

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

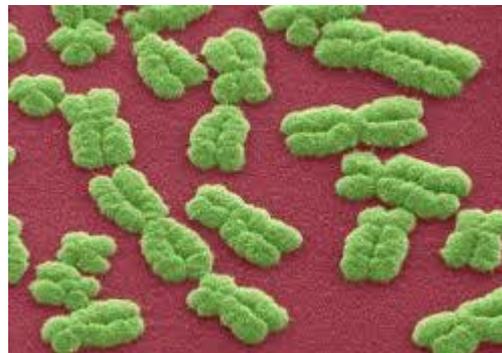
Genom (Winkler, 1920) - gesamtes genetisches Material einer Zelle oder eines Individuums
(nukleares, mitochondriales,... Genom)

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

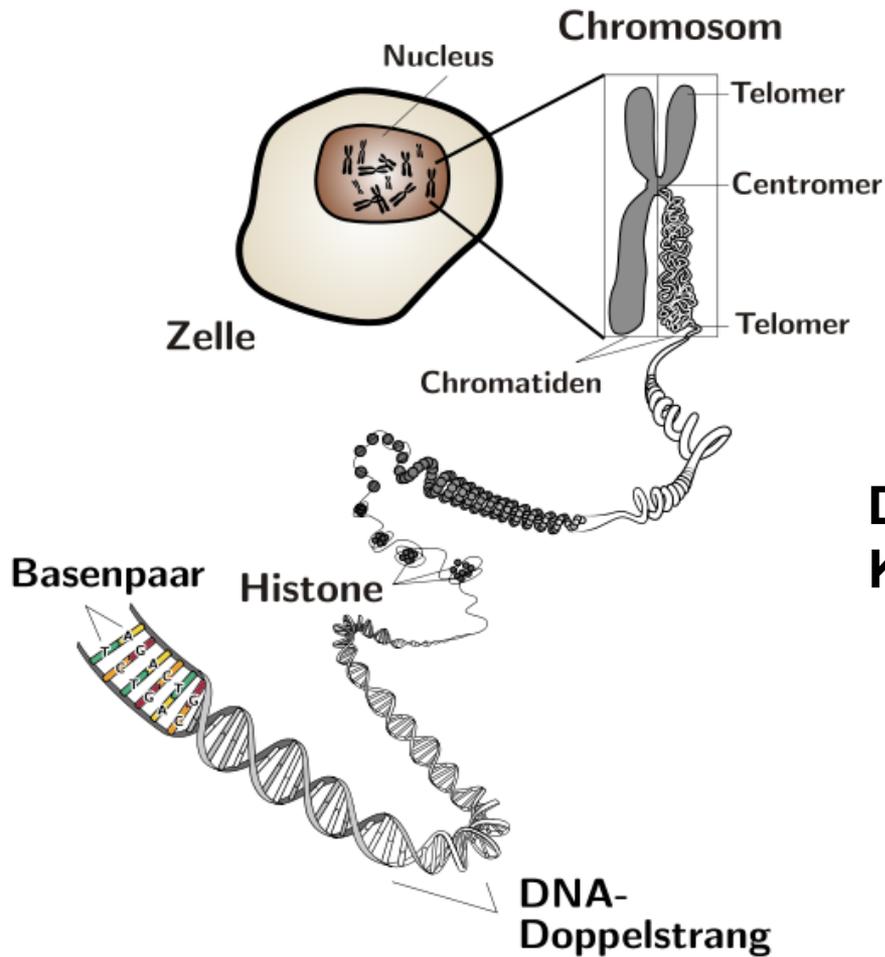
Talmud:

