Genomanalyse

Vorlesung Masterstudiengang 2011
Jena 9.11.2011

Klaus Huse
Leibniz-Institut für Altersforschung
Fritz-Lipmann-Institut Jena

Genom

Begriffsbestimmung

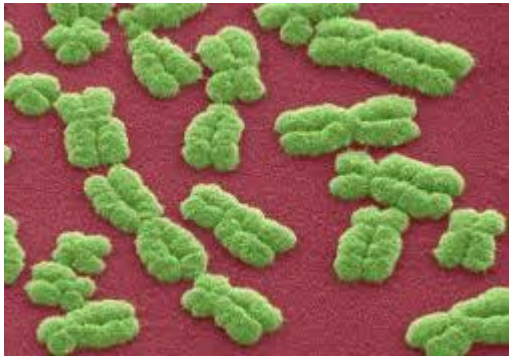
Gen (Johannsen, 1909) - Erbfaktor, der eine einzelne Einheit hereditären Materials darstellt

Genom (Winkler, 1920) - gesamtes genetisches Material einer Zelle oder eines Individuums
(nukleares, mitochondriales,… Genom)

Genomanalyse

Was ist das und wozu dient sie?

Strukturanalyse
Teilgebiet der chemischen Analyse zur Aufklärung des räumlichen Baus von Molekülen (vor allem kompliziert gebauter organischer Verbindungen wie den Chromosomen)



Funktionsanalyse
Zuordnung von Eigenschaften (Verhalten unter bestimmten Bedingungen) zu Molekülen und Molekülteilen (Genotyp bestimmt Phänotyp)

Genomanalyse

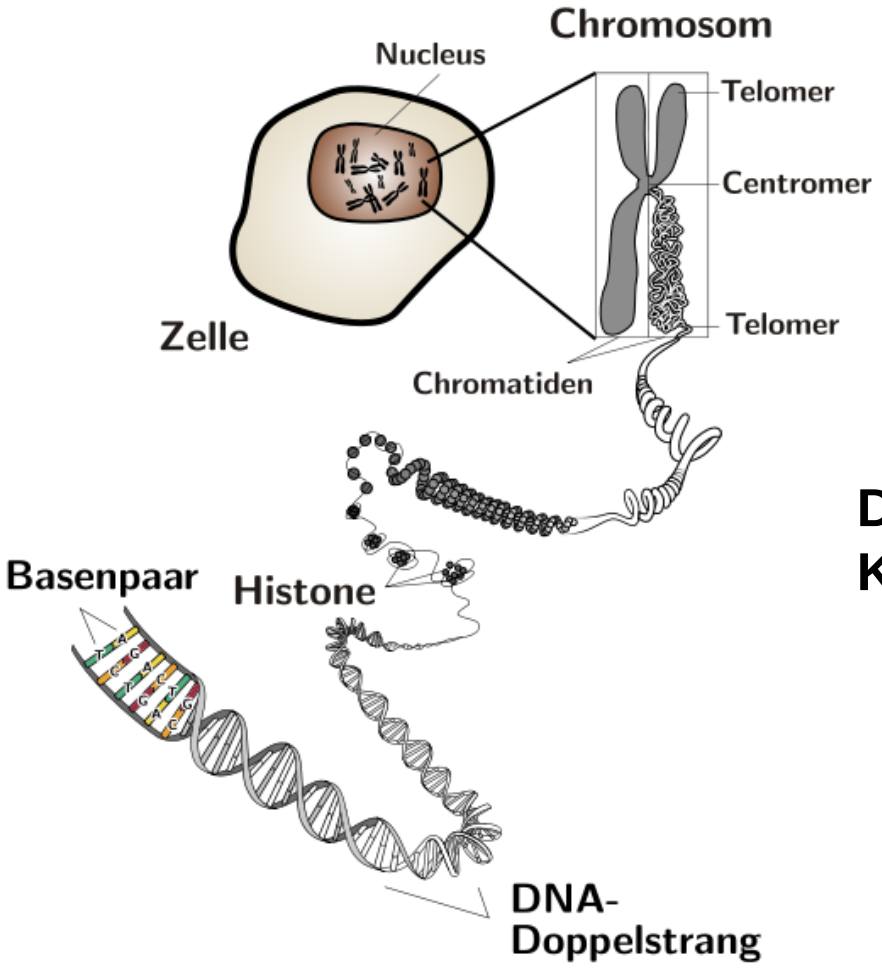
Vor 1700 Jahren

Talmud:
...wenn 2 Jungen bei der Beschneidung an Blutungen verstorben sind, brauchen alle weiteren Söhne der Mutter, ihrer Töchter und ihrer Schwestern nicht mehr beschnitten zu werden.

= X-chromosomal rezessiver Erbgang einer Hämophilie

Genomanalyse

Targets



Zellorganellen

DNA-RNA-Protein-Komplexe

Isolierte Komponenten

Genomanalyse

Metaphase-Chromosomen: Karyotypisierung

Anzahl und Art der Chromosomen:

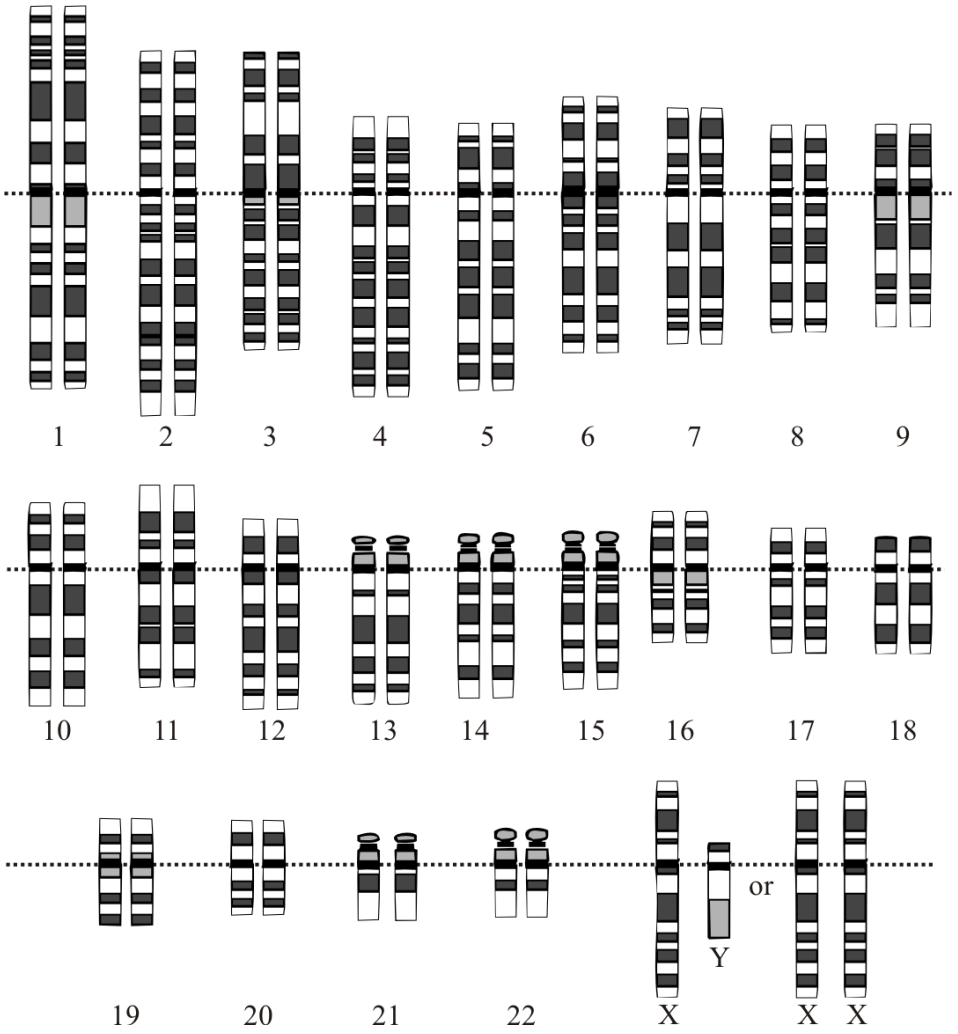
Größe, Lage des Centromers, Bandenmuster nach Färbung

Aussagen zu

- Spezies
- Geschlecht
- Anomalien (Krankheitsassoziationen)

Genomanalyse

Metaphase-Chromosomen: Karyotypisierung

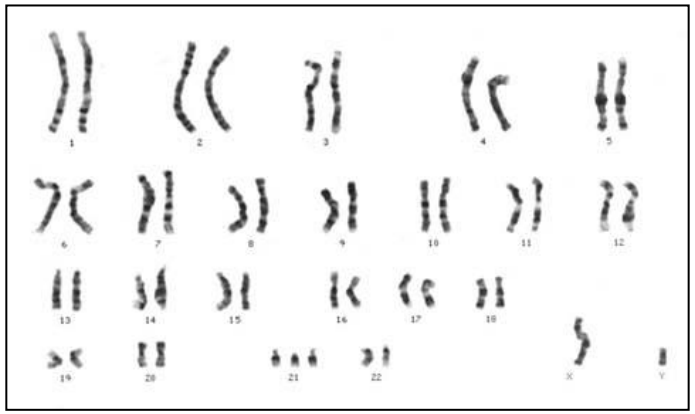


46, XX

46, XY

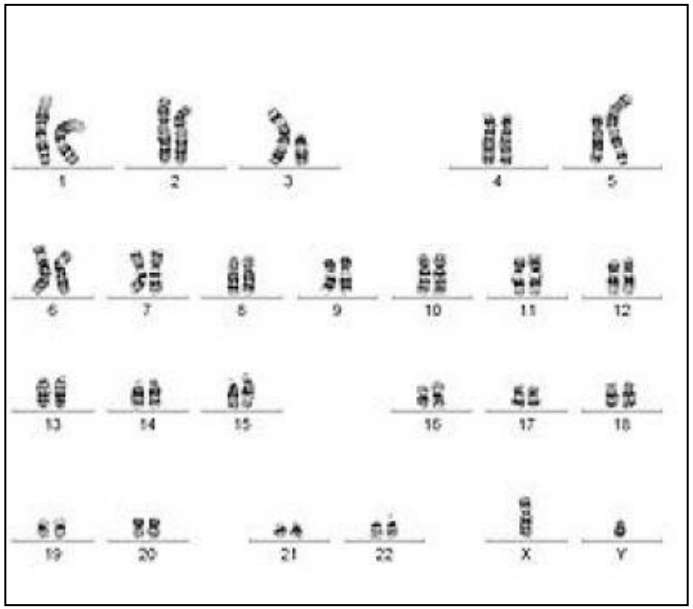
Genomanalyse

Metaphase-Chromosomen: Karyotypisierung



47, XY + 21

Down Syndrom



46, XY, t(3;5)(q25;q34)

