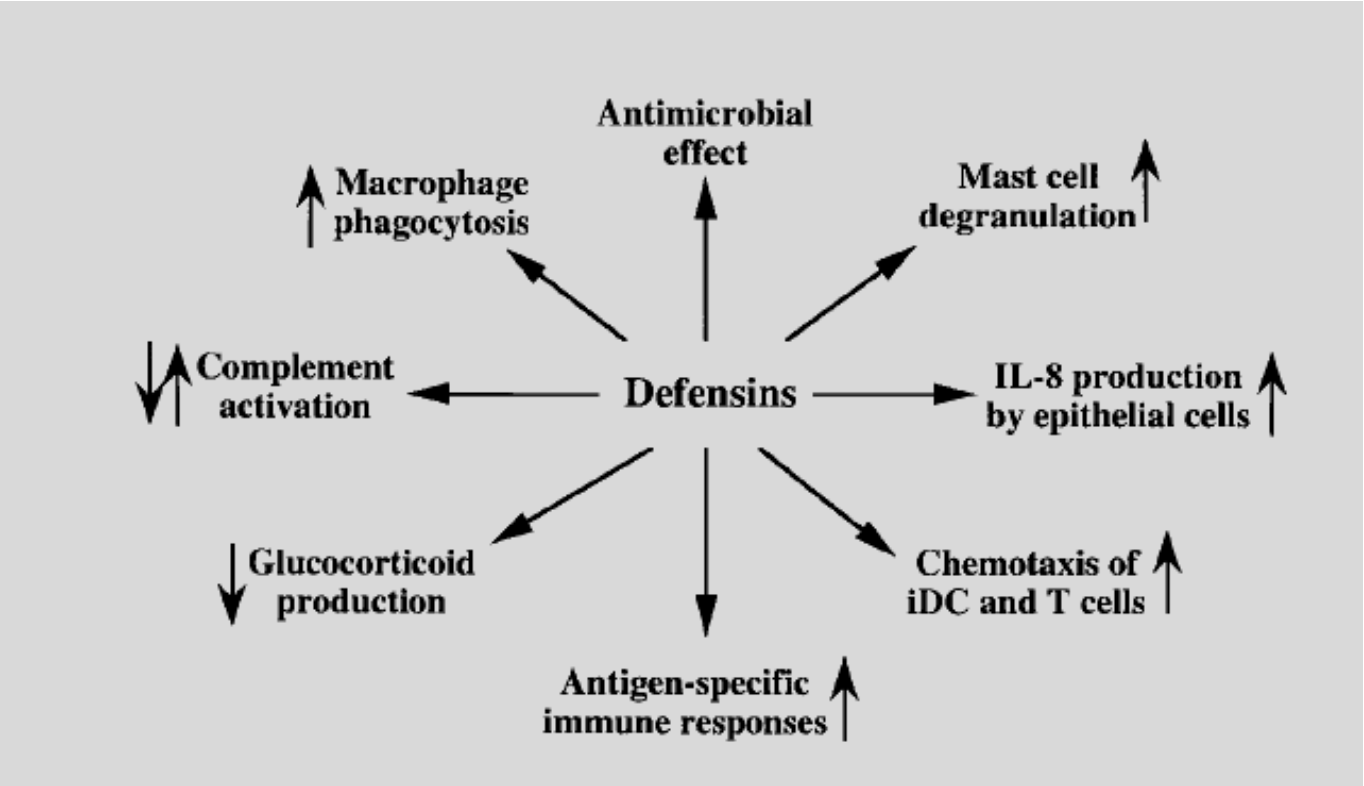
Hypothetical organisation



Taudien et al., BMC Genomics 5:92 (2004)

Defensine (DEF)

Multiple Funktionen



Immunität & Krebs

genetische Assoziation mit CED 2007, Psoriasis 2008, Prostata-Karzinom 2008

Yang et al., 2001. Cell. Mol. Life Sci. 58, 978-989

Complex phenotypes / diseases

Structural variations

Chr 17 900 kb inversion polymorphism & female fertility and recombination rates in humans

Nat Genet 37:129 (2005)

CCL3L1 copy number & AIDS susceptibility in humans

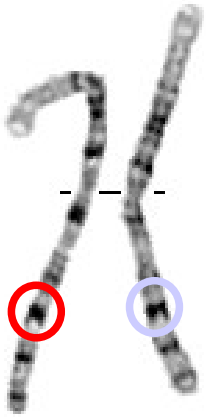
Science 307:1434 (2005)

FCGR3 copy number & glomerulonephritis in humans and rats

Nature 439:851 (2006)

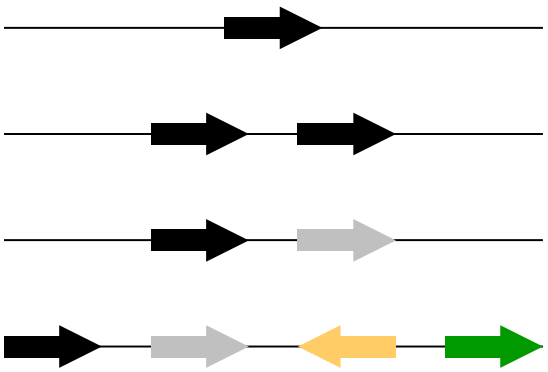
Humangenom-Dynamik

Zusammenfassung



Chromosom A

Chromosom B

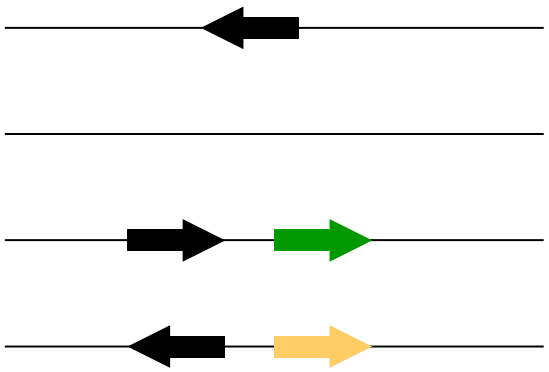


Inversion

InDel

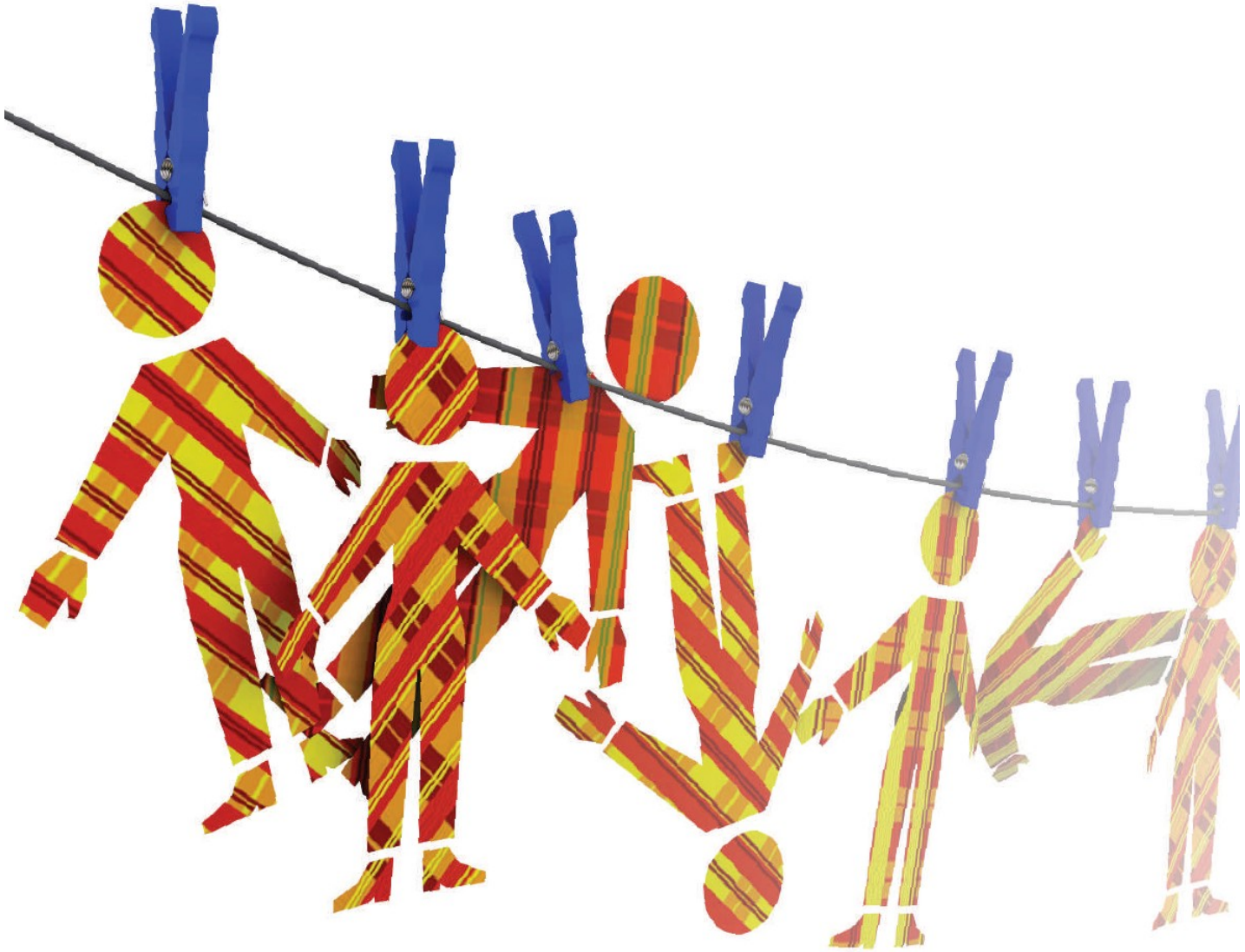
Allel-Variation

Kombination



Humangenom-Dynamik

Patchwork people ?