...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

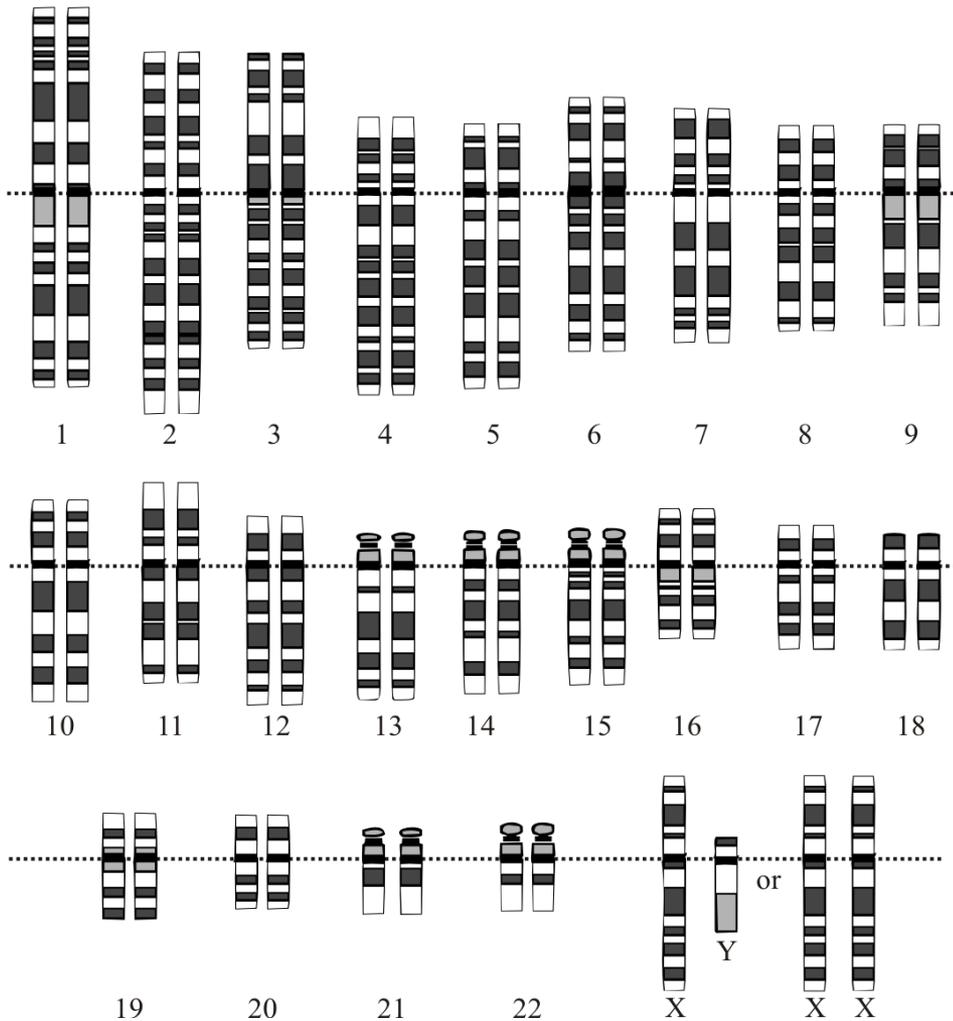
Spezies

Geschlecht

Anomalien (Krankheitsassoziationen)