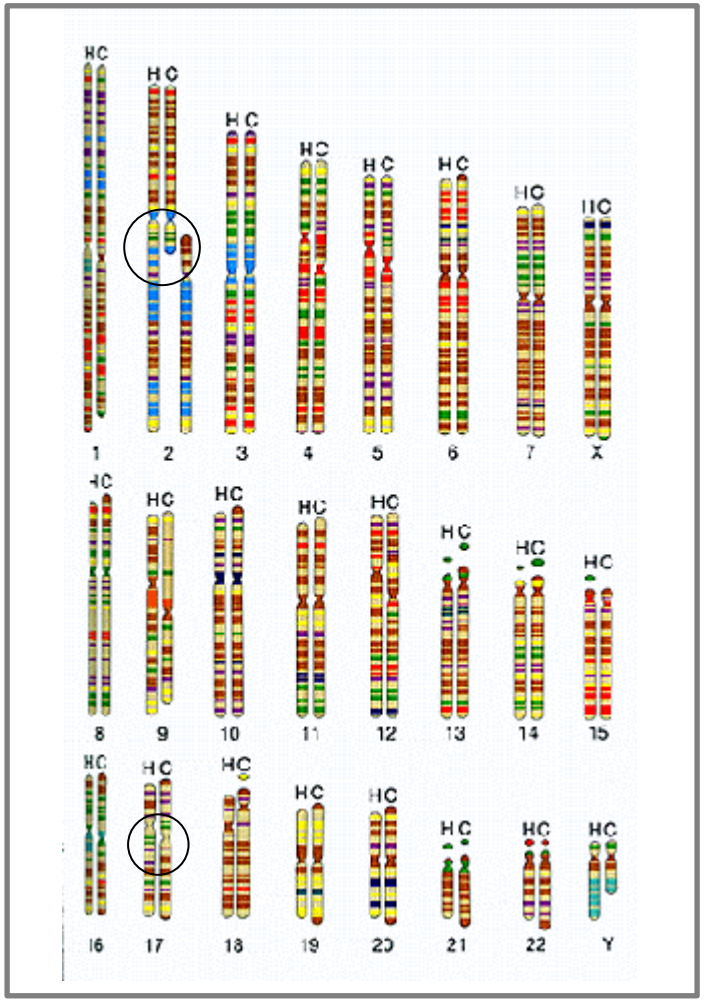
akute myeloische Leukämie

Genomanalyse

Chromosomensatz

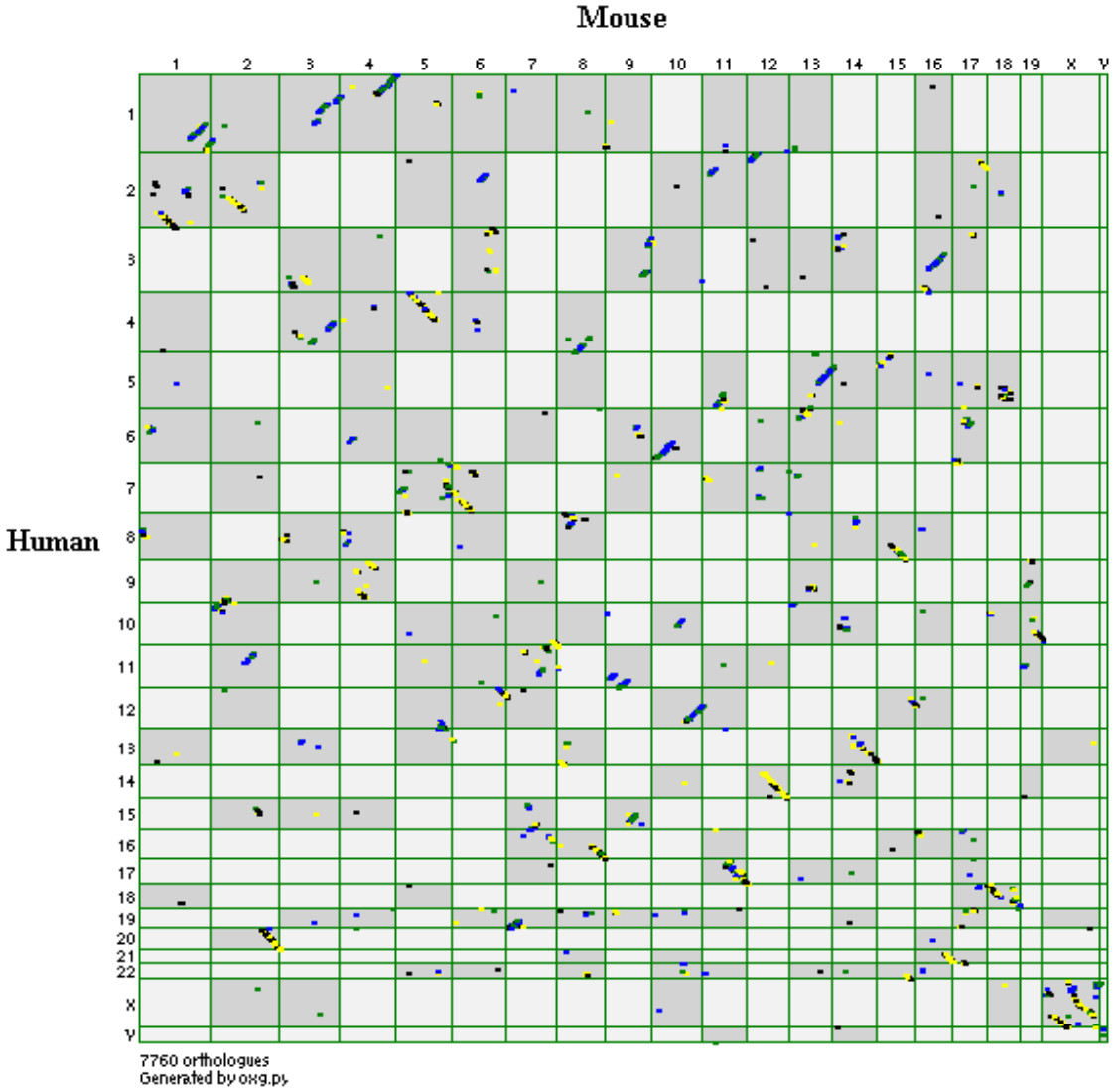
Homo sapiens: 46
Pan troglodytes: 48

9 intrachromosomale Inversionen
1 Chromosomenfusion



Genomanalyse

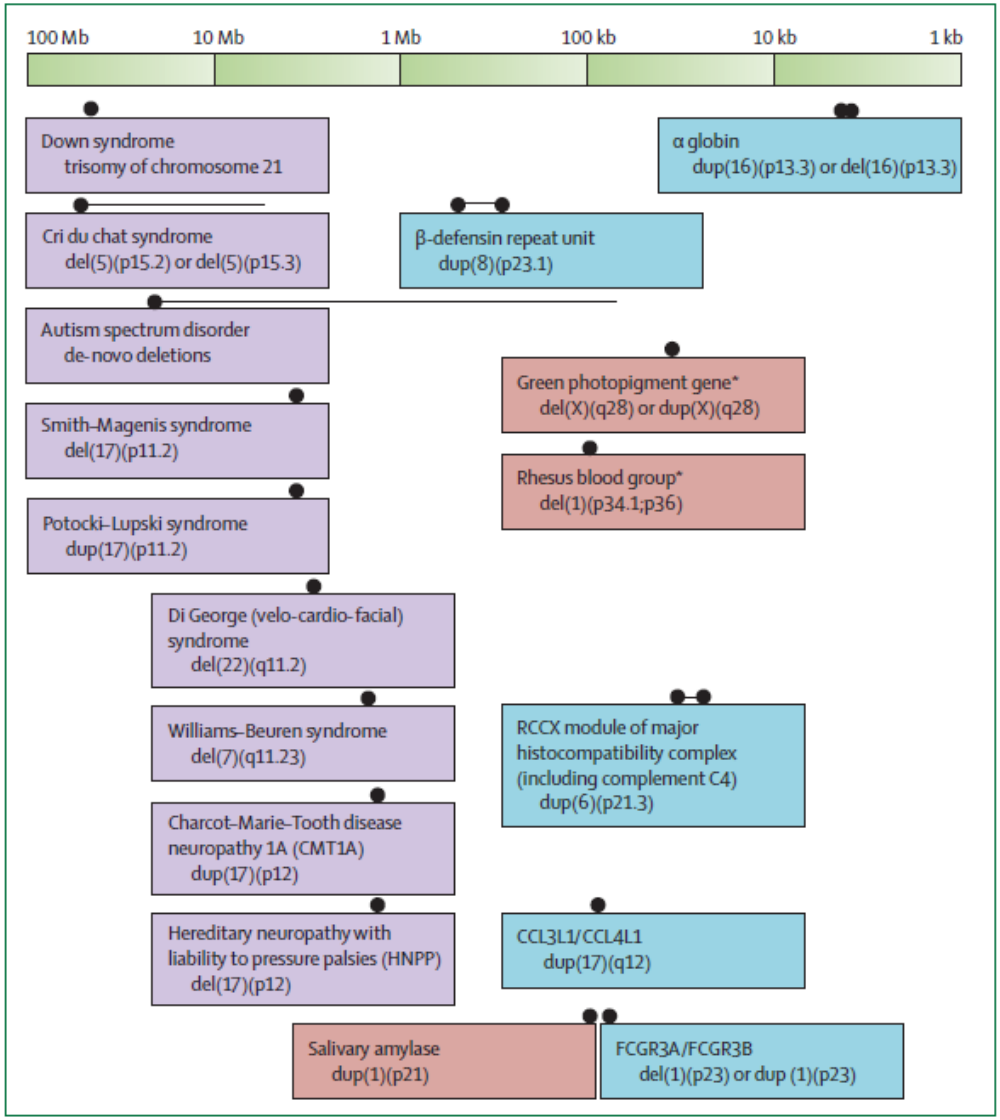
Comparative Genomanalyse: Syntenie



7760 orthologues

Genomanalyse

Strukturvarianten

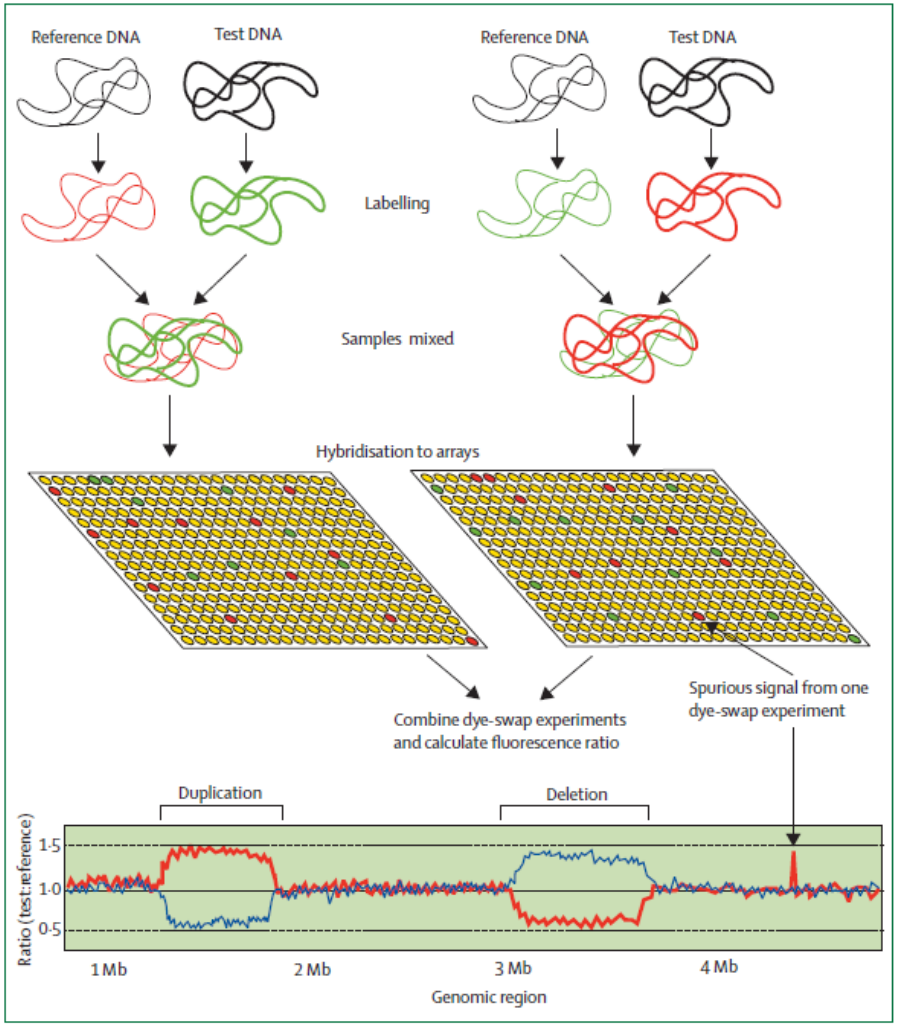


Genomanalyse

Bi- und multiallelische Strukturvarianten

Genomanalyse

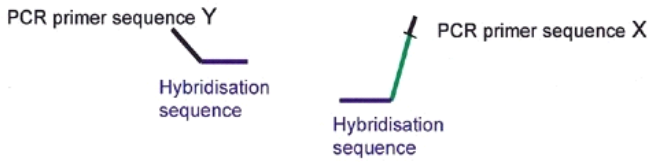
Comparative Genom-Hybridisierung



Genomanalyse

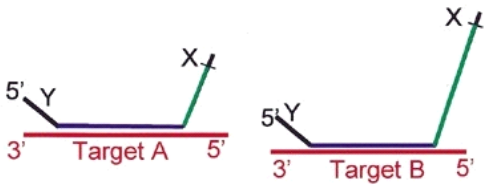
Multiplex Ligation-dependent Probe Amplification (MLPA)

Probe
(target-specific)



Annealing

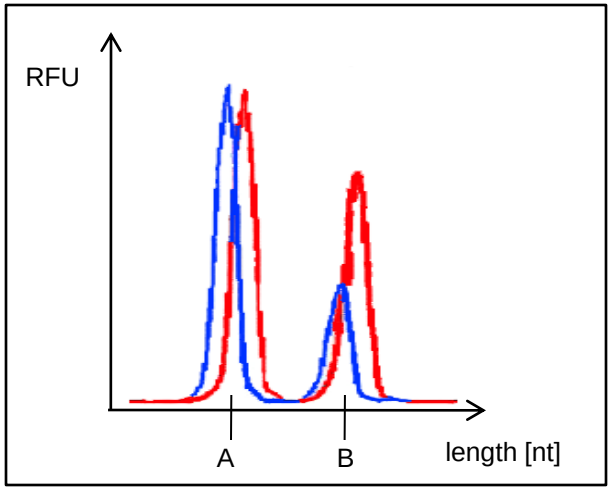
Ligation



Amplifikation

adapted from Schouten et al., Nucleic Acids Res. 2002 Jun 15;30(12):e57.