News Feature, Nature 437:1084 (2005)

Genetische Variation

Zusammenfassung

Genomes of any two individuals in the human population differ more at the structural level than at the nucleotide sequence level.

Differences between individuals

- CNV: >4 Mb >1/800 bp > 0.12 %
- SNP: 2.5 Mb 1/1,200 bp 0.08 %

Sebat, Nat Gen Suppl 39:s3 2007

Gene und Krankheit

monogene - polygene - multifaktorielle Erkrankungen

We used to think our fate
was in our stars.
Now we know,
in large measure,
our fate is in our genes

James Watson, 1989

Relatives Erkrankungsrisiko für Geschwister (λ_s)

750 Mukoviszidose

40 chronisch-entzündliche Darmerkrankungen
multiple Sklerose

20

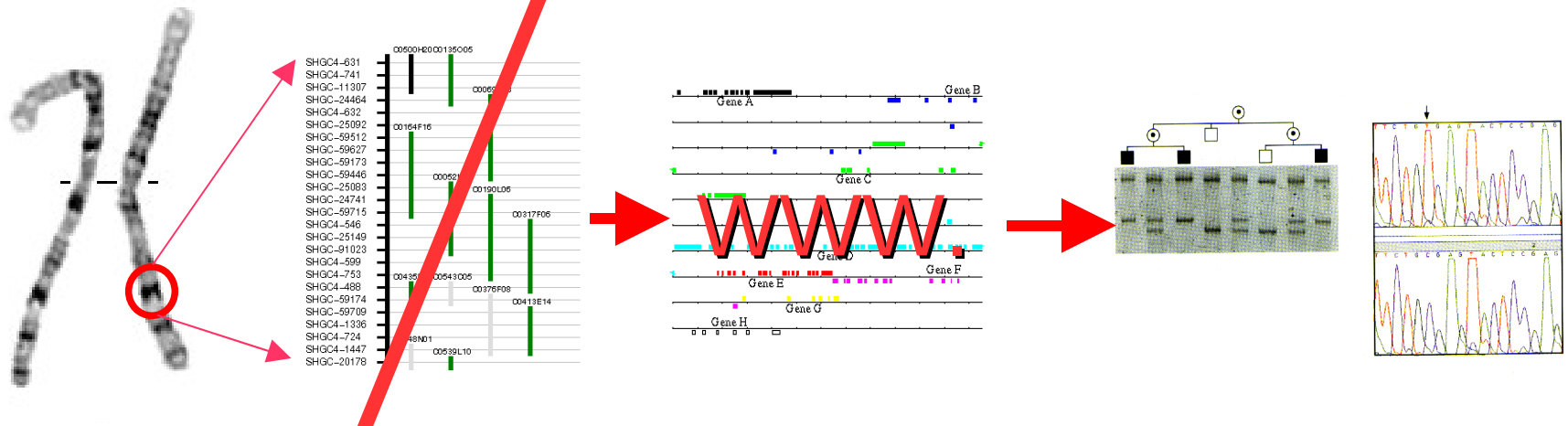
Schizophrenie
10 Psoriasis
Diabetes Typ I

5 Asthma

2,5 Bluthochdruck

Krankheitsgene

klassische Positionsklonierung



Kopplungsanalyse
in betroffenen Familien

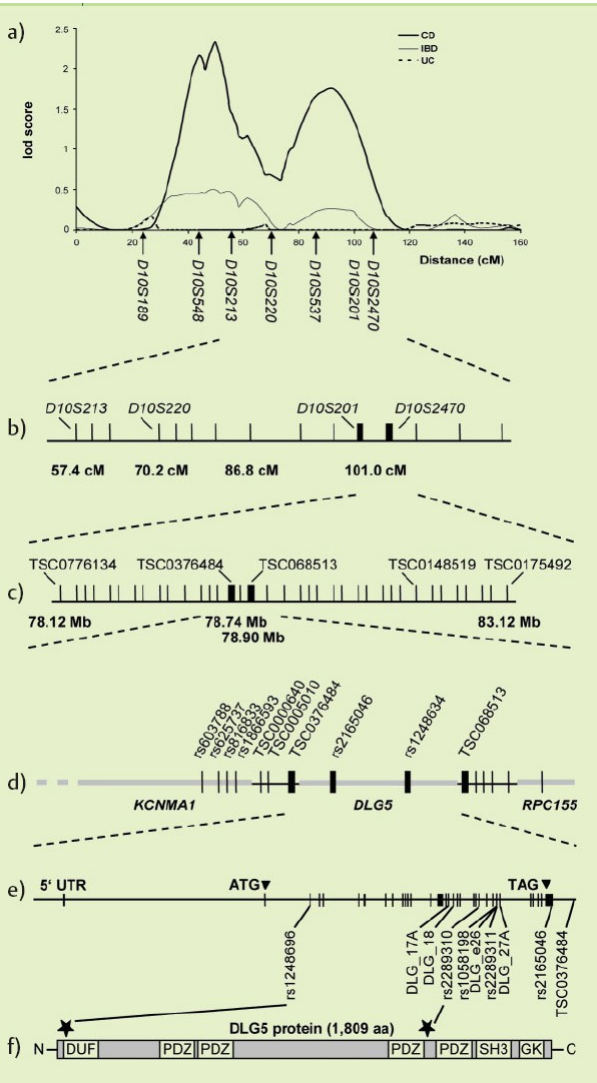
Physische Kartierung
der Region

Identifikation
aller Gene

Mutationsanalyse von
Kandidatengenen
in Betroffenen und Kontrollen

Polygene Erkrankungen

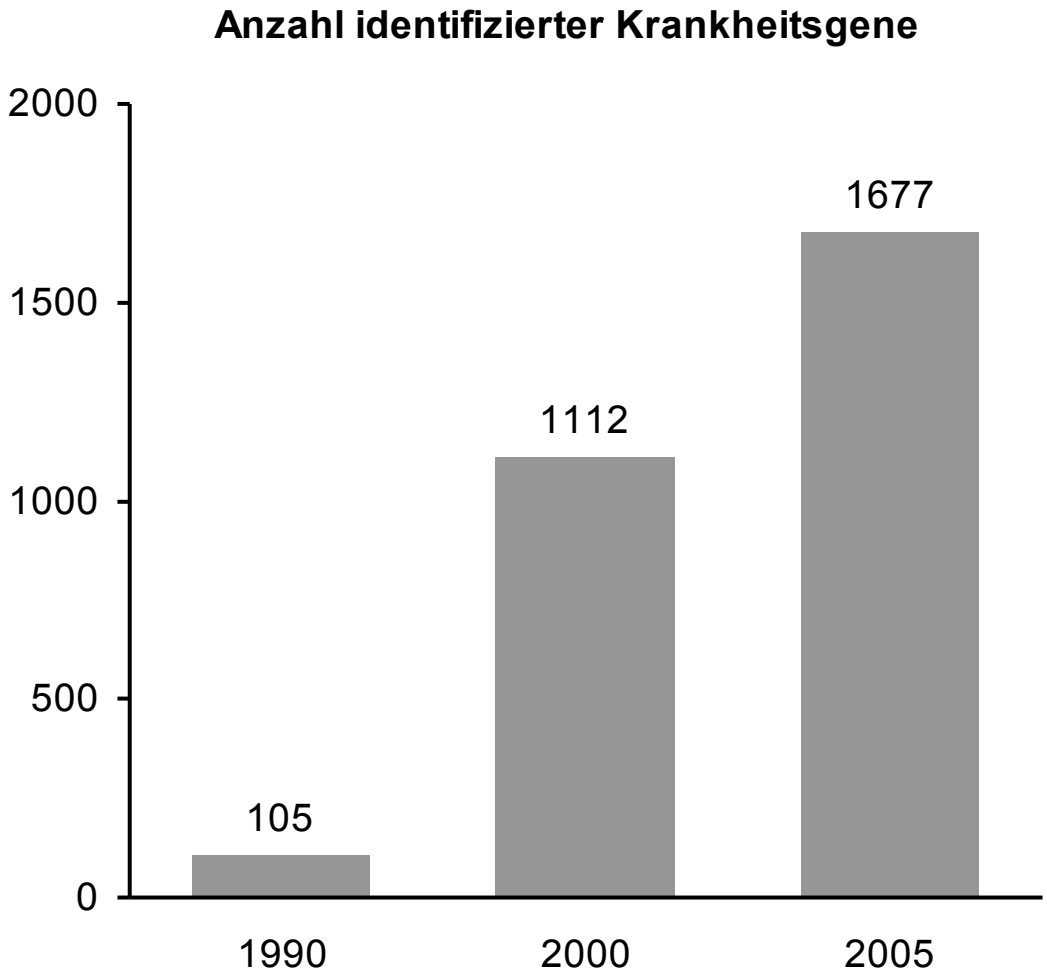
CED: Chromosom 10, *DLG5*



1. initiale Kopplungsanalyse mit Mikrosatelliten
2. Feinkartierung mit weiteren Mikrosatelliten
3. Feinstkartierung mit SNPs
4. Assoziation mit SNPs im *DLG5*
5. Identifizierung von 2 proteinverändernden SNPs