Genomanalyse

Metaphase-Chromosomen: Karyotypisierung



46, XX



46, XY

Genomanalyse

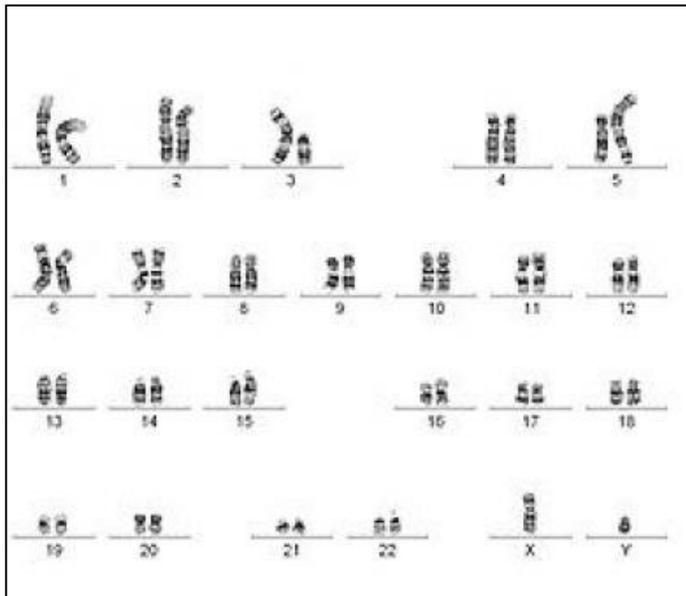
Metaphase-Chromosomen: Karyotypisierung



47, XY + 21



Down Syndrom



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