Separation of amplicons & determination of peak areas

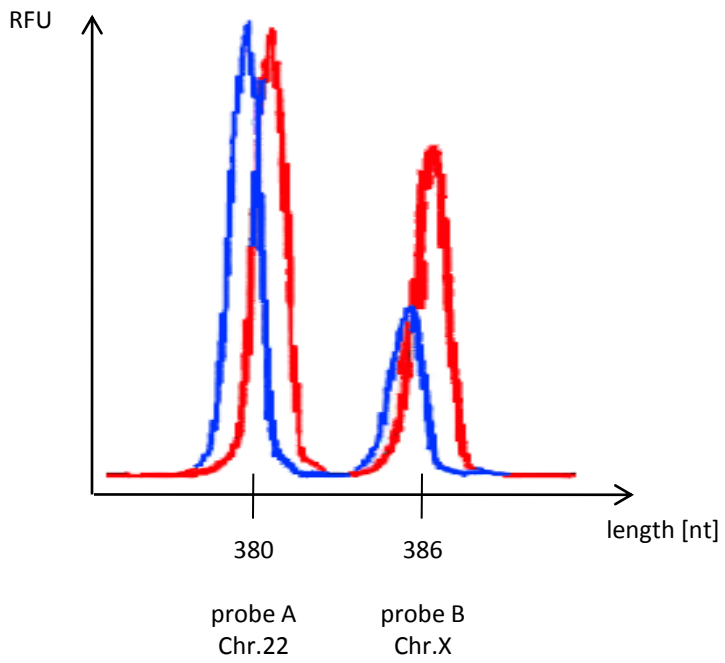


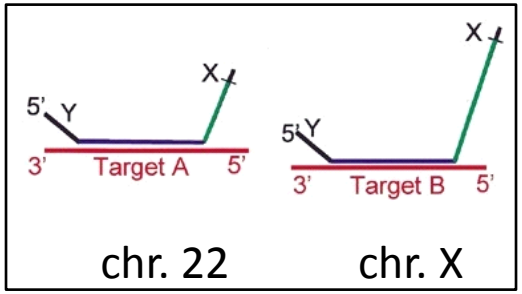
Genomanalyse

Multiplex Ligation-dependent Probe Amplification (MLPA)

MLPA: DNA-Sonden spezifisch für Autosom und X-Chromosom

Male/Female



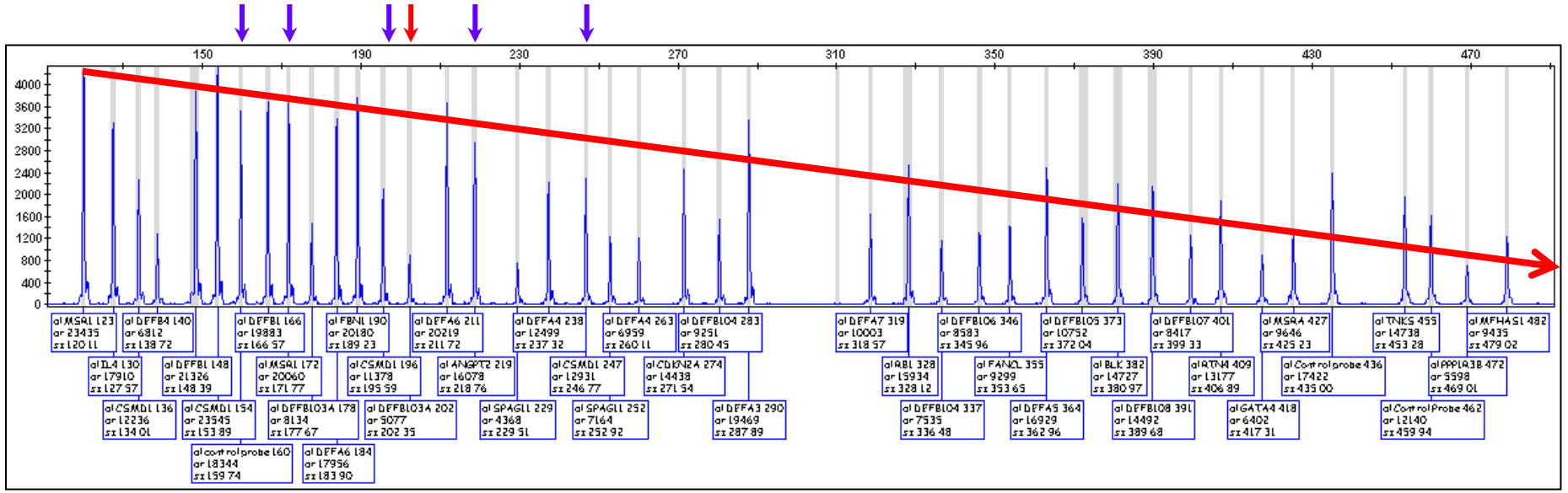


adapted from Schouten et al., Nucleic Acids Res. 2002 Jun 15;30(12):e57.

Genomanalyse

Multiplex Ligation-dependent Probe Amplification (MLPA)

- raw data: length of PCR product and respective peak area

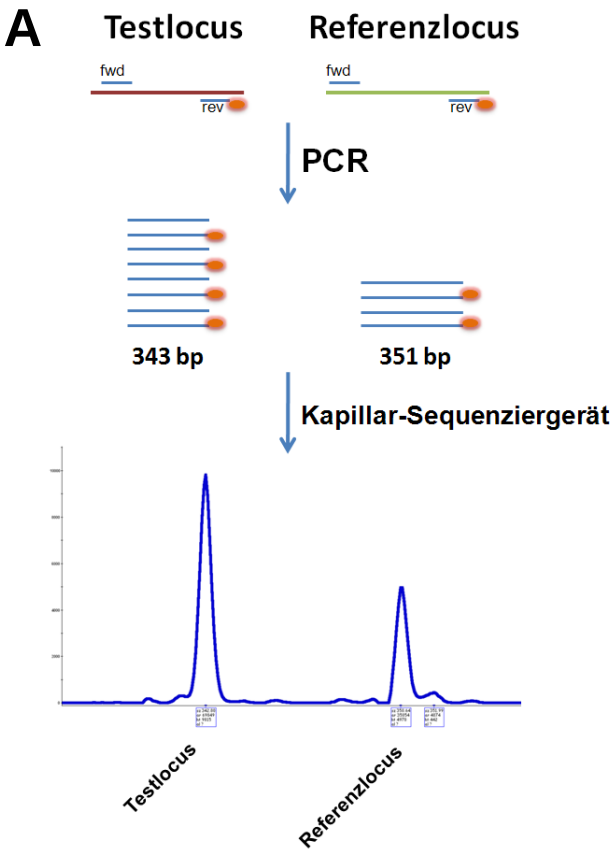


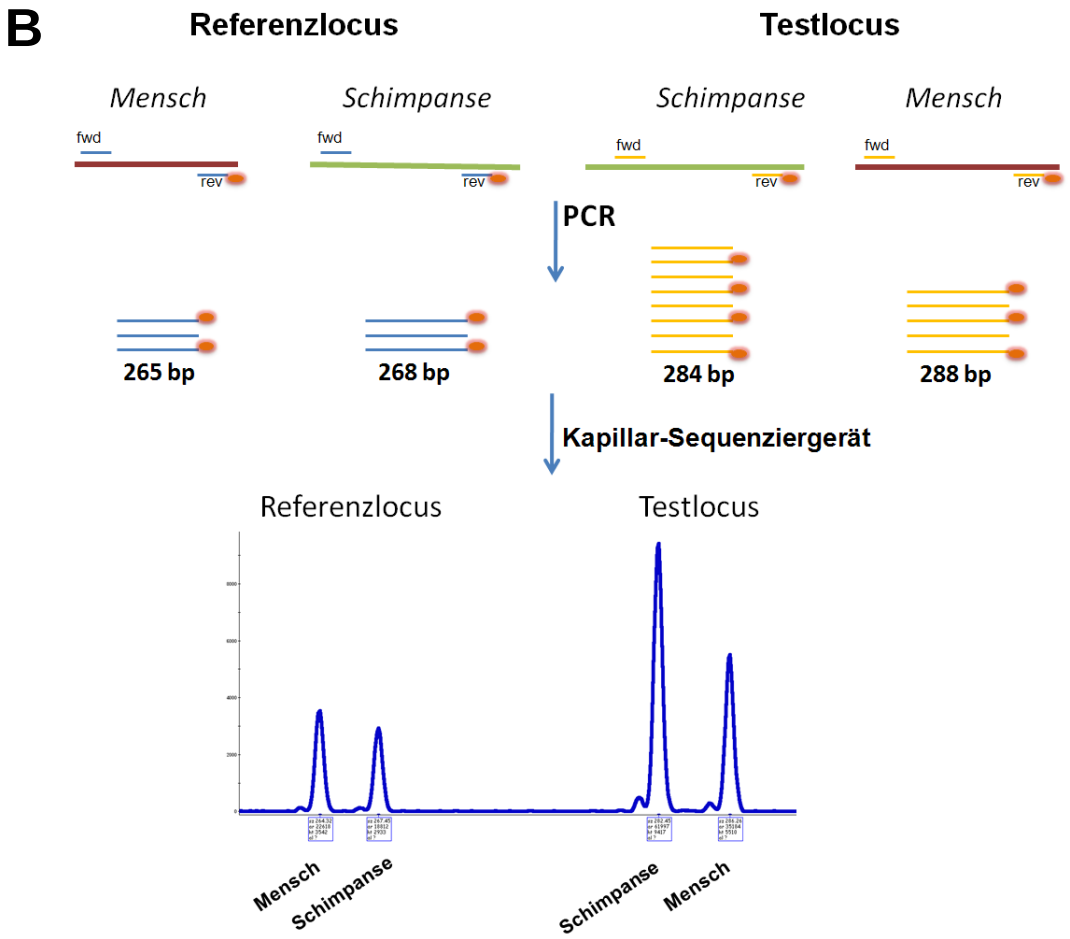
- to correct for fading for longer PCR amplicons: five nearest single-copy neighbours (5nn) method used