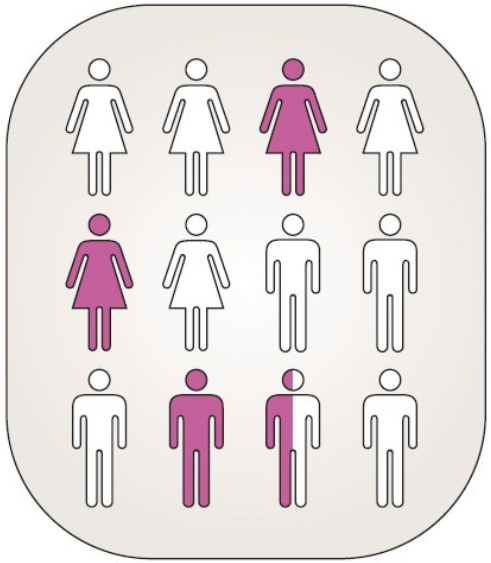
Krankheitsgene

Aufklärungsrate

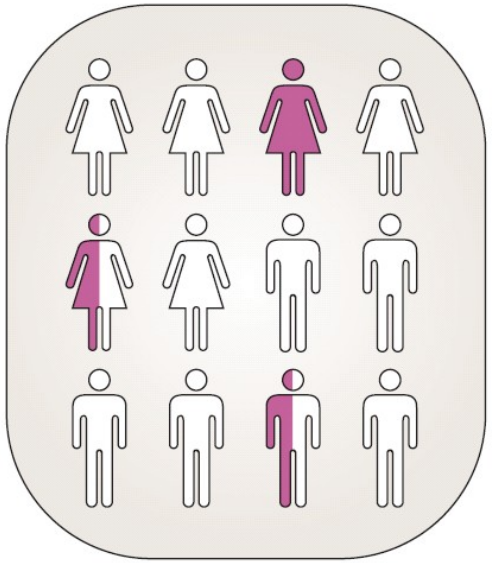


GenomWeite AssoziationsStudien

GWAS



Disease cases



Controls

Mathew. Nat Rev Genet 9:9 2008

GWAS prerequisites

Large, well-phenotyped study groups

WTCCC Wellcome Trust Case Control Consortium (GB)

17,000 samples

2,000 from each of seven diseases

type 1 diabetes, type 2 diabetes, coronary heart disease, hypertension, bipolar disorder, rheumatoid arthritis, Crohn's disease

3,000 controls also from England, Scotland and Wales

KORA Kooperative Gesundheitsforschung in der Region Augsburg

20,000 samples since 1985 *coronary heart disease*

POPGEN Schleswig-Holstein Biobank für eine Medizin der Zukunft

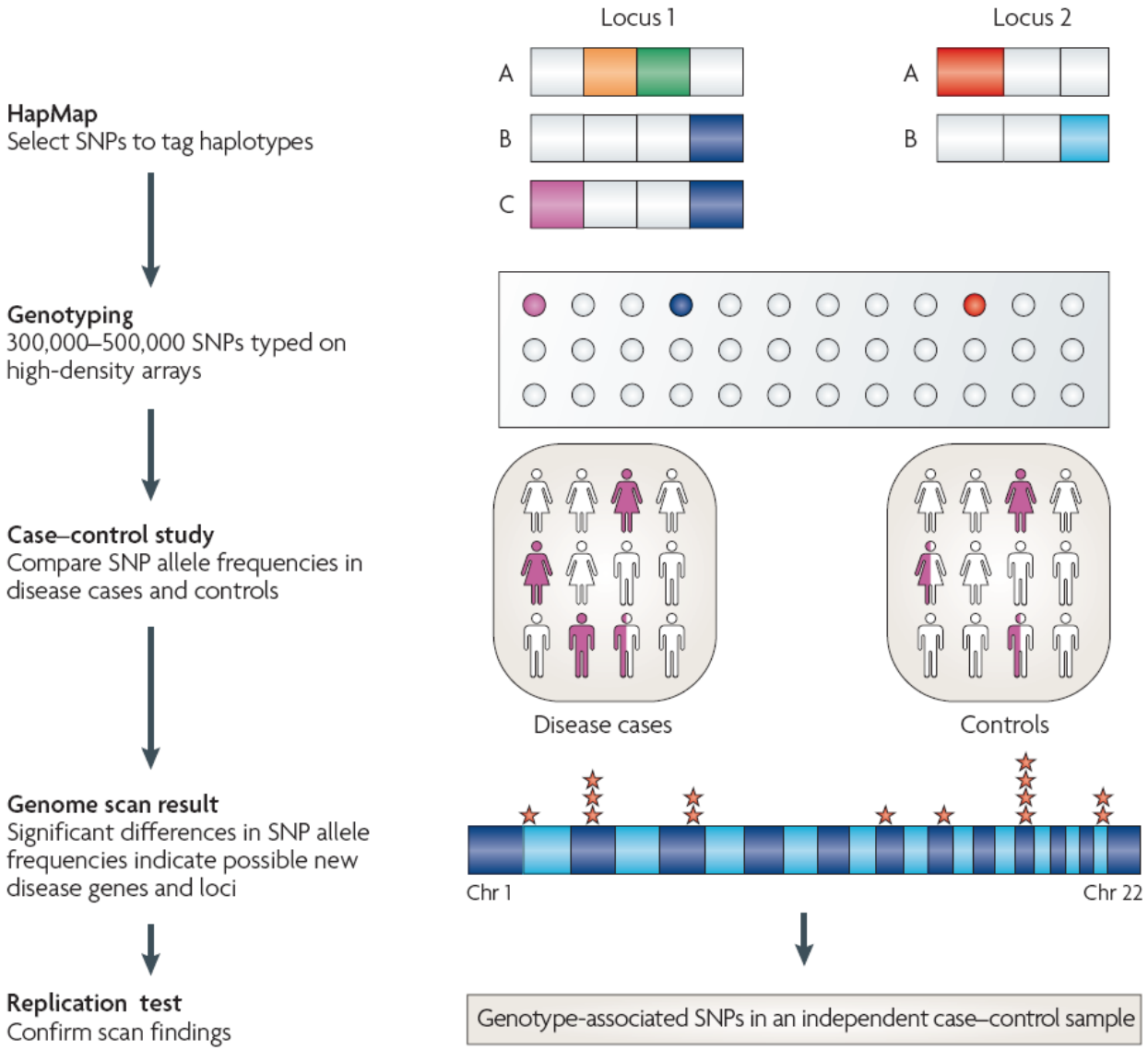
since 2003 aiming at 30,000 controls + study groups for:

aging *3,000 centenarians*

diseases *inflammation, heart, cancer, nervous system*

GWAS

Schema



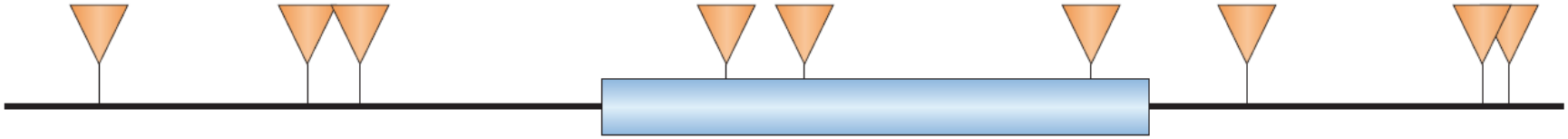
Mathew. Nat Rev Genet 9:9 2008

GWAS

Alternative designs

Direct:
catalogue and test all functional variants for association

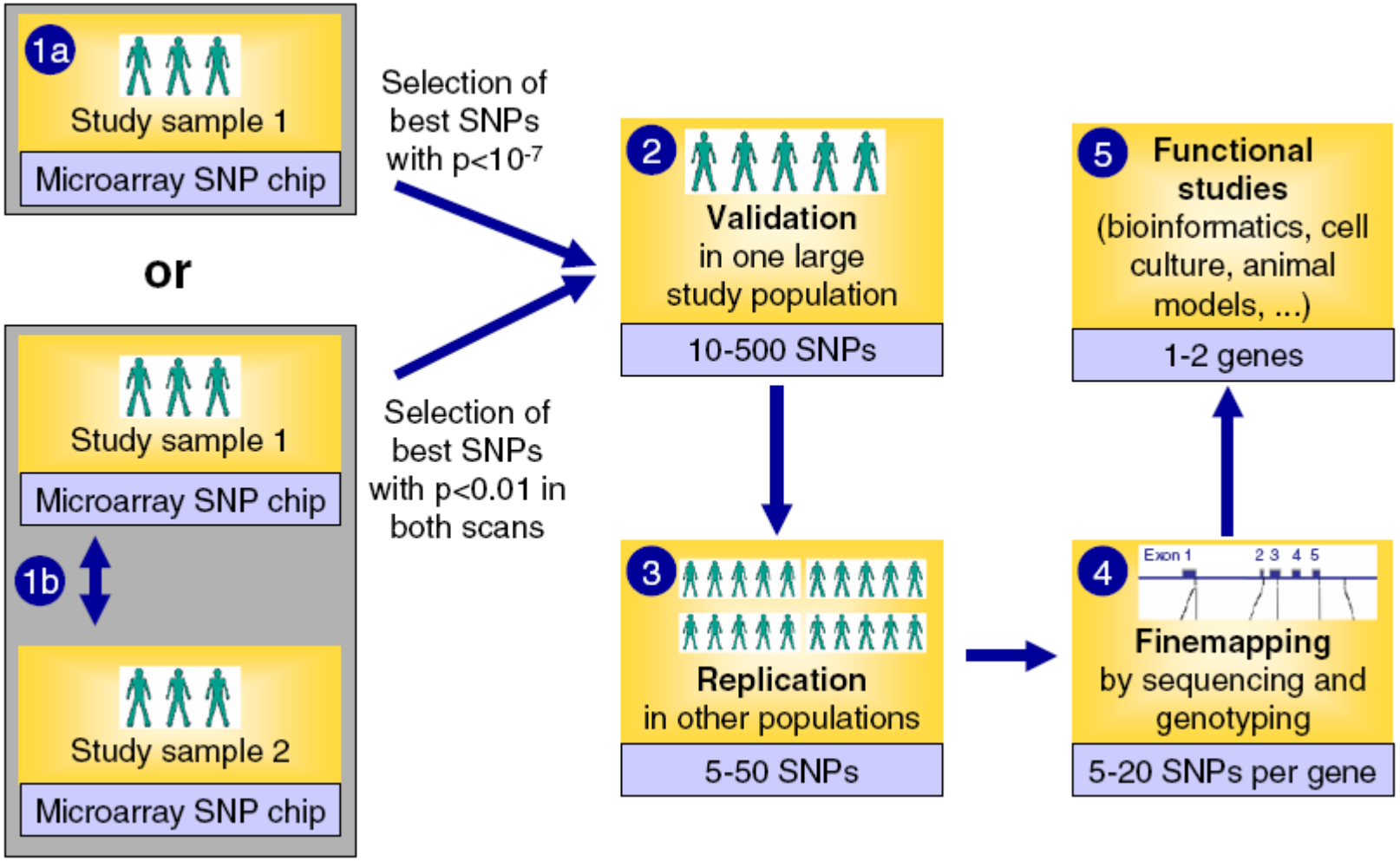
Indirect:
use a dense SNP map and test for linkage disequilibrium



Kruglyak. Nat Rev Genet 9:314 2008

GWAS

Statistical design & follow-up



Kronenberg. Exp Geront 43:39 2008

GWAS

Current stage 06/2009

carried-out GWAS: 180
within last 12 month: 100

Identified many novel genes involved in complex diseases.

Some genes are associated with several phenotypes.