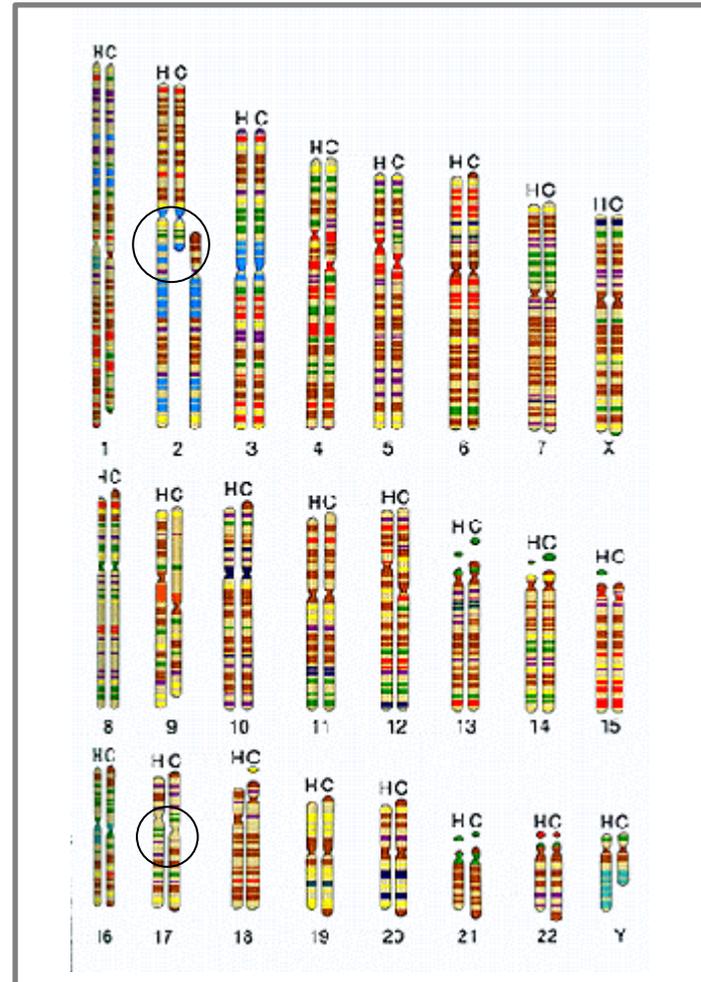
akute myeloische Leukämie

Genomanalyse

Chromosomensatz

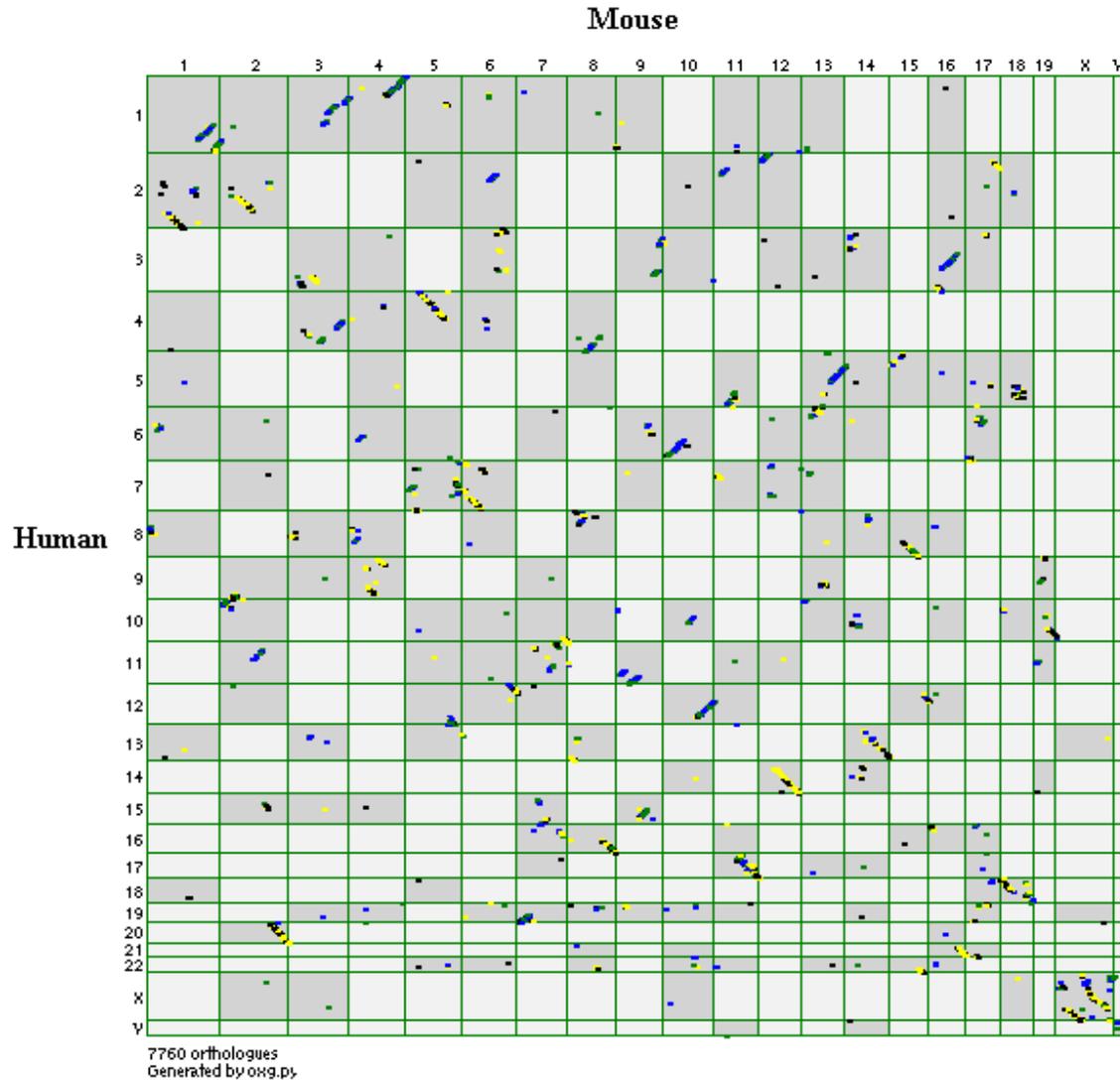
Homo sapiens: 46
Pan troglodytes: 48

9 intrachromosomale Inversionen
1 Chromosomenfusion