aus Promotion Marco Groth

Genomanalyse

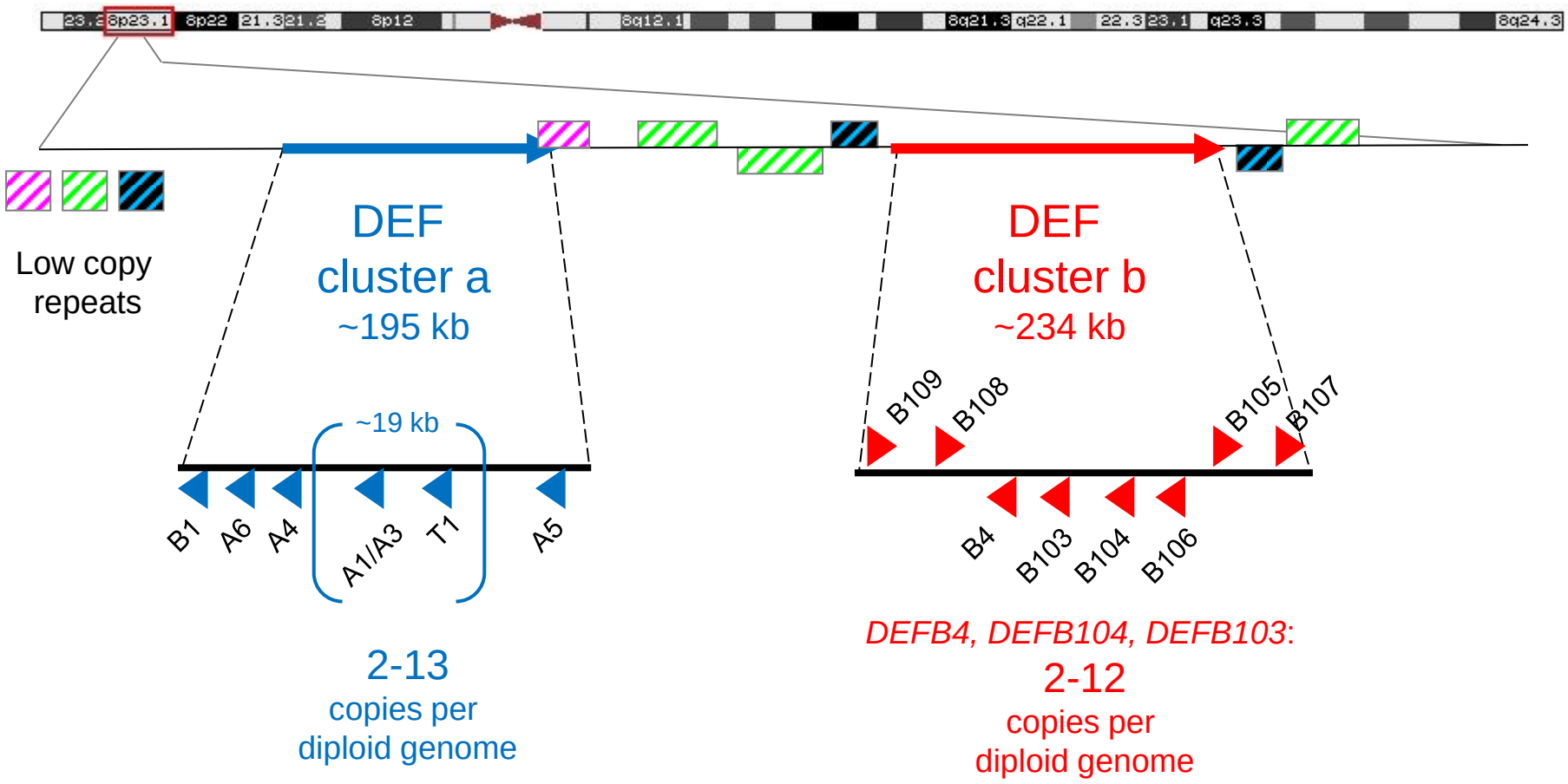
Paralog-Ratio-Test/Ortholog-Ratio-Test





Genomanalyse

Defensin cluster auf 8p23

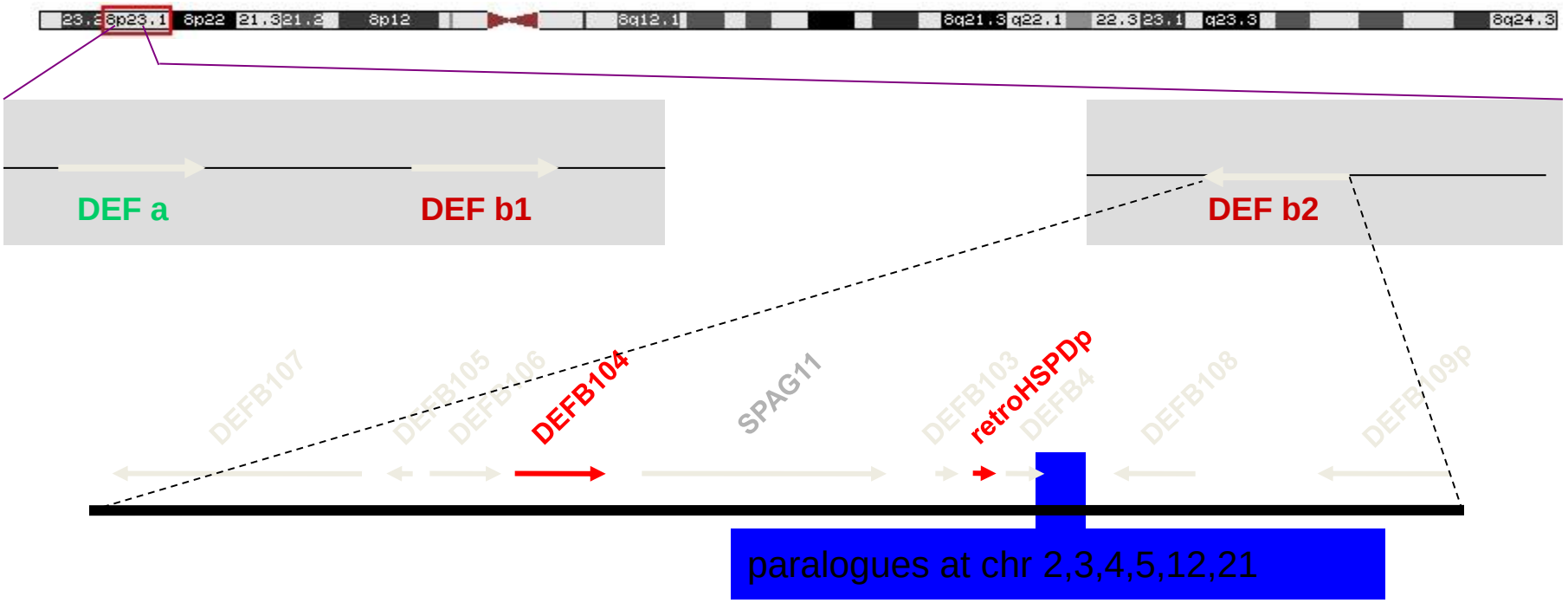


Taudien et al. BMC Genomics 5:92 (2004)

Genomanalyse

Paralog Ratio Test

Pseudogene sind im Genom häufig und weisen oft detektierbare.Sequenzunterschiede auf

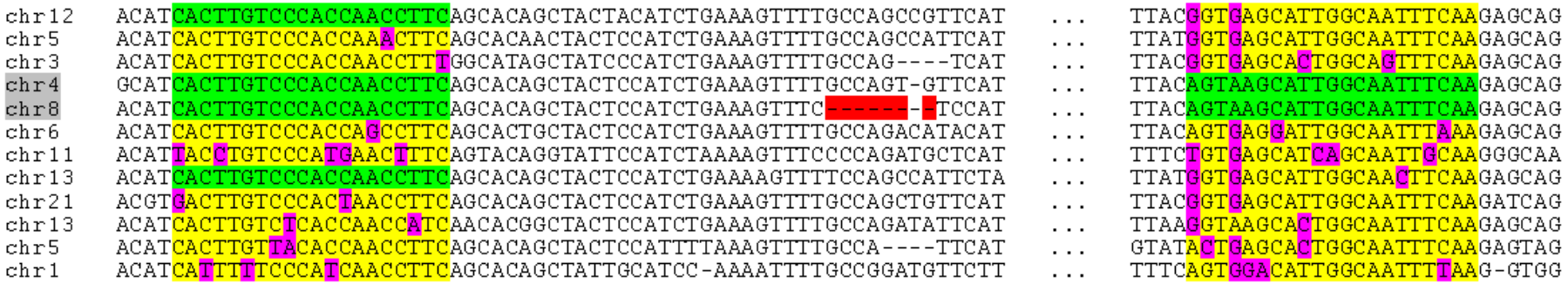


Ein oder mehrere Paraloge können als Referenz verwendet werden.

Genomanalyse

Paralog Ratio Test

Sorgfältiges Primerdesign erforderlich

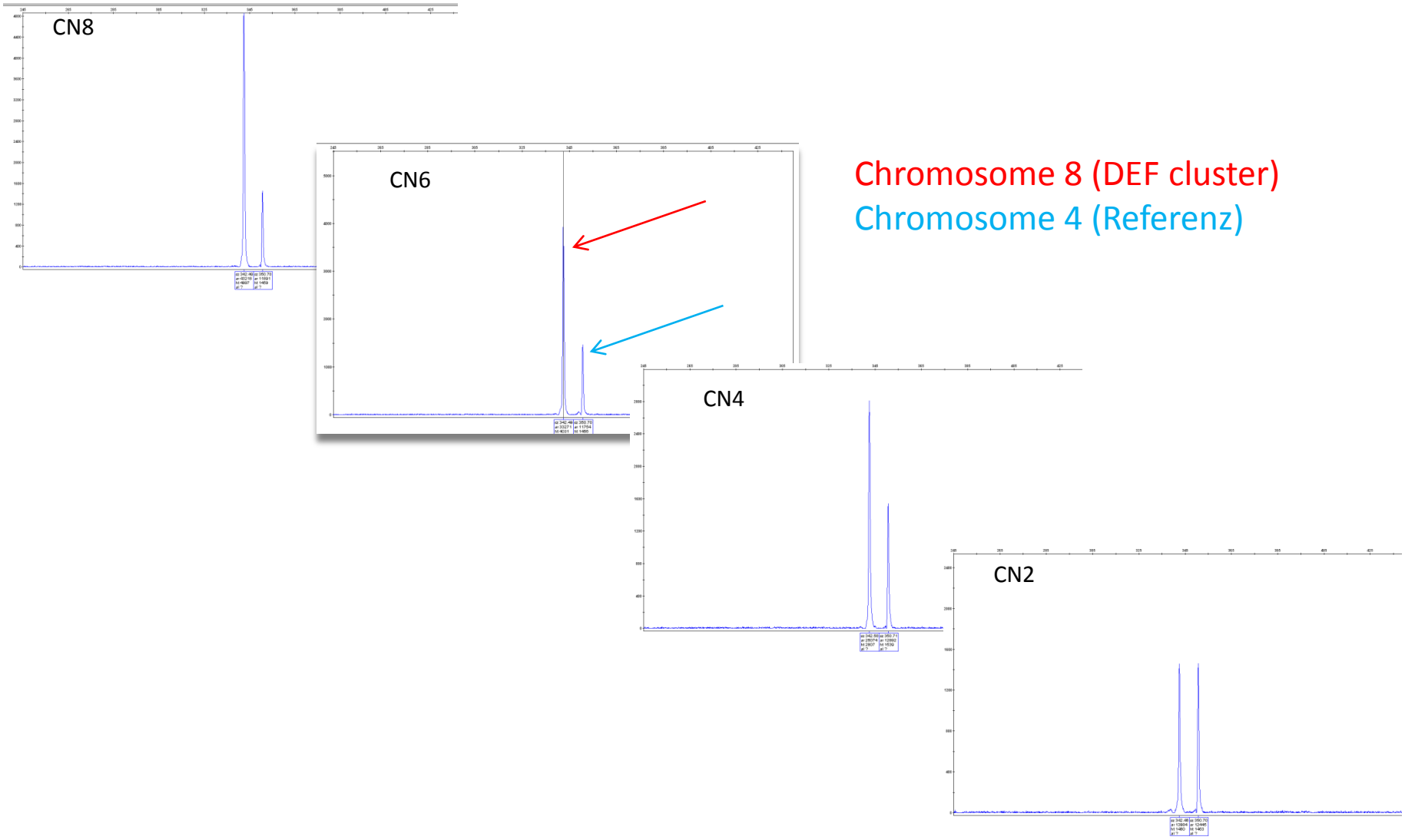


fw primer

rev primer

Genomanalyse

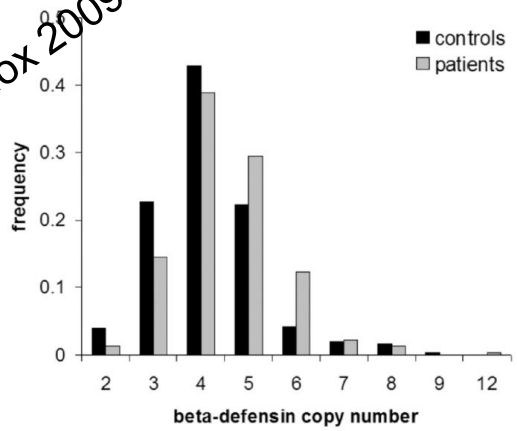
Paralog-Ratio-Test für Defensincluster auf Chromosom 8



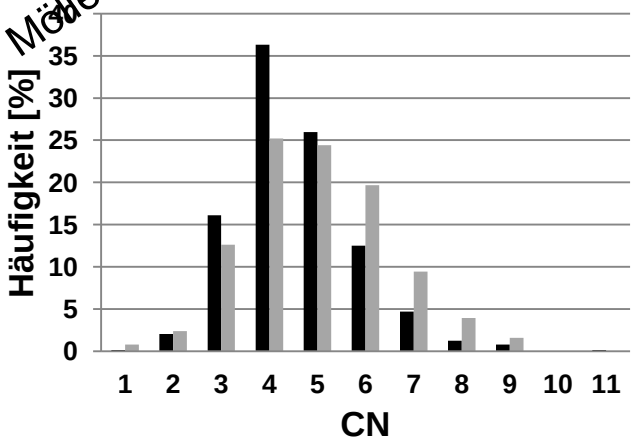
Genomanalyse

Assoziation von Kopienzahl und Psoriasis

Hollox 2009



Möller 2009



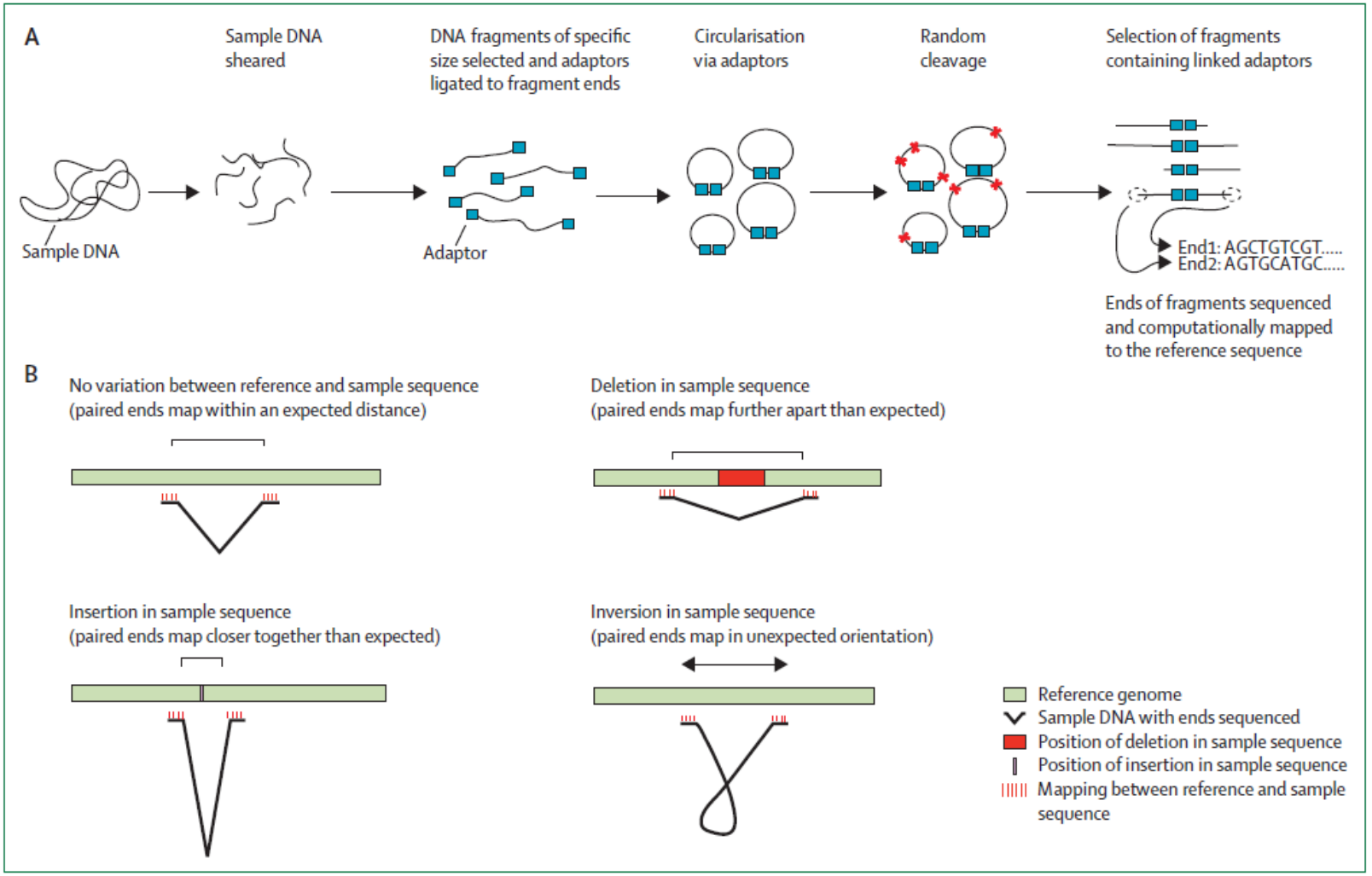
Patienten mit Psoriasis weisen höhere Kopienzahlen auf

Morbus Crohn
Prostatakarzinom
Pankreaskarzinom
Arthritis

Defensine sind an der Regulation der angeborenen Immunantwort beteiligt!