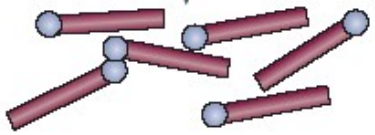
DNA-Microarray-(Chip)Technologie

Genotypisierung & Expressionsanalyse

Genomische DNA

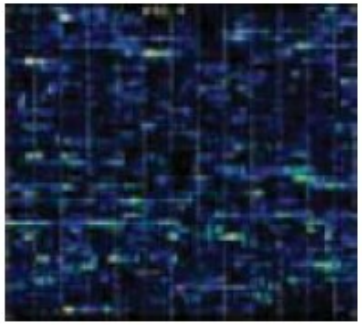
Vervielfältigung
(PCR)
Markierung

Fragmente



Hybridisierung

Analyse



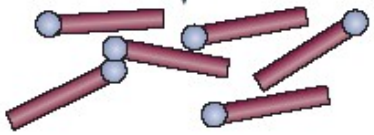
mRNA

cDNA-Synthese
reverse Transkription

cDNA

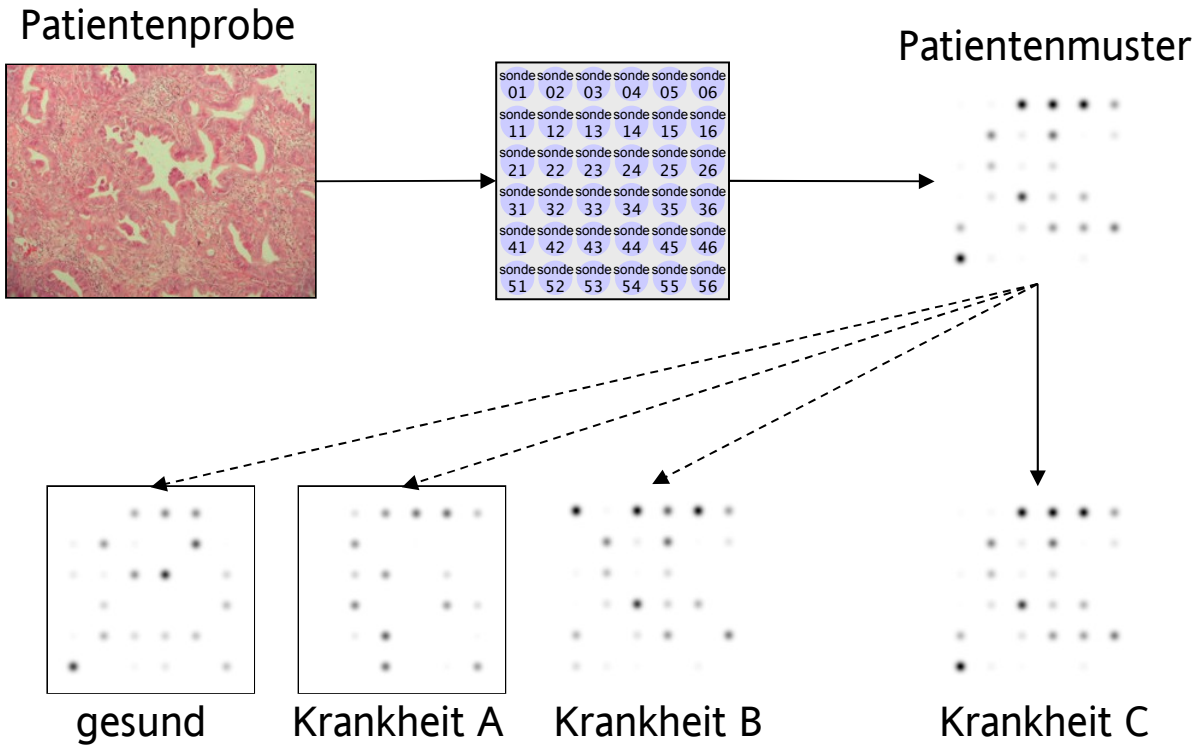
Vervielfältigung
(PCR)
Markierung

Fragmente



DNA-Microarray-(Chip)Technologie

Genexpression in Diagnostik & Therapie



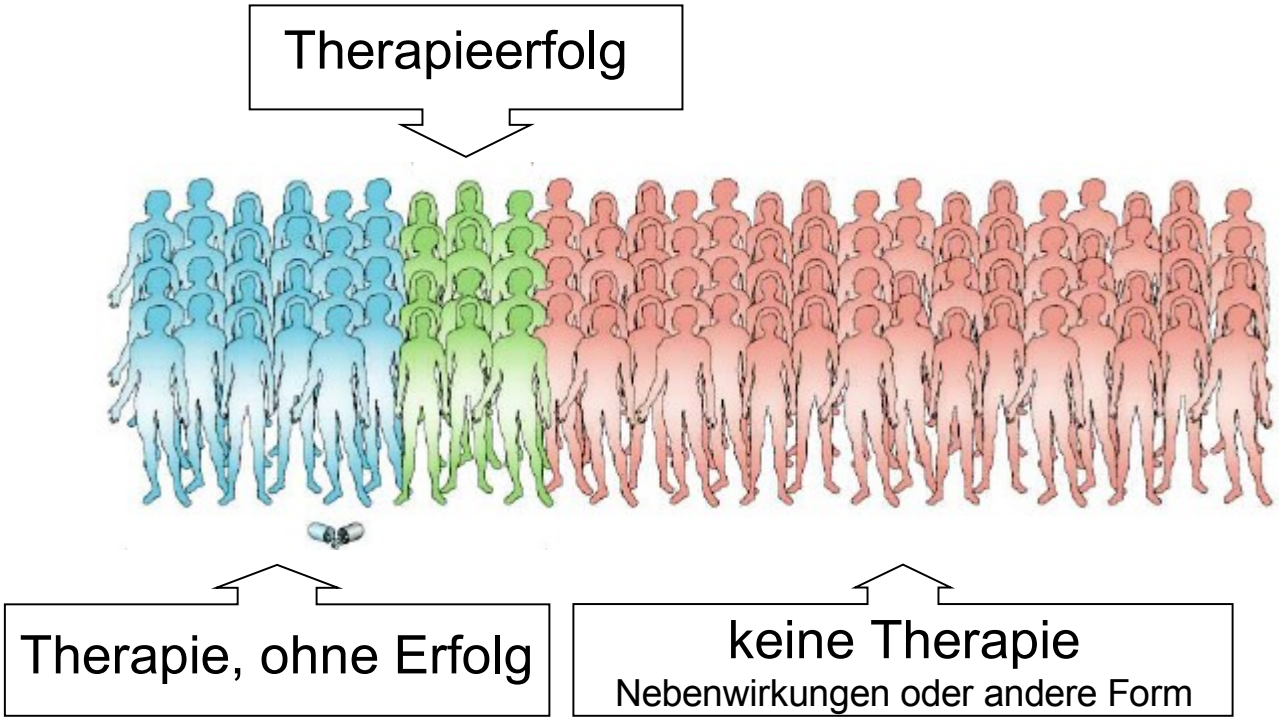
Mustervergleich

Diagnose

- Identifizierung von Zielgenen für Therapie, Verlaufskontrolle

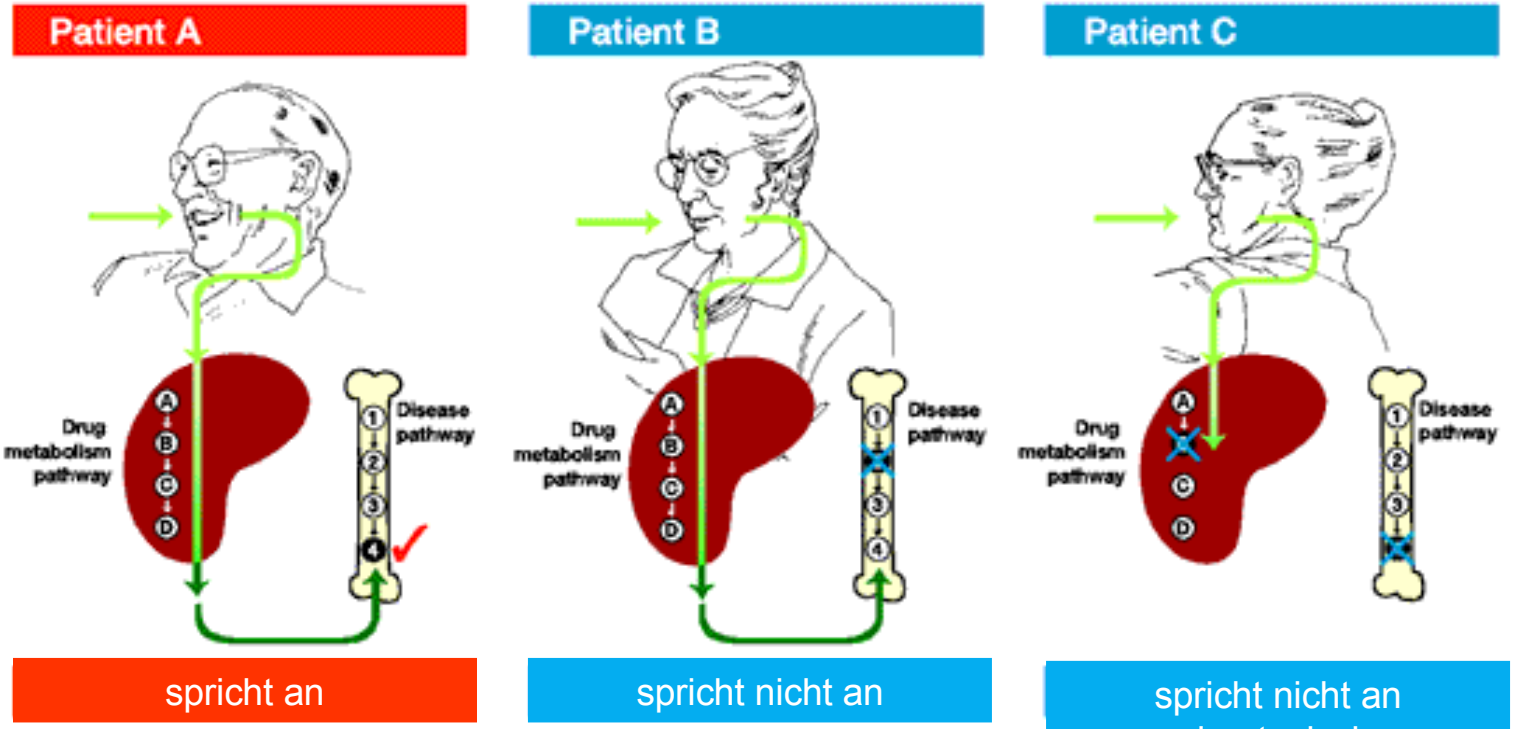
Pharmakogenetik

Problem



Pharmakogenetik

Mechanismen variabler Arzneimittelwirkungen



spricht an

spricht nicht an

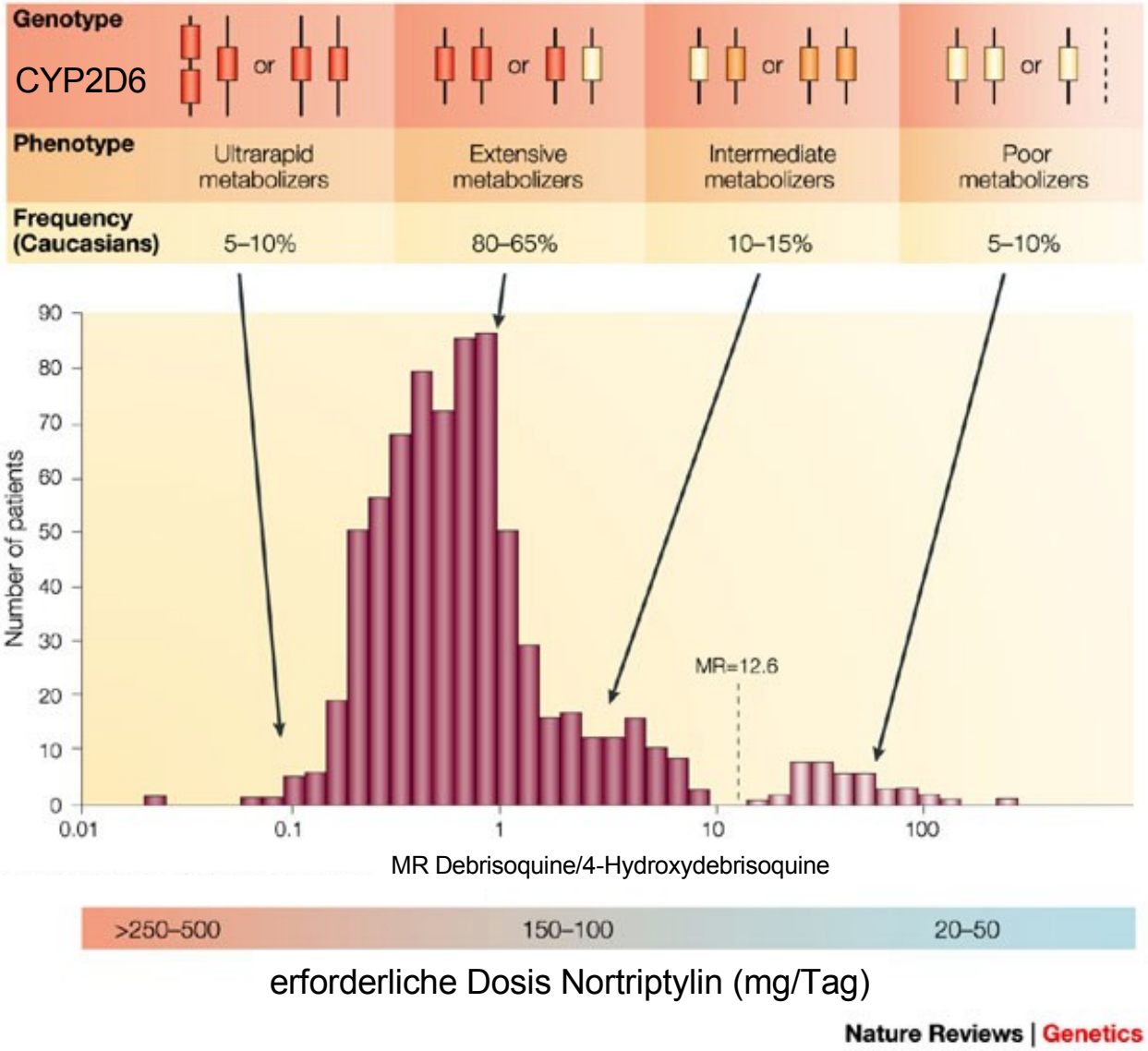
Gen im Wirkmechanismus des Arzneimittels
- Transport
- Rezeptor