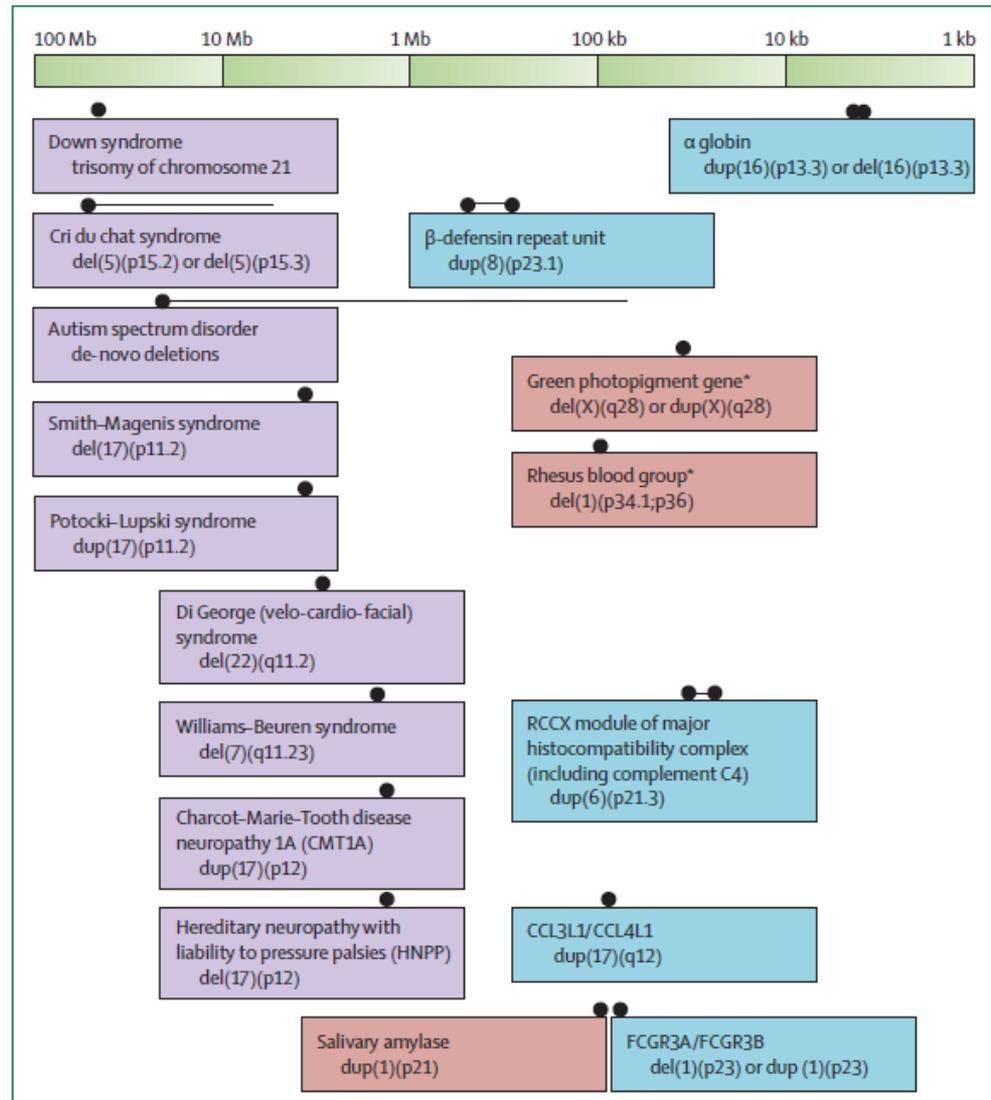
Genomanalyse

Comparative Genomanalyse: Syntenie



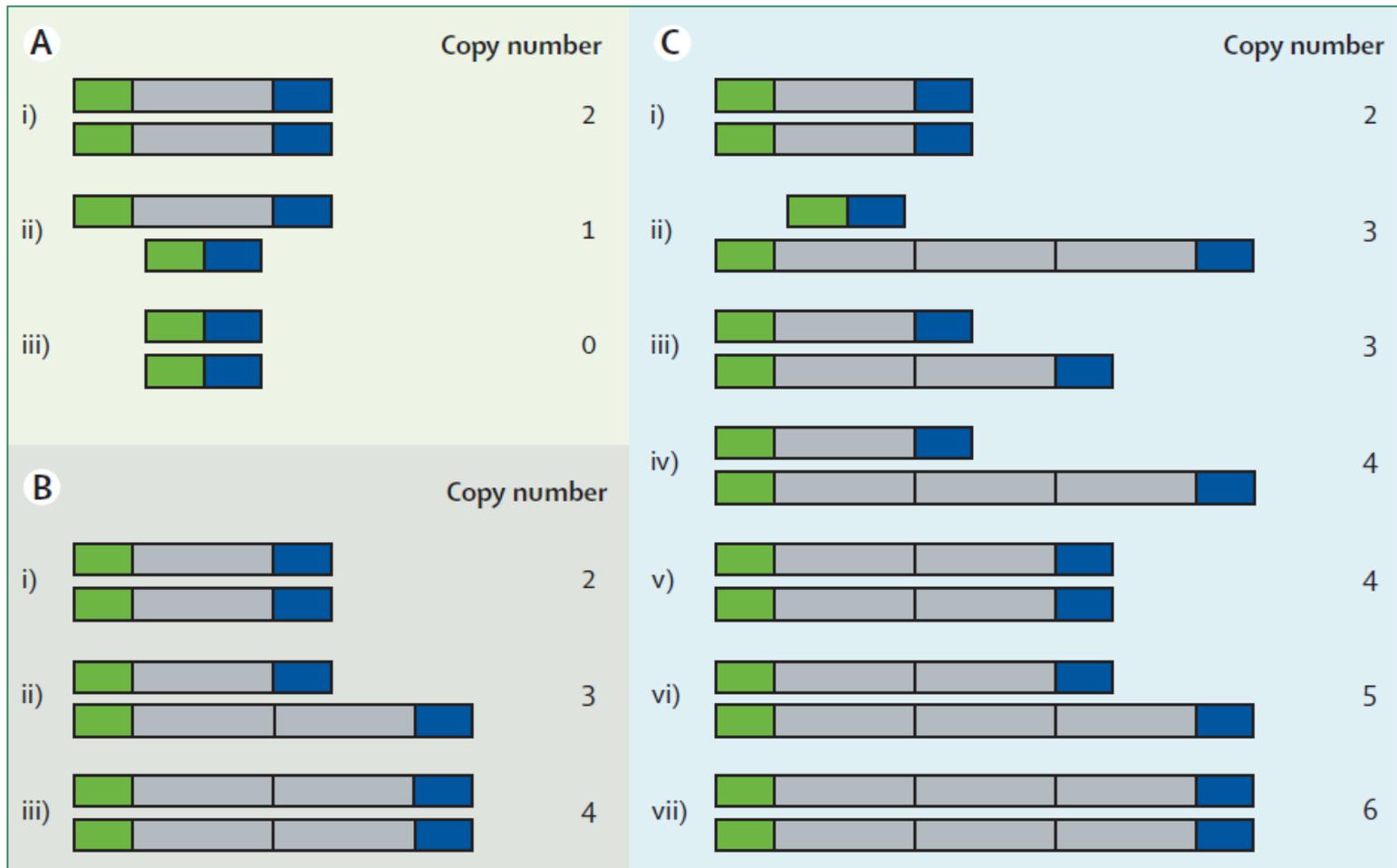
Genomanalyse

Strukturvarianten



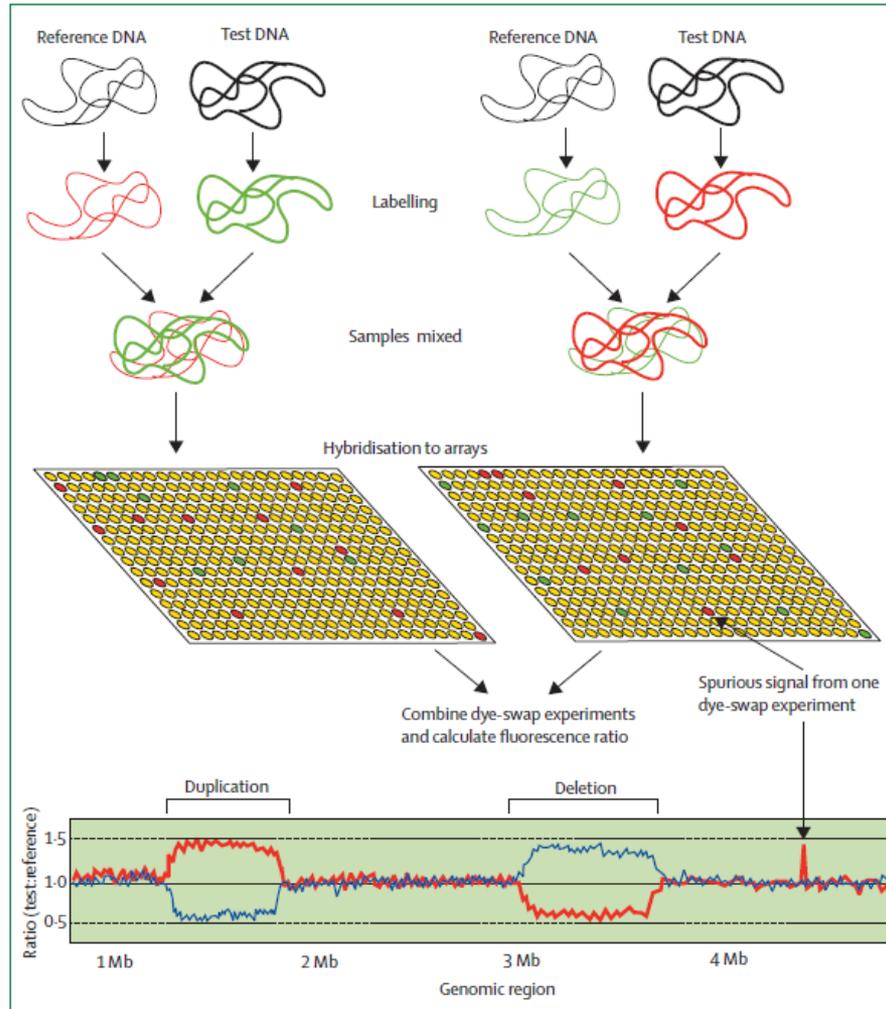
Genomanalyse

Bi- und multiallelische Strukturvarianten