Genomanalyse

NextGen Sequencing und Strukturvarianten



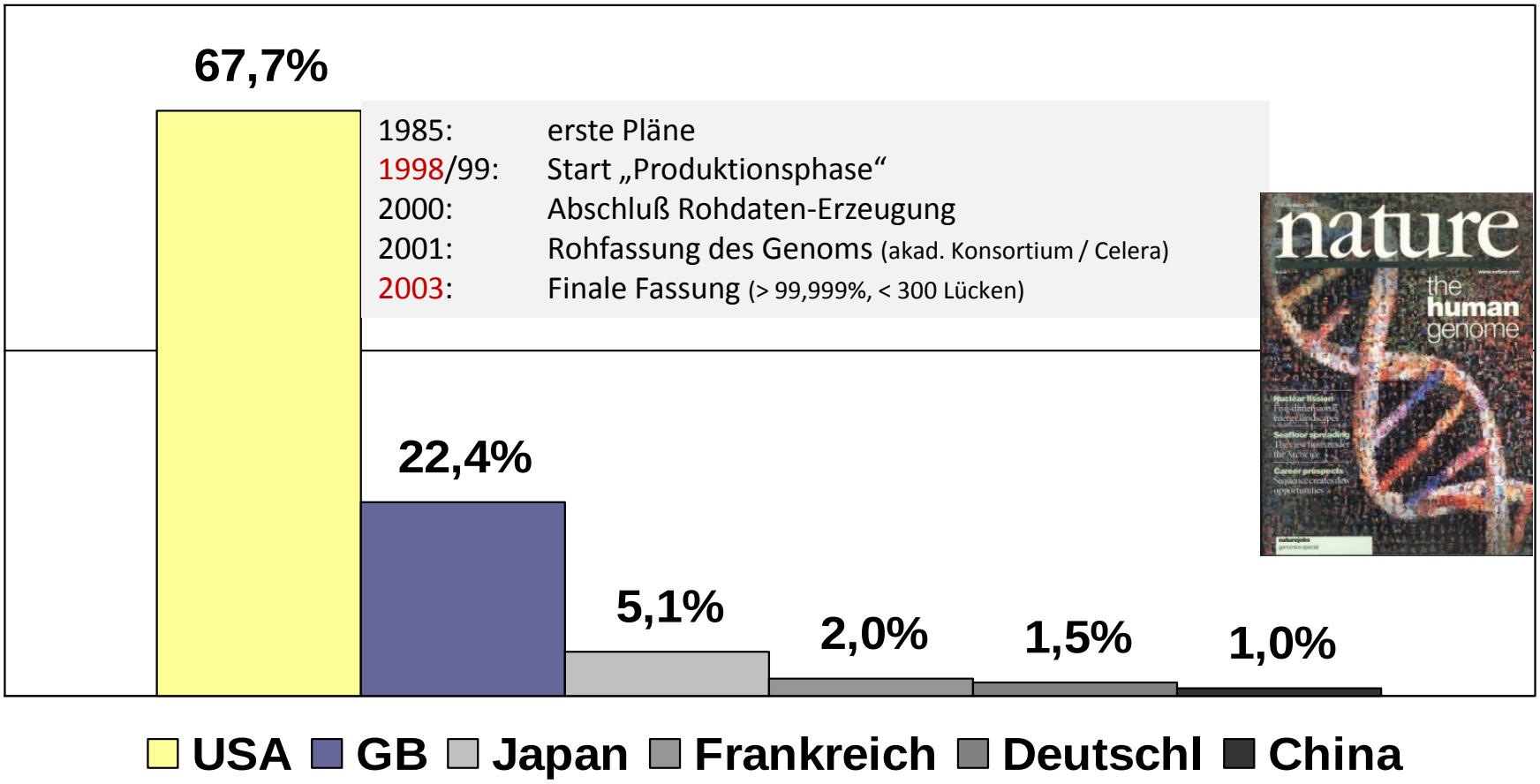
Genomanalyse

Eine „ultimative“ Genomstruktur: Die Nukleotidsequenz

AAGAGCCAGTCACACGAGGACTGGCTGTGTCTACCTTGCCCCGCGCCGTGGGCATCCCCAGCACCAAGCATGGTGCCT
GGCATGCAGTGAACCTTCAGTGTCACCTGACCGAGTAAACACAGGACTGAAGGAAGGGAAGGCGAGATGGATGCAAAR
GGCCCAGTGTGCCTCAGTCTTCTTTGGGAGGCAGCGGGGACCAGGGACATTTTTGAGGTGGTTTCCTCTCCAGGGATT
AACCGAGTTTATCCCTCCTCAGTTCTCCCAGGCTATGGGATGGATGCCAAGGGACCTTCCAGACAAACTATTCTGAAG
TCTGGGCCTGTGAGAAGCATCTTGAGAGATCTTTATGCTGCTGTGAGGATTCTAGGACCTGGAGGTCTGTCAGGGCAA
TCAGGAATGCTCCCTGGAGGAAGTGGCAGTGAGGCAGAGCTTTGTGGAGGGGCAGCAGGGGAGAGAGTGCTGTGGCCA
GAGGAAGGGAAAGGGCCATGGCAGGGAAGCAGGCCAGACACTGTTAGGCCTATTAGGAGAGTGTGAAGGGTAATTGGA
TCTGAGTGGGAGGAGCGGCTGTGGTAACTCCTTGTCTCAACGCCTCTGACTAAAGTGGCCAGGCACACCACCCAGGCA
CCTGCCGCCCTTGCCCCAGGGGCTGCTGCAACACCCAACTCGGCTTTAGCCAGAGGGCACTGGGCCAGGCAGATATGG
GGTCCTTCTCCCTGAAAGACAGGGAAGGAAGGTTCCTGCCCCAACCCCAGGCTCTGCCCAGGCTCGTGGAGAAGCCTG
TCCCCGAGGCAGTGGGCGGAACCAGGATGCTCTCAGTCACAGGGCAGCCAAGCGGCTGGGGCGGCAGCTCTTCATCAT
CTTCAGTTGCATTTGACAGCTGTGGCTGACCAGGAGGGGGCAGCTAGGCAGAAGTGTCATCCCGGTCACACCAGTGTC
ACCTGAGGCTCATCAGGGCAGTGACTTGCCCGAGGTCTCCCGGAGCAGCTCAGTAAGGATTGAGCATTTGTTTACTCA
TCTCCATCGCTGACAGCTGAAACTGGATTCATCCTGCCCTGCCCAGGGGGCTGGGAATCAGGGATCGGGGCCAGCTCA
GCTGAGATGTGAGGGTCCCAAATAGTATTTCCACCCCTGACCCCAAAATAGGCCCCAATCCCCTGTCCCAAAGACTTG
ACGTGCCCATTGATGTGTCTGTCGTTCCTGACAGGGTGTCCTCTTGGTTCCTCAGGTCAGAAACCCAAGGGTCATCTC
TGTCCCTGGGCTCTTCCTCACTGCCCACATCCAGTCCATCAGCGAAGTGTGTTAGTTGGCTACCTGCATGCGTCCTGA
ATCCGTCTCCTTCTCCTCCTCCATCTCTCTGCTGGACCACAGCCGTGAACTCATGCCCCAGCTTCCAGGCCTTCCTGG
TCCAGTCCATCTTCCTCTCAAAAGCCCATTGATCTCCAAAACACAAGTCTGAGCAAGGCTCTCCCCTCTTGAAATCCT
TCAATGCCTCCCCGTTACTCTCAGGACGAAGCCCACATTTCTGCCTACTTCTCCAGGCTCAACTTTTTCCTCCAGCTC
CATCACCCACTATATCATTTTGTCCCAGCCACATGGACCTTCTTTCAGTTCCTTGAAAATGCCGGGCCTTTACTGATG
CCCAAGCCTTCGTACAAGCTGTTCTTCCTATCTGGAACATTCTTTCCTGTTCCCCCGTCTCACTTCCTTTCACTCCTA
CTYTTATTTTAGAGTGCAGCTTGAACCCTTACTCCTTCCTTCCTCTATGTCCCCAGAGTTCACTGACTTCCCTCATCT

Genomanalyse

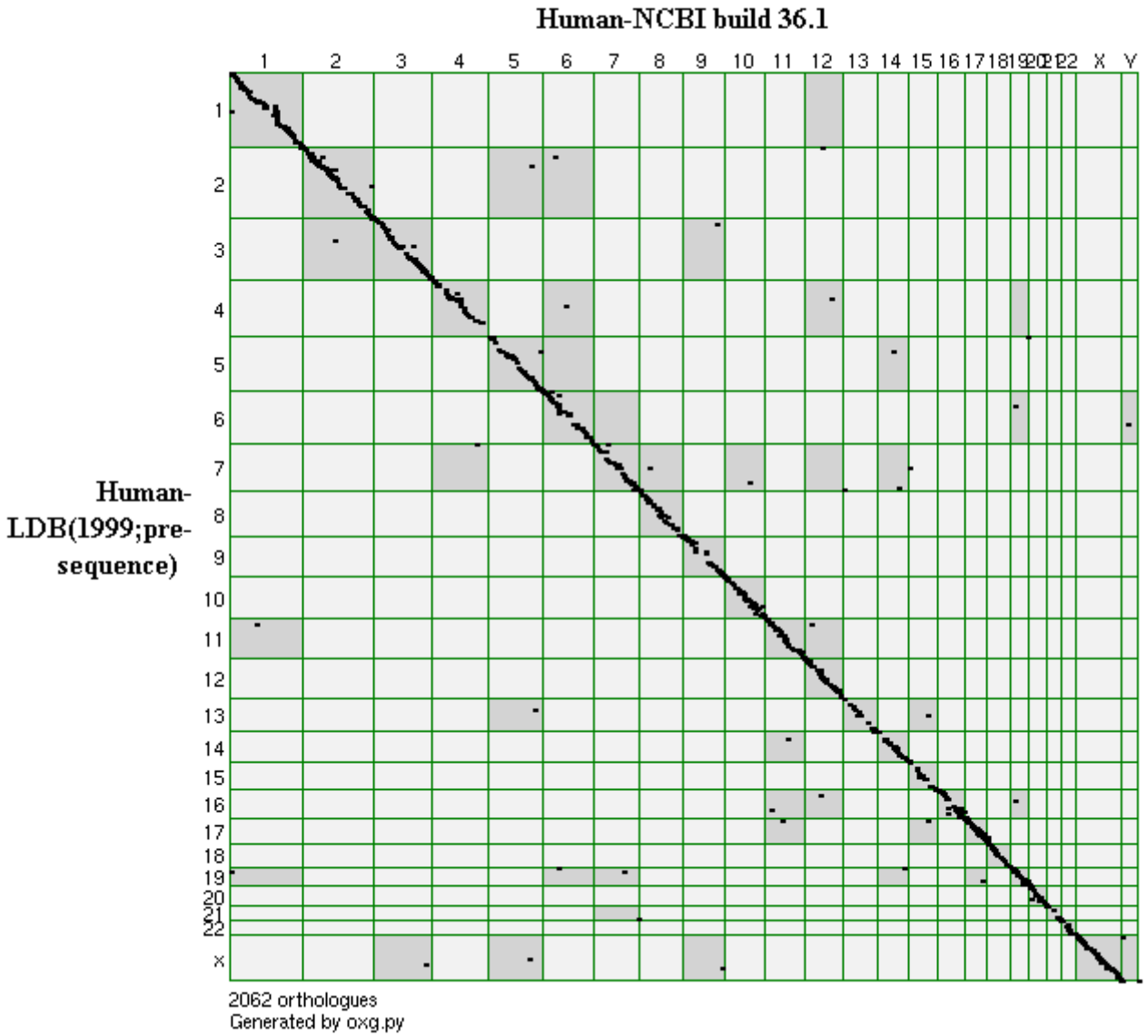
Eine „ultimative“ Genomstruktur: Die Nukleotidsequenz



1985: erste Pläne
1998/99: Start „Produktionsphase“
2000: Abschluß Rohdaten-Erzeugung
2001: Rohfassung des Genoms (akad. Konsortium / Celera)
2003: Finale Fassung (> 99,999%, < 300 Lücken)