spricht nicht an oder toxische Nebenwirkungen

Gen im Arzneimittelmetabolismus

Pharmakogenetik

Zytochrom P450



AmpliChip von Roche

Bestimmung der Allele von 2 Zytochrom P450 Genen

Berechnung der individuellen optimalen Arzneimitteldosis

Individualisierte Medizin

Nutzen & Risiken

Pharmaka für Subpopulationen (Genotypen)
optimale Dosis - geringe Nebenwirkungen

Einengung der Pharmamärkte
Benachteiligung von Trägern seltener Genotypen

HGP: Was bleibt zu tun?

Grundlagenforschung

Funktionsanalyse aller Gene

Identifizierung von Krankheitsgenen

Vergleichende [Epi-]Genomik

zwischen Spezies

Sequenzierung von weiteren Modell-Genomen

zwischen Populationen

Variabilität - Internationales Projekt HGDiversityP

zwischen Individuen

Variabilität - Internationales Projekt HapMap

zwischen Zellen/Geweben

Bedeutung der 5. Base mC für Differenzierung & Krankheiten-Human Epigenom Project

Technologieentwicklung

Sequenzierung & Typisierung

HGP: Was bleibt zu tun?

Anwendungsforschung

Testentwicklung & Validierung

genbezogen
krankheitsbezogen
genomweit

Technologietransfer

Sequenzierung & Typisierung
in der ärztlichen & klinischen Routine

HGP: Was bleibt zu tun?

Gesellschaftlicher Rahmen & Konsens

Recht auf Nichtwissen

Recht auf Nichtinformiertheit

Recht auf Wissen, dessen Datenschutz und Finanzierung

Verhinderung genetischer Fremdbestimmung

Bundesrat **Drucksache 374/09**

24.04.09

G

Gesetzesbeschluss
des Deutschen Bundestages

Gesetz über genetische Untersuchungen bei Menschen (Gendiagnostikgesetz - GenDG)

Der Deutsche Bundestag hat in seiner 218. Sitzung am 24. April 2009 aufgrund der Beschlussempfehlung und des Berichts des Ausschusses für Gesundheit – Drucksache 16/12713 – den von der Bundesregierung eingebrachten

Entwurf eines Gesetzes über genetische Untersuchungen bei Menschen (Gendiagnostikgesetz – GenDG)
– Drucksachen 16/10532, 16/10582 –

in beigefügter Fassung angenommen.

Gediagnostikgesetz

GenDG

§ 1

Zweck des Gesetzes

Zweck dieses Gesetzes ist es, die Voraussetzungen für genetische Untersuchungen und im Rahmen genetischer Untersuchungen durchgeführte genetische Analysen sowie die Verwendung genetischer Proben und Daten zu bestimmen und eine Benachteiligung auf Grund genetischer Eigenschaften zu verhindern, um insbesondere die staatliche Verpflichtung zur Achtung und zum Schutz der Würde des Menschen und des Rechts auf informationelle Selbstbestimmung zu wahren.

Gediagnostikgesetz

GenDG

§ 4

Benachteiligungsverbot

(1) Niemand darf wegen seiner oder der genetischen Eigenschaften einer genetisch verwandten Person, wegen der Vornahme oder Nichtvornahme einer genetischen Untersuchung oder Analyse bei sich oder einer genetisch verwandten Person oder wegen des Ergebnisses einer solchen Untersuchung oder Analyse benachteiligt werden.

Gediagnostikgesetz

GenDG

§ 9

Aufklärung

5. das Recht der betroffenen Person auf Nichtwissen einschließlich des Rechts, das Untersuchungsergebnis oder Teile davon nicht zur Kenntnis zu nehmen, sondern vernichten zu lassen,

Gediagnostikgesetz

GenDG

§ 11

Mitteilung der Ergebnisse genetischer Untersuchungen und Analysen

(3) Die verantwortliche ärztliche Person darf das Ergebnis der genetischen Untersuchung oder Analyse anderen nur mit ausdrücklicher und schriftlicher Einwilligung der betroffenen Person mitteilen.

Gediagnostikgesetz

GenDG

§ 18

Genetische Untersuchungen und Analysen im Zusammenhang mit dem Abschluss eines Versicherungsvertrages

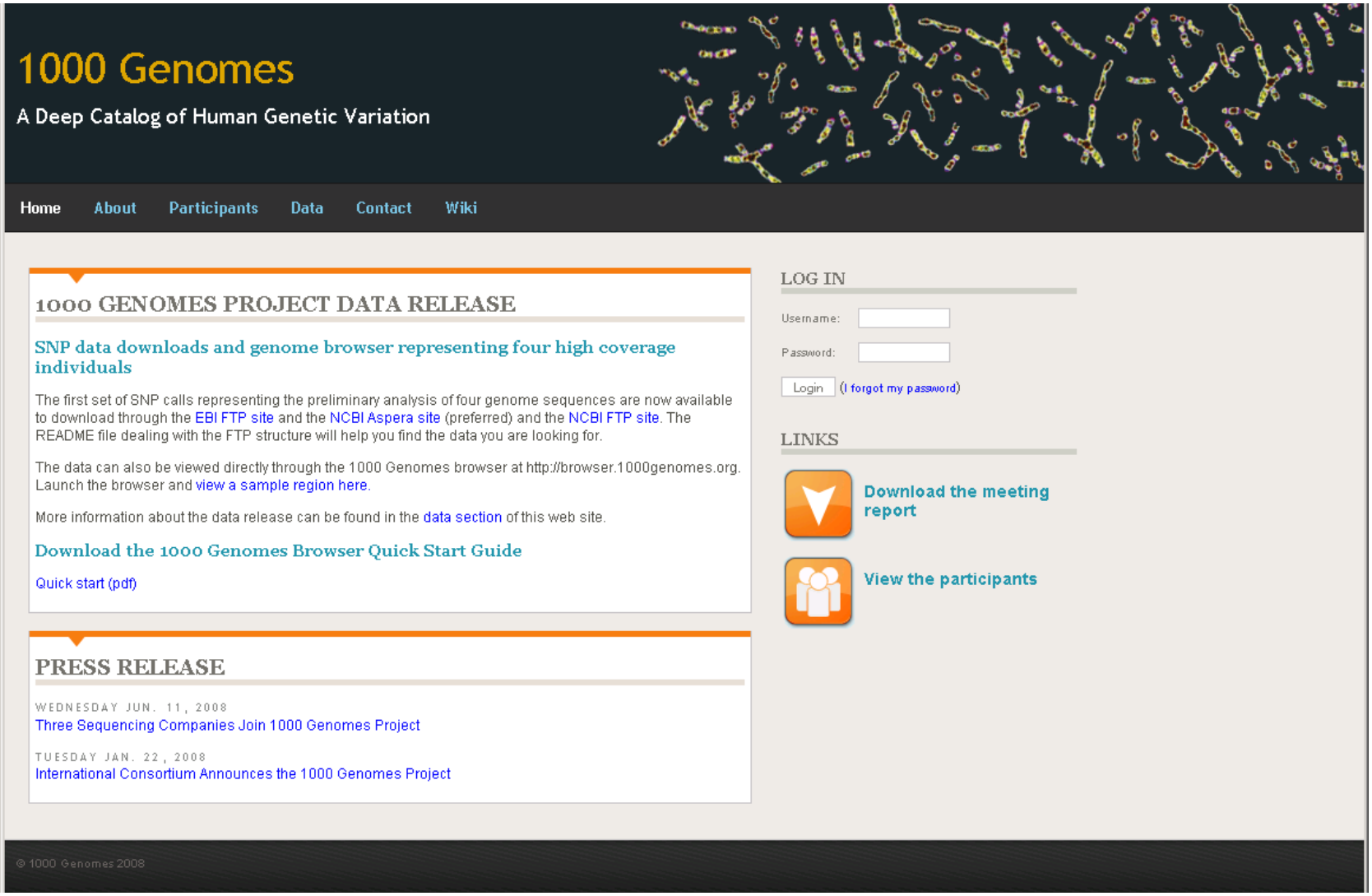
(1) Der Versicherer darf von Versicherten weder vor noch nach Abschluss des Versicherungsvertrages

1. die Vornahme genetischer Untersuchungen oder Analysen verlangen oder
2. die Mitteilung von Ergebnissen oder Daten aus bereits vorgenommenen genetischen Untersuchungen oder Analysen verlangen oder solche Ergebnisse oder Daten entgegennehmen oder verwenden.