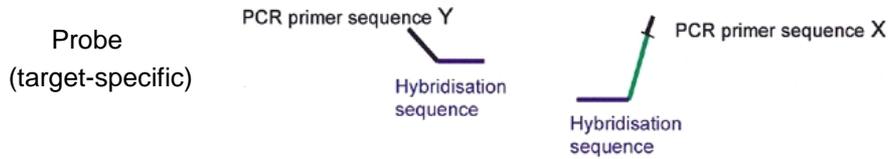
Genomanalyse

Comparative Genom-Hybridisierung

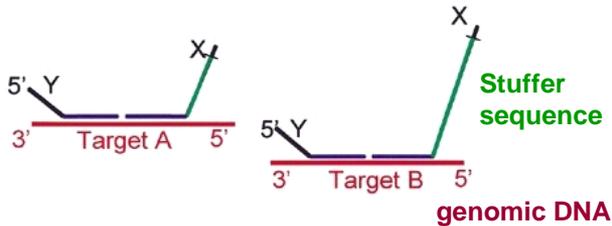


Genomanalyse

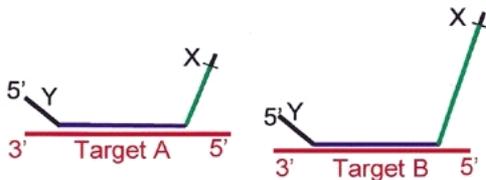
Multiplex Ligation-dependent Probe Amplification (MLPA)