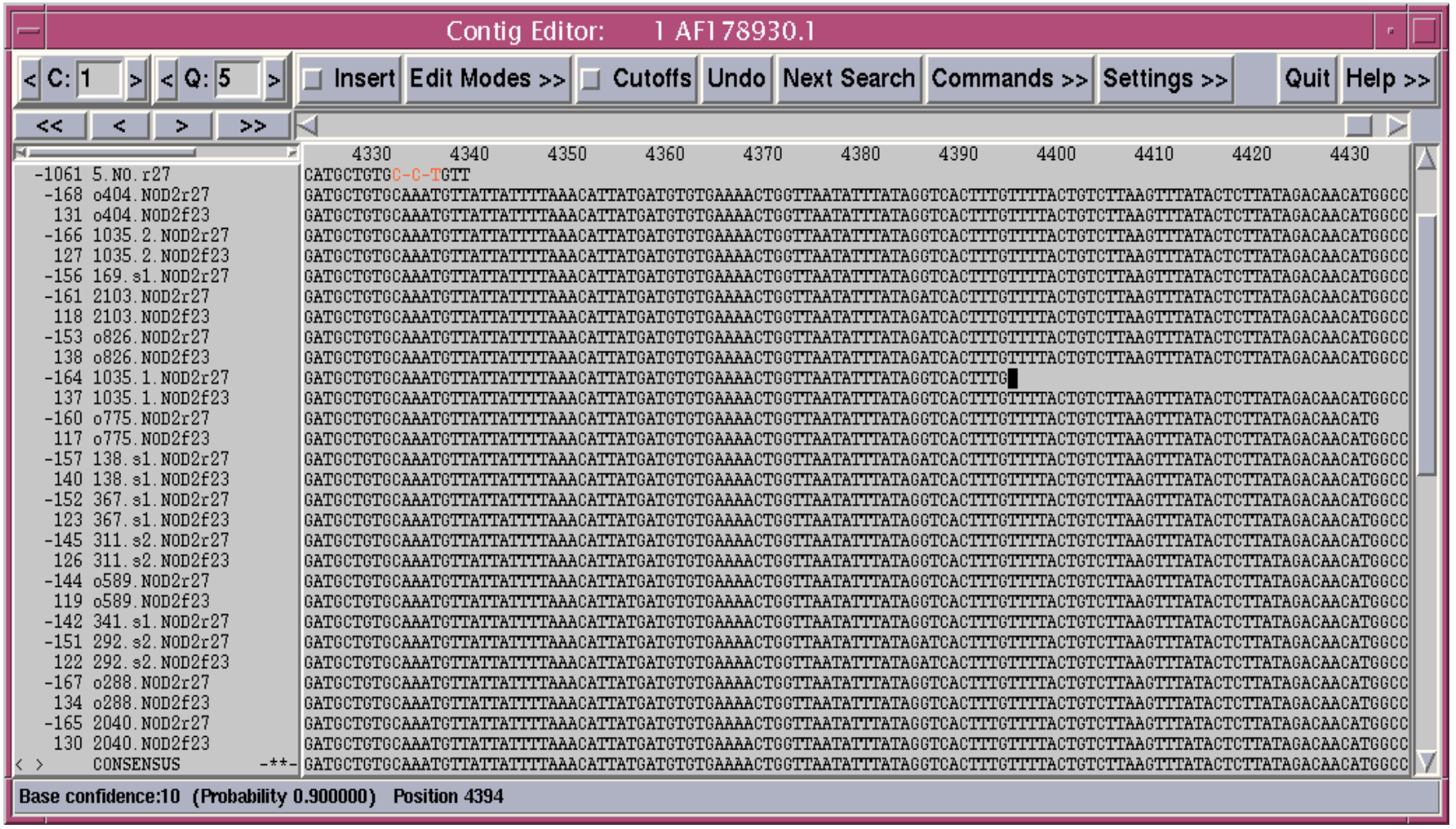
Genomanalyse

Klassische Kopplungskarten und Sequenzen differieren (ein wenig)