Für die Lebensversicherung, die Berufsunfähigkeitsversicherung, die Erwerbsunfähigkeitsversicherung und die Pflegerentenversicherung gilt Satz 1 Nr. 2 nicht, wenn eine Leistung von mehr als 300 000 Euro oder mehr als 30 000 Euro Jahresrente vereinbart wird.

1000-Genome-Projekt

http://www.1000genomes.org



Persönliche-Genome-Projekt

http://www.personalgenomes.org

Personal Genome Project

Home | Project Overview | Participation Overview | PGP Community | DONATE

Volunteers from the general public working together with researchers to advance personal genomics.

We believe individuals from the general public have a vital role to play in making personal genomes useful. We are recruiting volunteers who are willing to share their genome sequence and many types of personal information with the research community and the general public, so that together we will be better able to advance our understanding of genetic and environmental contributions to human traits and to improve our ability to diagnose, treat, and prevent illness. Learn more about how to participate in the Personal Genome Project.

Project Overview. The PGP hopes to make personal genome sequencing more affordable, accessible, and useful for humankind. Learn more about our mission.

Want to participate? We aim to enroll 100,000 informed participants from the general public. Learn more about participation in the PGP and how you can get involved.

Meet our volunteers. Participants may volunteer to publicly share their DNA sequence and other personal information for research and education. Meet the "PGP-10".

Documentary Film about PGP. Two-time Emmy Award-winning documentary producer Marilyn Ness is making a film about the PGP. Watch webisode 1, 2, and 3.

CC0. We are committed to making research data from the PGP freely available to the public. Read about PersonalGenomes.org's use of the CC0 universal waiver.

Help Us Build. Are you a Rails developer interested in supporting the PGP through volunteer web application development? See our new project page at RailsBridge.org.

Project News

Keep up-to-date with PersonalGenomes.org news and events. Subscribe to our newsletter.

October 13, 2009: Our newsletter highlights recent advancements and preparations for the next phase of the PGP. Read.

August 18, 2009: Science NovaNOW will feature a segment on the Personal Genome Project. Founder, George Church, will also be participating in an "Ask the Expert" Q&A feature addressing questions from the public. More info.

August 8, 2009: Eligibility screening results will be issued to all prospective participants by August 15th. Complete your eligibility screening application for the PGP-100 now, if you haven't already. Applications will still be accepted after August 15th, but time is running out to be considered for the PGP-100.

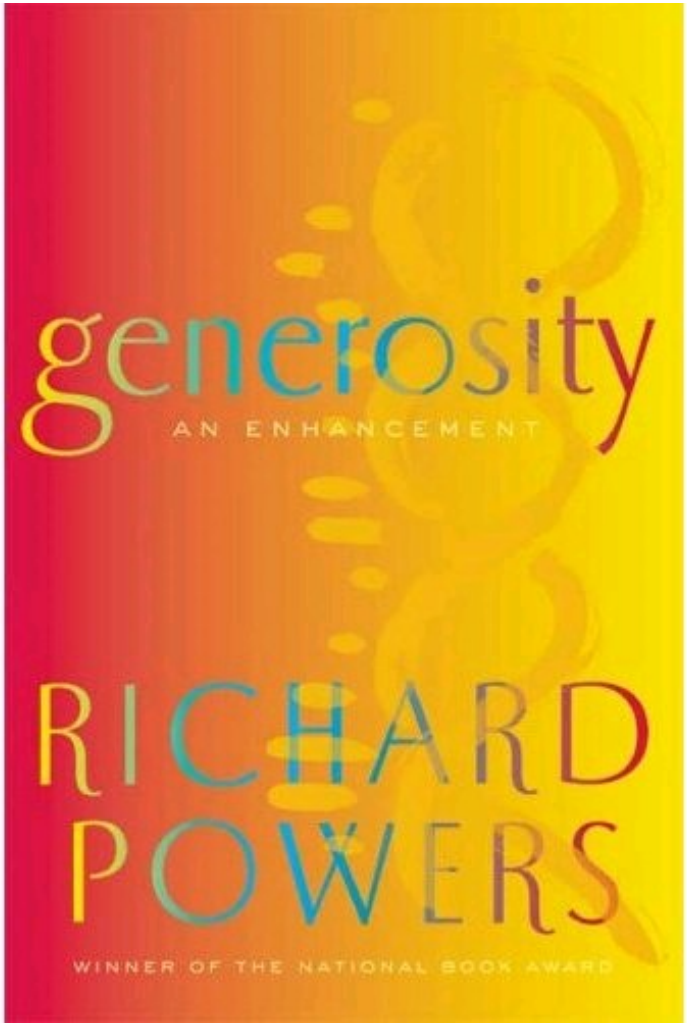
July 30, 2009: Two-time Emmy-award winning

Complete Genomics

http://www.completegenomicsinc.com

2009: 1.000 … 2010: 20.000 Genome

Richard Powers: "Das größere Glück", 432 Seiten, Fischer



SPIEGEL ONLINE 14.10.2009

Wissen kann unfrei machen

Für seinen neuen Roman ließ Richard Powers sein Erbgut entschlüsseln

…

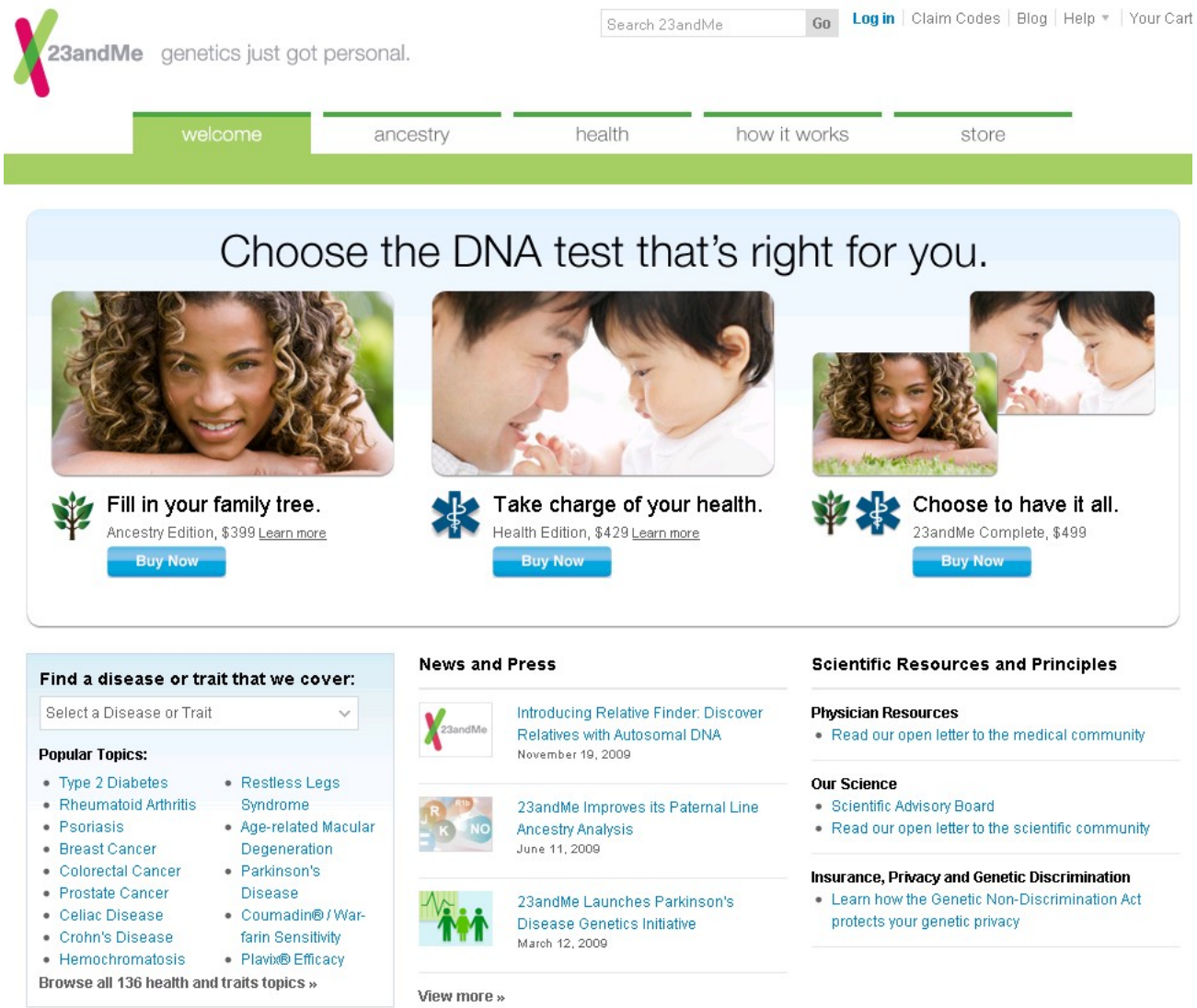
Powers: Nein, derzeit bringt uns die Gen-Diagnostik nicht weiter auf dem Weg zur Selbsterkenntnis. Einige Ergebnisse mögen ähnlich nützlich sein wie ein Cholesterin-Test. Sie könnten als Impuls dienen, ein wenig gesünder zu leben. Aber die Vielzahl von Informationen, die wir gewinnen, verwirren eher als dass sie uns nützen. Die meisten Menschen würden dadurch ängstlicher werden.