Annealing



Ligation

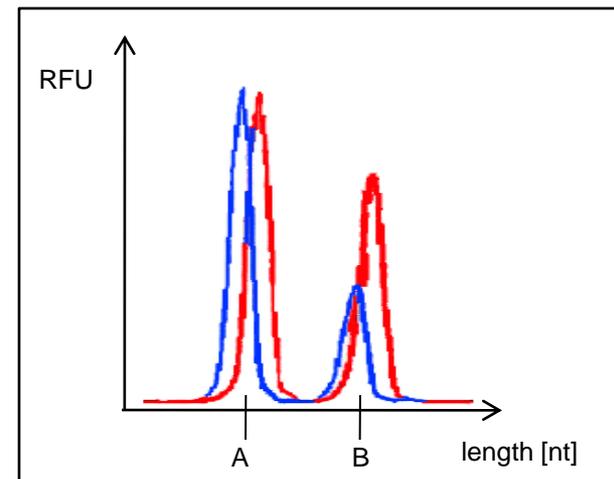


Separation of amplicons & determination of peak areas

Amplifikation



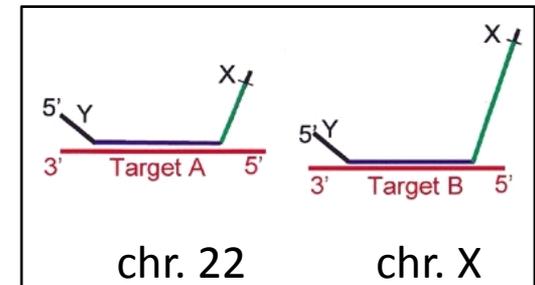
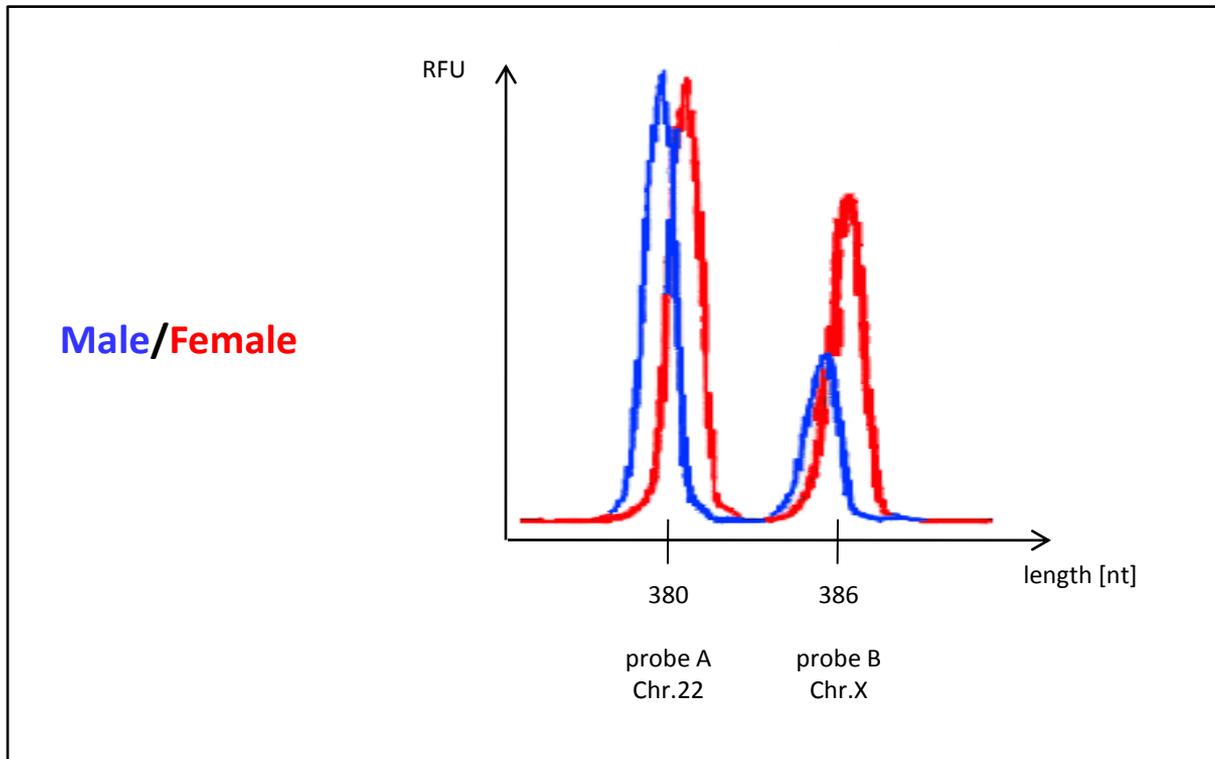
adapted from Schouten et al., Nucleic Acids Res. 2002 Jun 15;30(12):e57.



Genomanalyse

Multiplex Ligation-dependent Probe Amplification (MLPA)

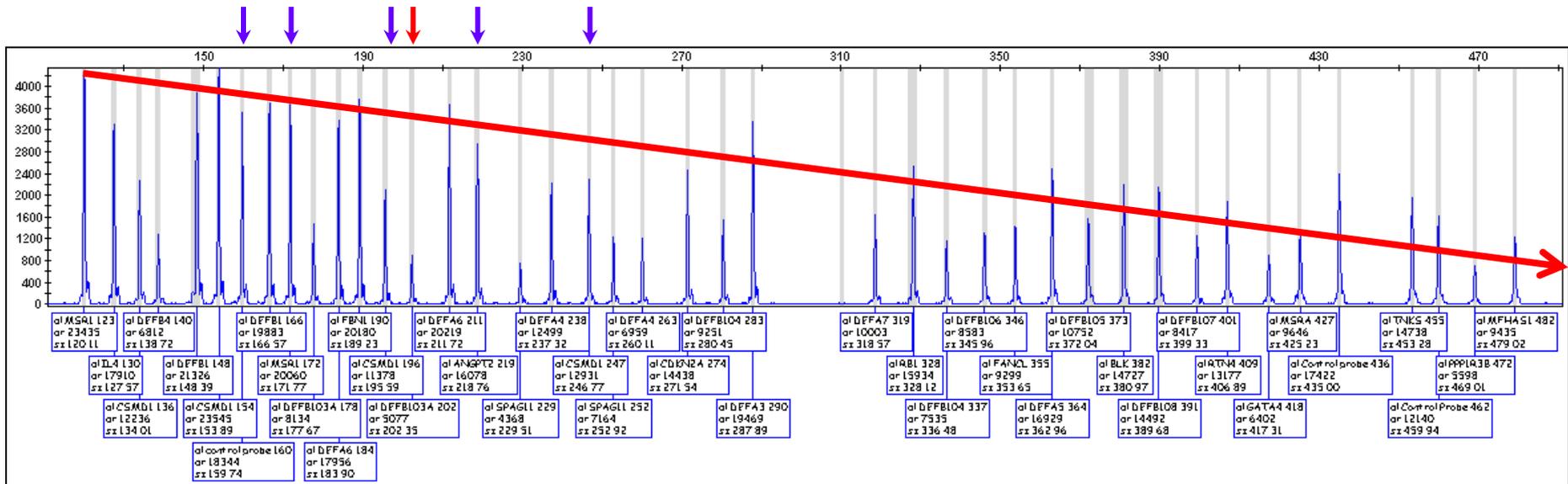
MLPA: DNA-Sonden spezifisch für Autosom und X-Chromosom