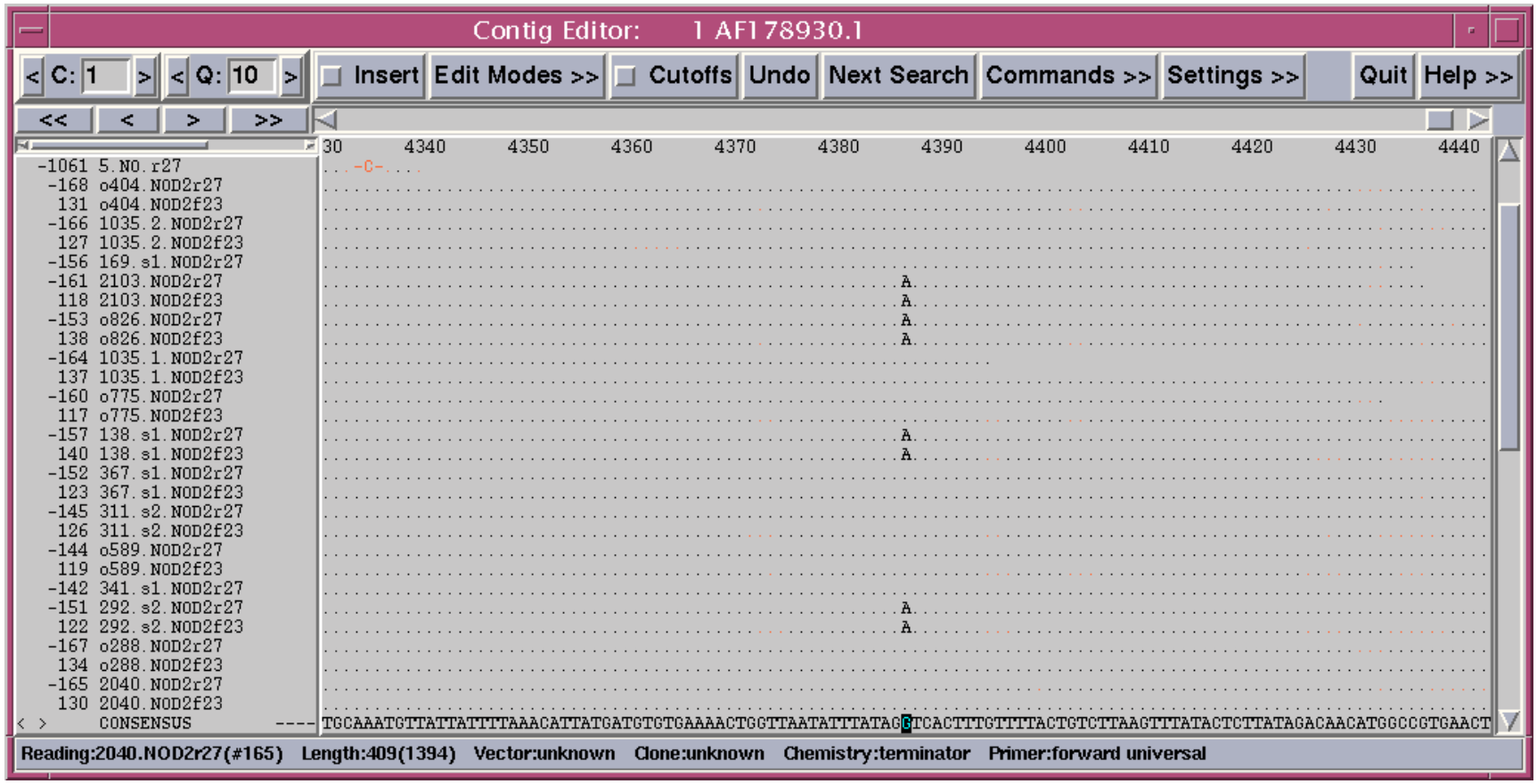
Genomanalyse

Eine „ultimative“ Genomstruktur: Die Nukleotidsequenz



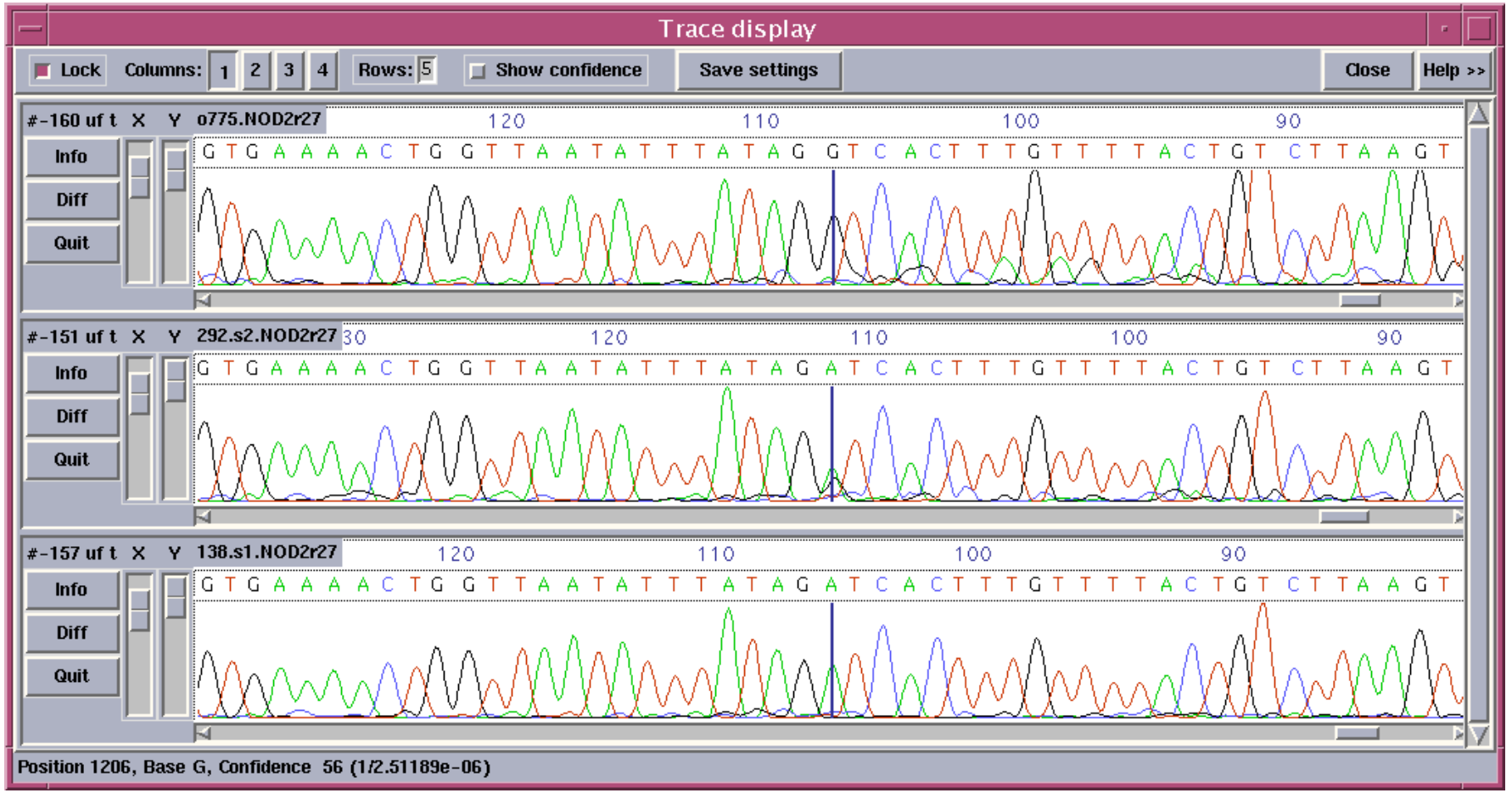
Genomanalyse

Eine „ultimative“ Genomstruktur: Die Nukleotidsequenz



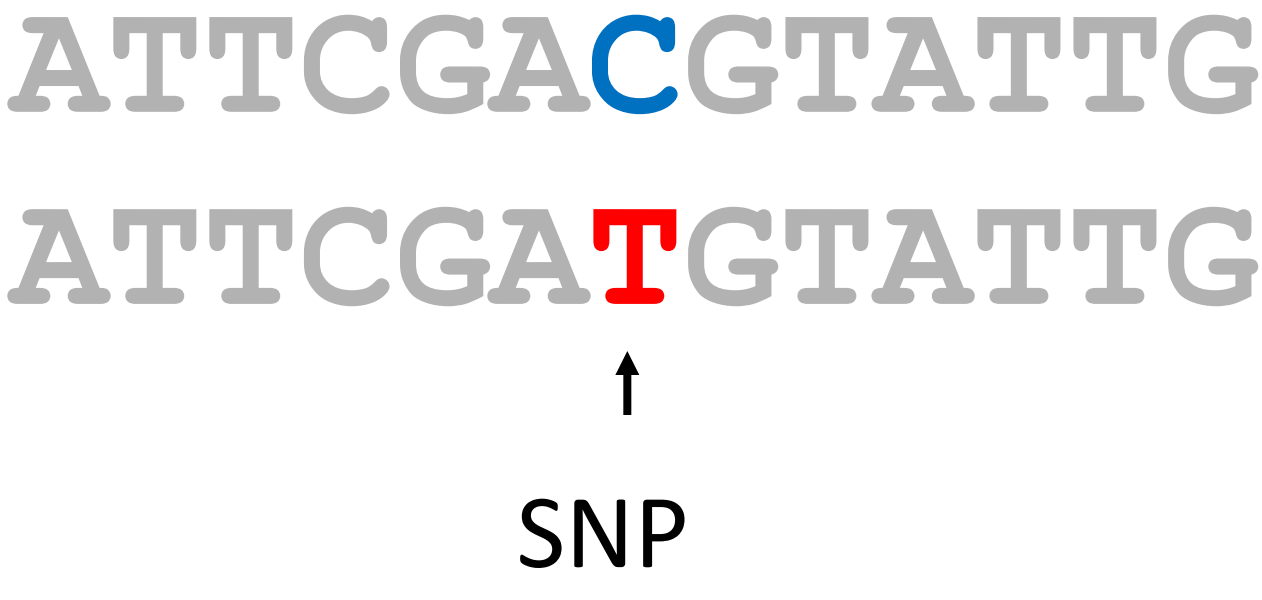
Genomanalyse

Genomsequenzen sind individuell



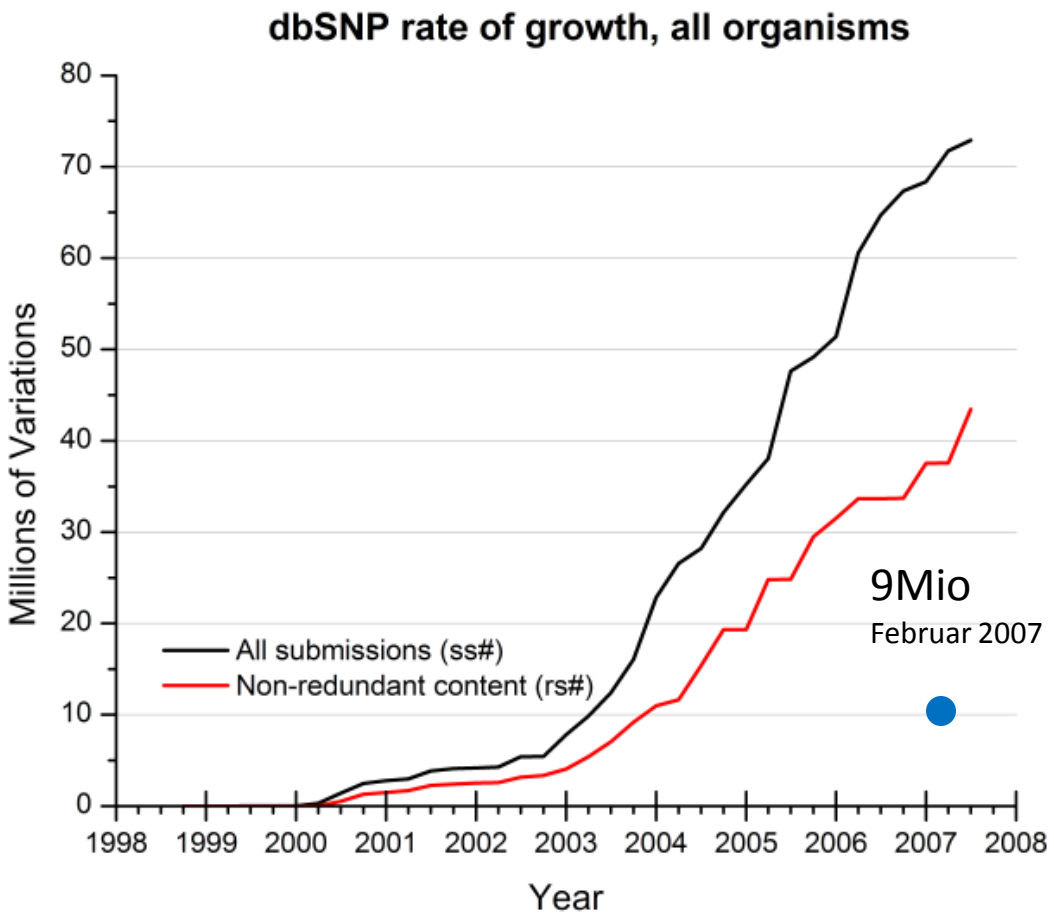
Genomanalyse

Einzelnukleotidvariationen: SNPs



Genomanalyse

SNP Datenbanken



Genomanalyse

Die „ultimative“ Genomstruktur: Sequenz *und* SNP-Katalog

AAGAGCCAGTCACACGAGGACTGGCTGTGTCTACCTTGCCCCGCGCCGTGGGCATCCCCAGCACCAAGCATGGTGCCT
GGCATGCAGTGAACCTTCAGTGTCACCTGACCGAGTAAACACAGGACTGAAGGAAGGGAAGGCGAGATGGATGCAAAR
GGCCCAGTGTGCCTCAGTCTTCTTTGGGAGGCAGCGGGGACCAGGGACATTTTTGAGGTGGTTTCCTCTCCAGGGATT
AACCGAGTTTATCCCTCCTCAGTTCTCCCAGGCTATGGGATGGATGCCAAGGGAMCTTCCAGACAAAMTATTCTGAAG
TCTGGGCCTGTGAGAAGCATCTTGAGAGATCTTTATGCTGCTGTGAGGATTCTAGGACCTGGAGGTCTGTCAGGGCAA
TCAGGAATGCTCCCTGGAGGAAGTGGCAGTGAGGCAGAGCTTTGTGGAGGGGCAGCAGGGGAGAGAGTGCTGTGGCCA
RAGGAAGGGAAAGGGCCATGGCAGGGAAGCAGGCCAGACACTGTTAGGCCTATTAGGAGAGTGTGAAGGGTAATTGGA
TCTGAGTGGGAGGAGCGGCTGTGGTAACTCCTTGTCTCAACGCCTCTGACTAAAGTGGCCAGGCACACCACCCAGGCA
CCTGCCGCCCTTGCCCCAGGGGCTGCTGCAACACCCAACTCGGCTTTAGCCAGAGGGCACTGGGCCAGGCAGATATGG
GGTCCTTCTCCCTGAAAGACAGGGAAGGAAGGTTCCTGCCCCAACCCCAGGCTCTGCCCAGGCTCGTGGAGAAGCCTG
TCCCCGAGGCAGTGGGCGGAACCAGGATGCTCTCAGTCACAGGGCAGCCAAGCGGCTGGGGCGGCAGCTCTTCATCAT
CTTCAGTTGCATTTGACAGCTGTGGCTGACCAGGAGGGGGCAGCTAGGCAGAAGTGTCATCCCGGTCACACCAGTGTC
ACCTGAGGCTCATCAGGGCAGTGACTTGCCCGAGGTCTCCCGGAGCAGCTCAGTAAGGATTGAGCATTTGTTTACTCA
TCTCCATCGCTGACAGCTGAAACTGGATTCATCCTGCCCTGCCCAGGGGGCTGGGAATCAGGGATCGGGGCCAGCTCA
GCTGAGATGTGAGGGTCCCAAATAGTATTTCCACMCCTGACCCCAAAATAGGCCCCAATCCCCTGTCCCAAAGACTTG
ACGTGCCCATTGATGTGTCTGTCGTTCCTGACAGGGTGTCCTCTTGGTTCCTCAGGTCAGAAACCCAAGGGTCATCTC
TGTCCCTGGGCTCTTCCTCACTGCCCACATCCAGTCCATCAGCGAAGTGTGTTAGTTGGCTACCTGCATGCGTCCTGA
ATCCGTCTCCTTCTCCTCCTCCATCTCTCTGCTGGACCACAGCCGTGAACTCATGCCCCAGCTTCCAGGCCTTCCTGG
TCCAGTCCATCTTCCTCTCAAAAGCCCATTGATCTCCAAAACACAAGTCTGAGCAAGGCTCTCCCCTCTTGAAATCCT
TCAATGCCTCCCCGTTACTCTCAGGACGAAGCCCACATTTCTGCCTACTTCTCCAGGCTCAACTTTTTCCTCCAGCTC
CATCACCCACTATATCATTTTGTCCCAGCCACATGGACCTTCTTTCAGTTCCTTGAAAATGCCGGGCCTTTACTGATG
CCCAAGCCTTCGTACAAGCTGTTCTTCCTRTCTGGAACATTCTTTCCTGTTCCCCCGTCTCACTTCCTTTCACTCCTA
CTYTTATTTTAGAGTGCAGCTTGAACCCTTACTCCTTCCTTCCTCTATGTCCCCAGAGTTCACTGACTTCCCTCATCT

Genomanalyse

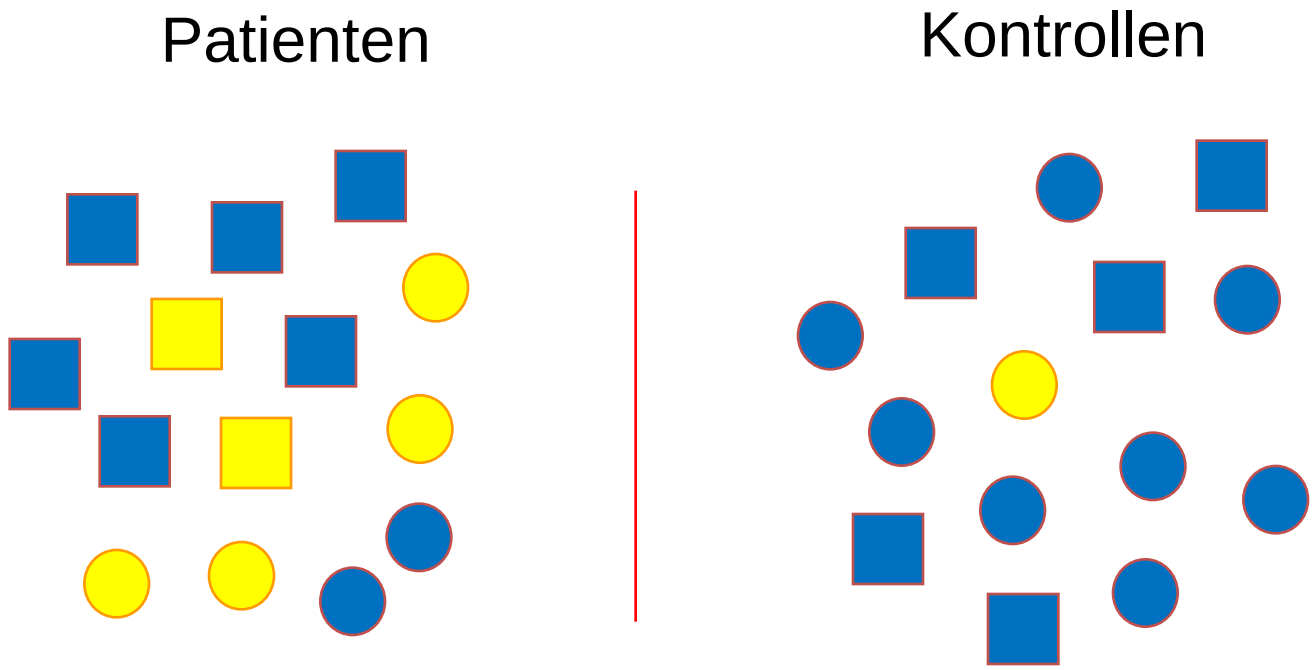
Die „ultimative“ Genomstruktur: Sequenz *und* SNP-Katalog

IUB code: M = aMino (A/C); R = puRin (G/A); Y = pYrimidin (C/T); …

Genomanalyse

SNPs: wertvolle Marker im Genom

Typisierung von SNPs in Assoziations- und Kopplungsstudien



Genomanalyse

SNPs: wertvolle Marker im Genom

Figure 2 - Manhattan plot for the 402,951 SNPs from the stage 1 genome-wide association meta-analysis of the WTCCC-T2D, DGI, WTCCC-HT, WTCCC-CAD, EPIC-Obesity and WTCCC-UKBS studies.

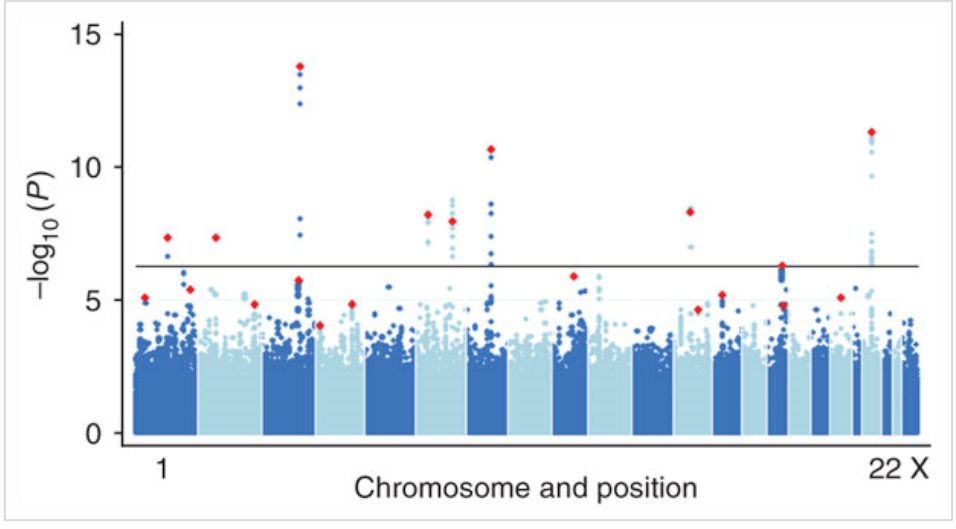
From the following article

Genome-wide association analysis identifies 20 loci that influence adult height

Michael N Weedon, Hana Lango, Cecilia M Lindgren, Chris Wallace, David M Evans, Massimo Mangino, Rachel M Freathy, John R B Perry, Suzanne Stevens, Alistair S Hall, Nilesh J Samani, Beverly Shields, Inga Prokopenko, Martin Farrall, Anna Dominiczak, Diabetes Genetics Initiative, The Wellcome Trust Case Control Consortium, Toby Johnson, Sven Bergmann, Jacques S Beckmann, Peter Vollenweider, Dawn M Waterworth, Vincent Mooser, Colin N A Palmer, Andrew D Morris, Willem H Ouwehand, Cambridge GEM Consortium, Mark Caulfield, Patricia B Munroe, Andrew T Hattersley, Mark I McCarthy & Timothy M Frayling