…

http://www.spiegel.de/kultur/literatur/0,1518,654627,00.html

23andMe

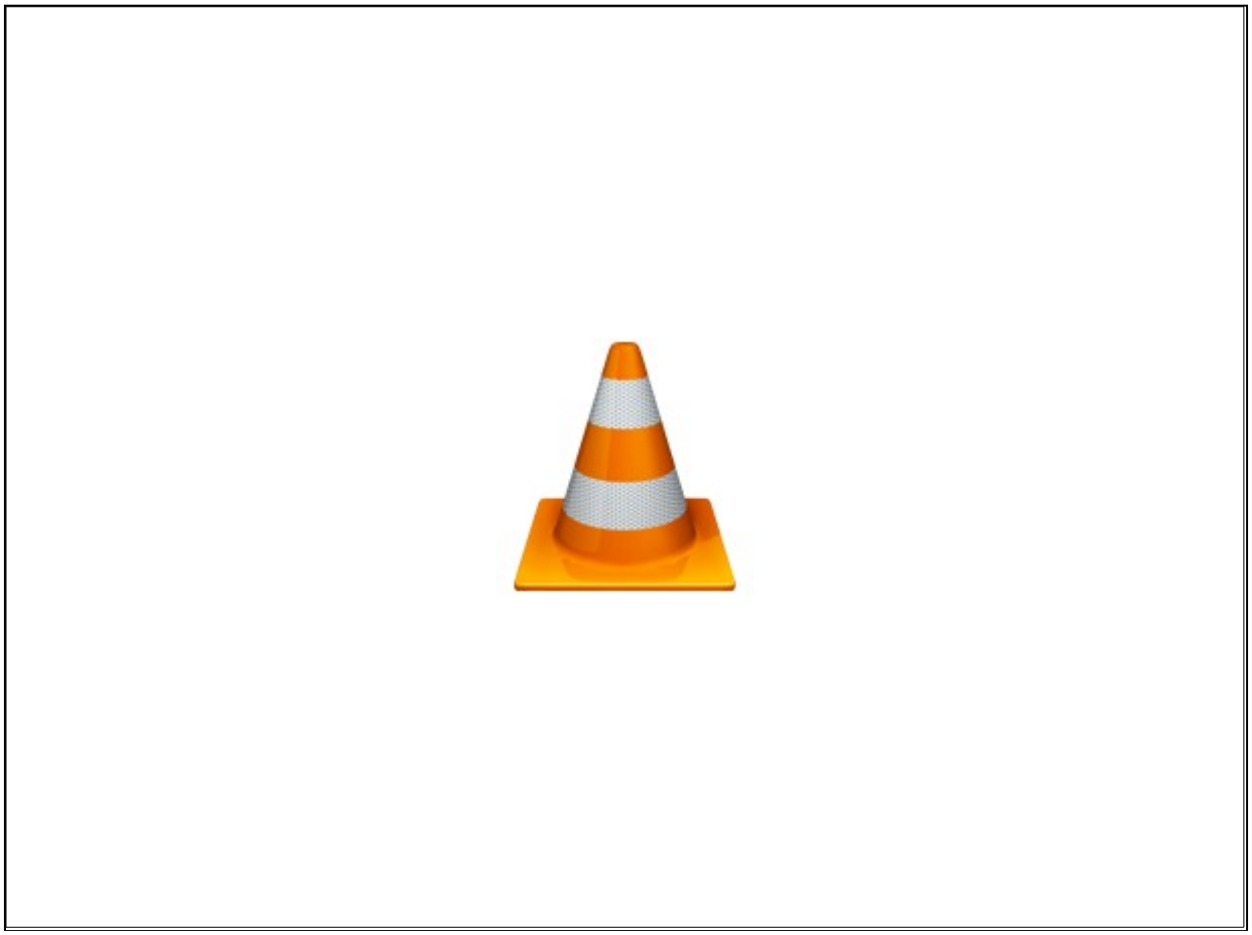
Genetics just got personal



www.23andme.com

23andMe

Genetics just got personal



www.23andme.com

Ausblick

Nature 21. Oktober 2004

End of the beginning

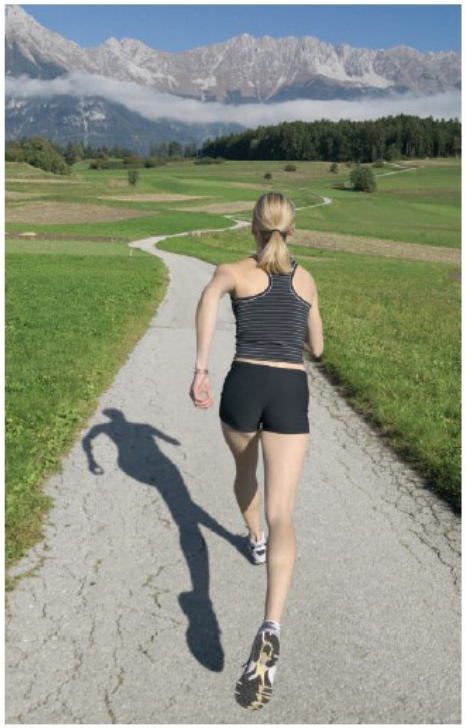
Lincoln D. Stein

Just over three years ago, it was announced that a first draft of the human genome sequence had been completed. Gaps and errors remained, but the job of fixing those problems is now largely done.

This issue of *Nature* features an article[1] entitled "Finishing the euchromatic sequence of the human genome". It has been authored by members of the International Human Genome Sequencing Consortium (IHGSC), and appears on page 931. The article marks the latest, but by no means the last, milestone in this historic project. But readers can be forgiven for being a bit confused by the announcement. Wasn't the human genome 'finished' several years ago?

The answer is 'yes' — and 'no'. Early in 2001, the duelling IHGSC (public) and Celera Corporation (private) groups published papers in *Nature*[2] and *Science*[3] describing the completion of so-called 'draft' sequences. These sequences have revolutionized molecular biology by largely eliminating the need to clone and sequence genes involved in human health and disease. Instead of going to the bench, biologists now go to the web to look up gene sequences in public online databases.

But despite their immediate usefulness, the draft sequences were far from perfect. Both drafts were missing some 10% of the so-called 'euchromatin' — the gene-rich portion of the genome — and some 30% of the genome as a whole (which includes the gene-



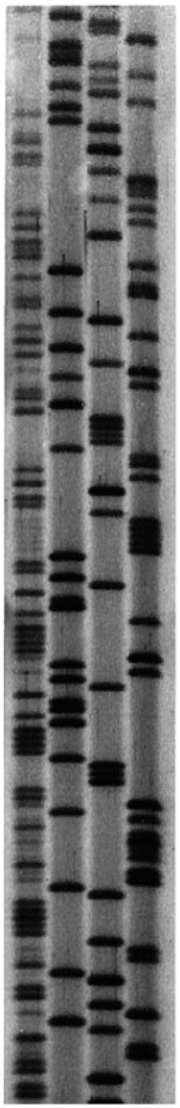
GETTYIMAGES

of DNA. The sequence announced today has just 341 gaps remaining, and consists

mutation that rendered the gene non-functional. The resulting pseudogene then slowly degrades and disappears.

In a second analysis, the authors use the finished sequence to map out segmental duplications — large regions of the genome that have duplicated in recent evolution. They find that 5% of the genome is involved in segmental duplications, and that the distribution of these regions varies widely across the chromosomes. Knowing the nature and extent of such duplications is important for understanding the evolution of the human genome, and for studying the many medically relevant disorders that are involved in segmental duplications, such as DiGeorge syndrome and Charcot–Marie–Tooth syndrome.

Another paper in this issue, by She *et al.*[4] (page 927), directly compares the outcomes of this second analysis with results obtained on an unfinished version of the human genome (an improved version of the Celera draft). She *et al.* find that the draft version artefactually 'simplifies' the genome by eliminating many duplicated regions. Their results bear on one of the highly publicized differences between the public and private genome projects. The public project used an older strategy in which the genome was first



genome.fli-leibniz.de
Teaching