Nature Genetics **40**, 575 - 583 (2008) Published online: 6 April 2008

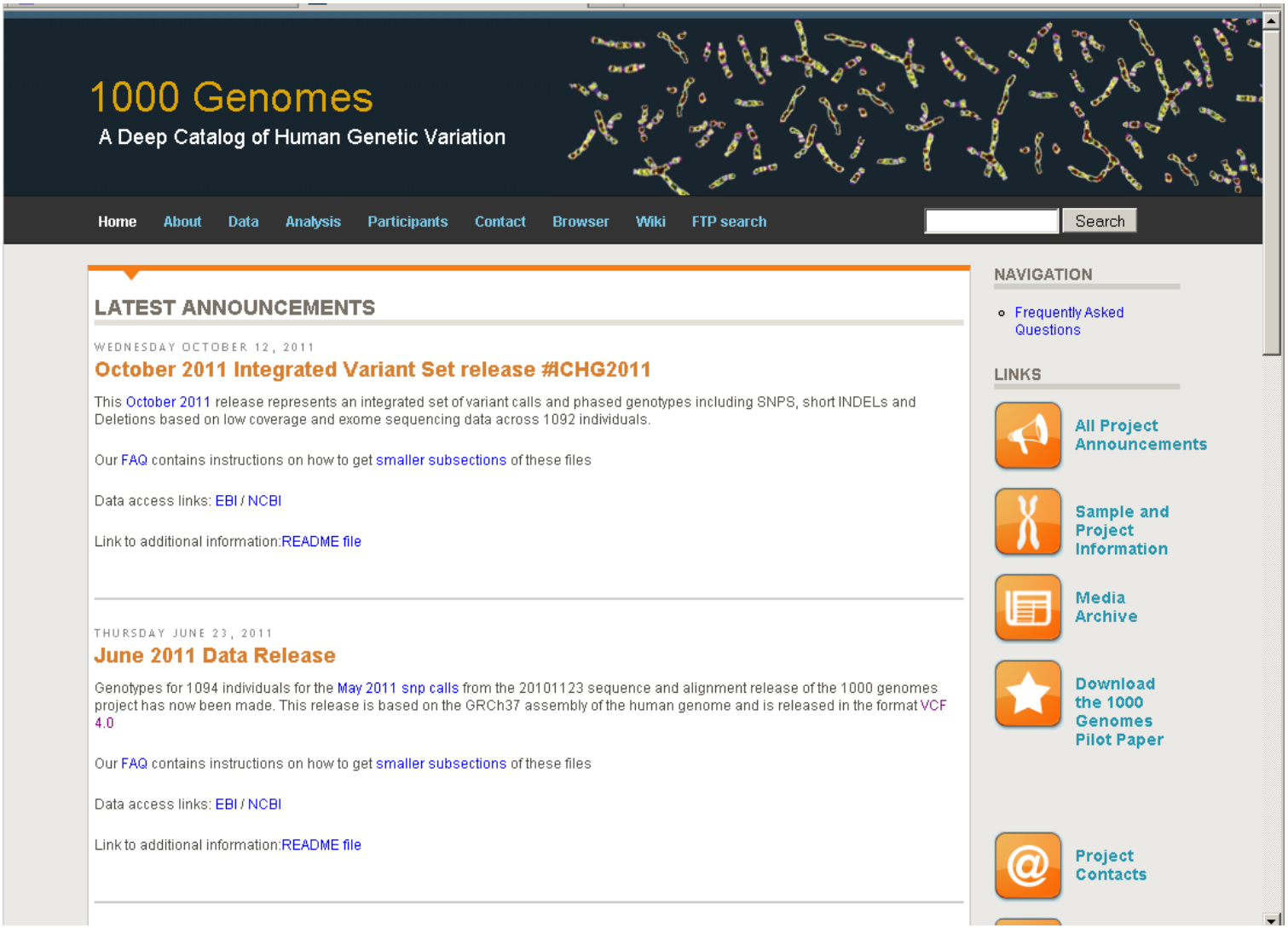
doi:10.1038/ng.121



The red dots represent the SNPs that reached a $P < 5 \times 10^{-7}$ in a joint analysis of stage 1 and stage 2 samples.

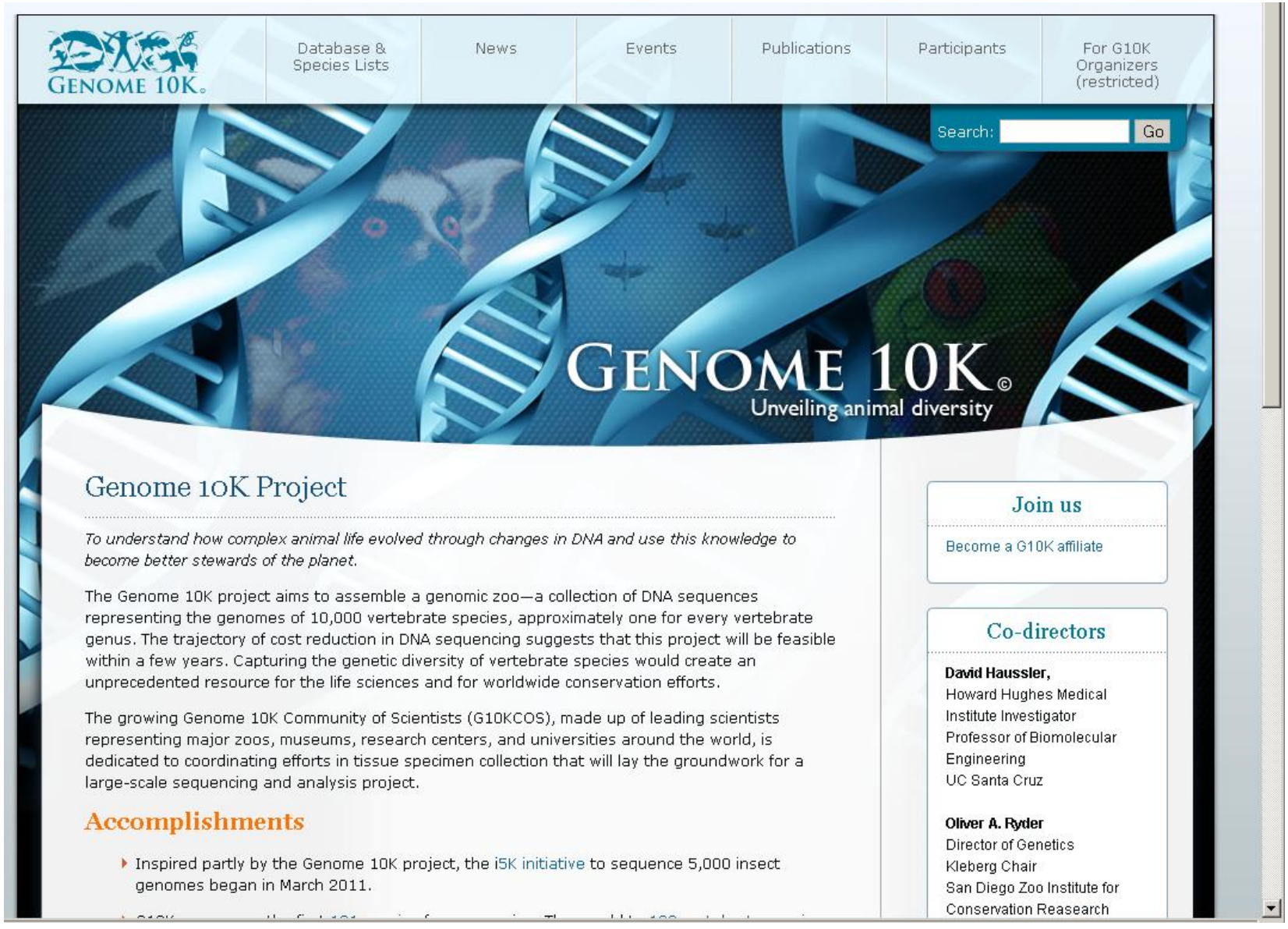
Genomanalyse

1000 Genome Project:: http://www.1000genomes.org/



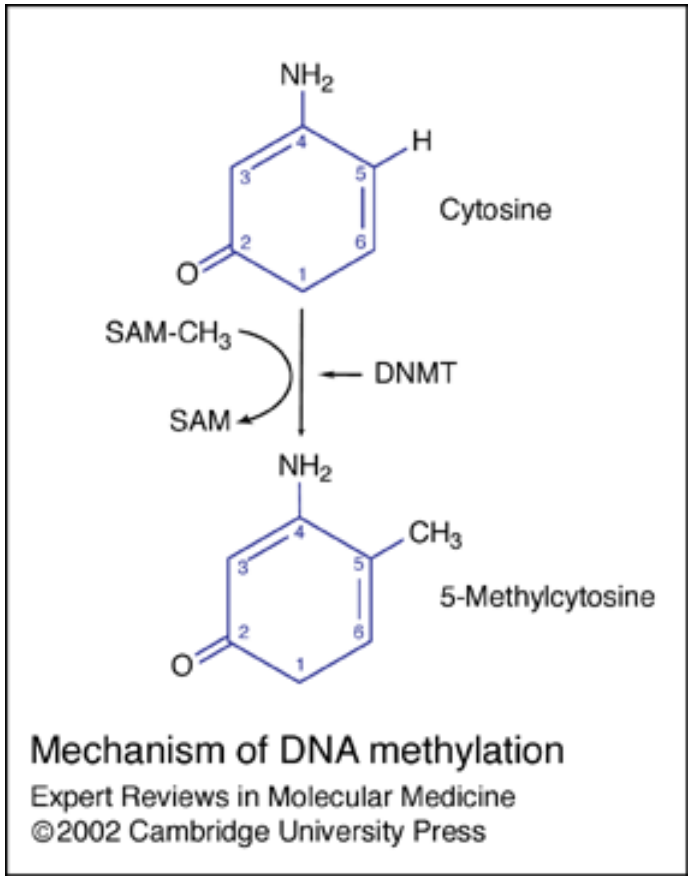
Genomanalyse

10000 Genome Project:: http://www.genome10k.org/



Genomanalyse

Sequenz ist nicht alles: Die fünfte Base



Mechanism of DNA methylation
Expert Reviews in Molecular Medicine

Durch DNA-Methylierung entsteht aus Cytosin die “fünfte Base”

Methylcytosin

Die Verteilung von Methylcytosin im Genom ist zell- und gewebespezifisch und stark umweltabhängig

selbst eineiige Zwillinge sind an ihrem Methylierungsmuster unterscheidbar

Genomanalyse

Epigenomik

Methylierung ist sinnvoll

sie dient zur Inaktivierung von Chromosomen (Lyonisation), zur Inaktivierung von in das Genom integrierten retroviralen Sequenzen und zur Inaktivierung einzelner Gene (Imprinting).

Methylierung ist fatal

wenn sie z.B. Tumorsupressorgene betrifft. Genau dies passiert in der Cancerogenese.

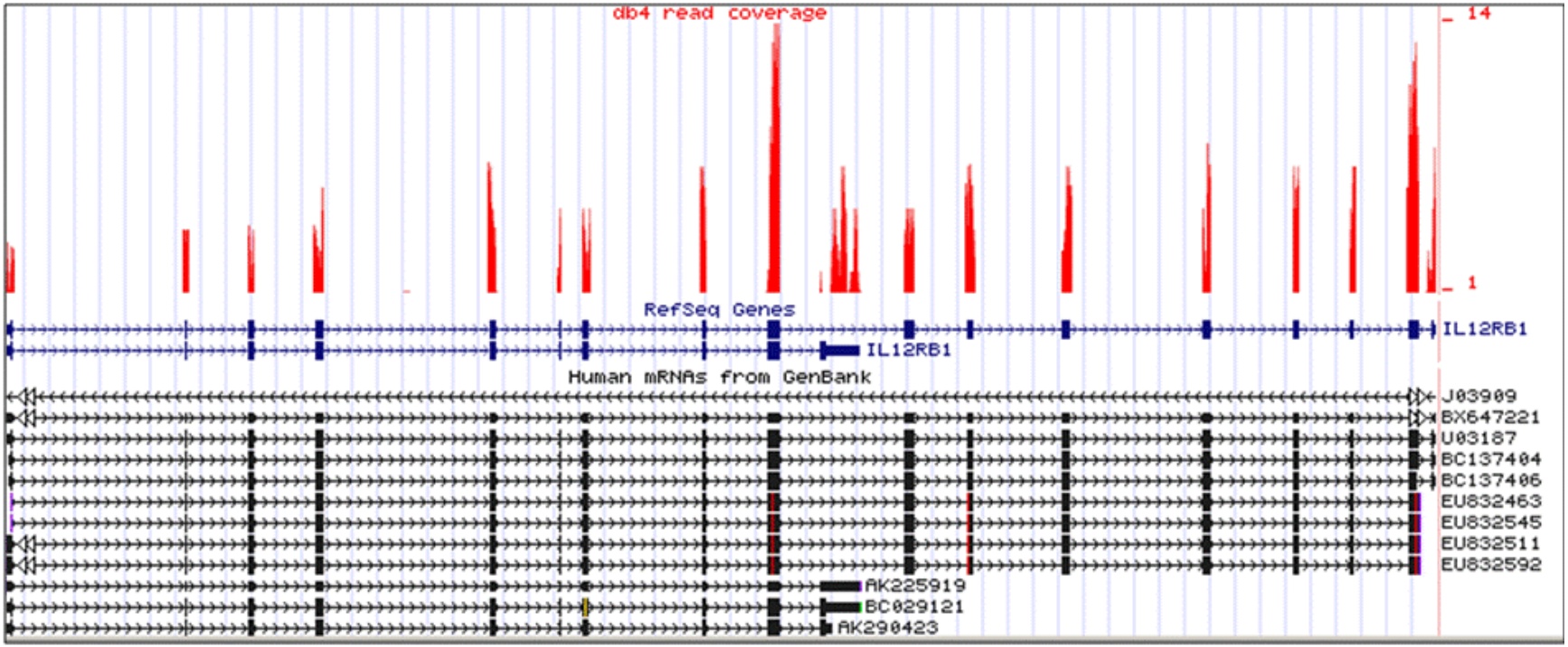
Die Kenntnis des Methylierungsmusters eines Genoms (des Methyloms bzw. Epigenoms) wird immer wichtiger!

Genomanalyse

Was nun?

Nach der Strukturanalyse die Funktionsanalyse

Genstrukturen – Genexpression - Genregulation



Vielen Dank!