Genomanalyse

Multiplex Ligation-dependent Probe Amplification (MLPA)

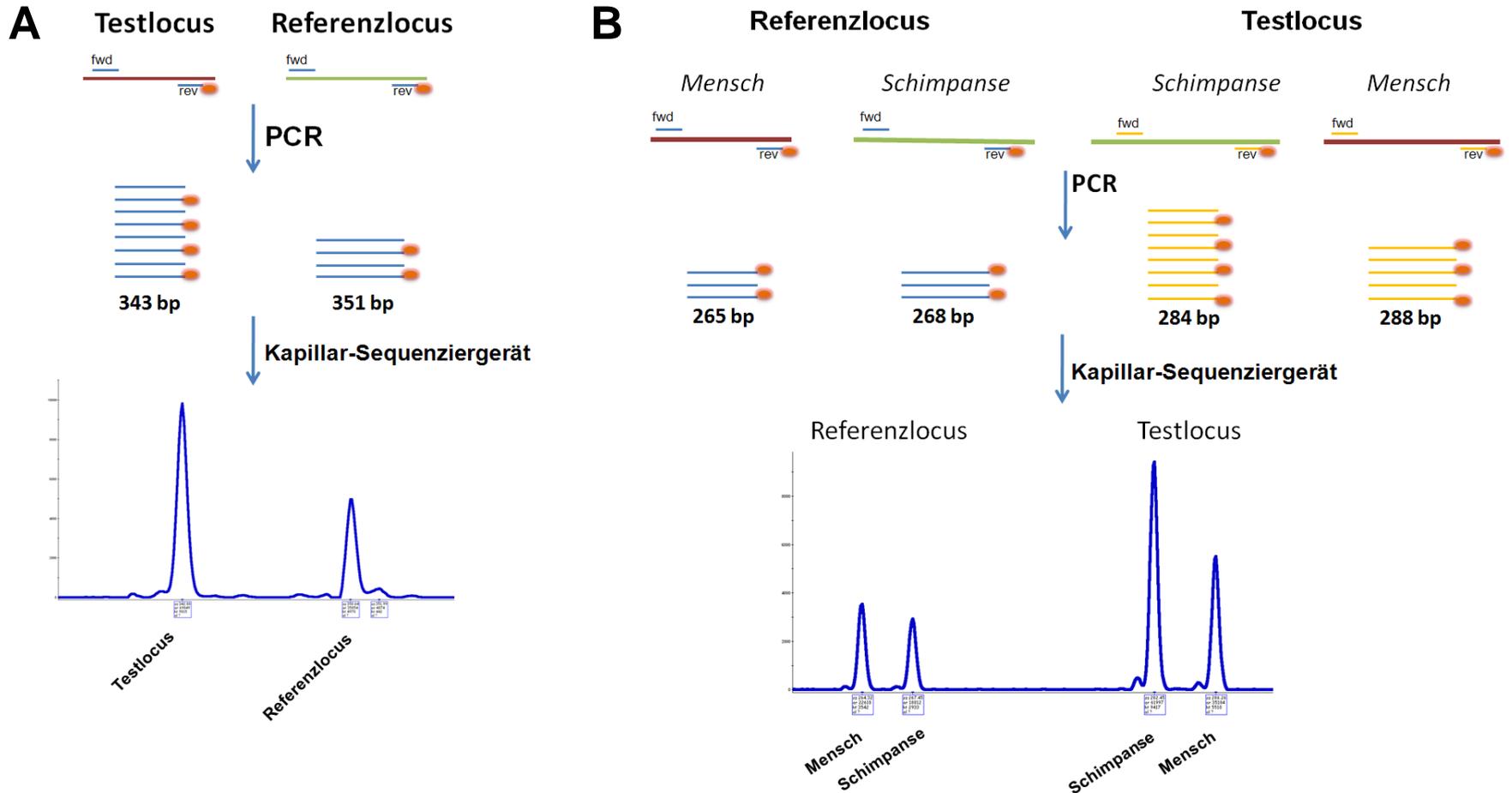
- raw data: length of PCR product and respective peak area



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

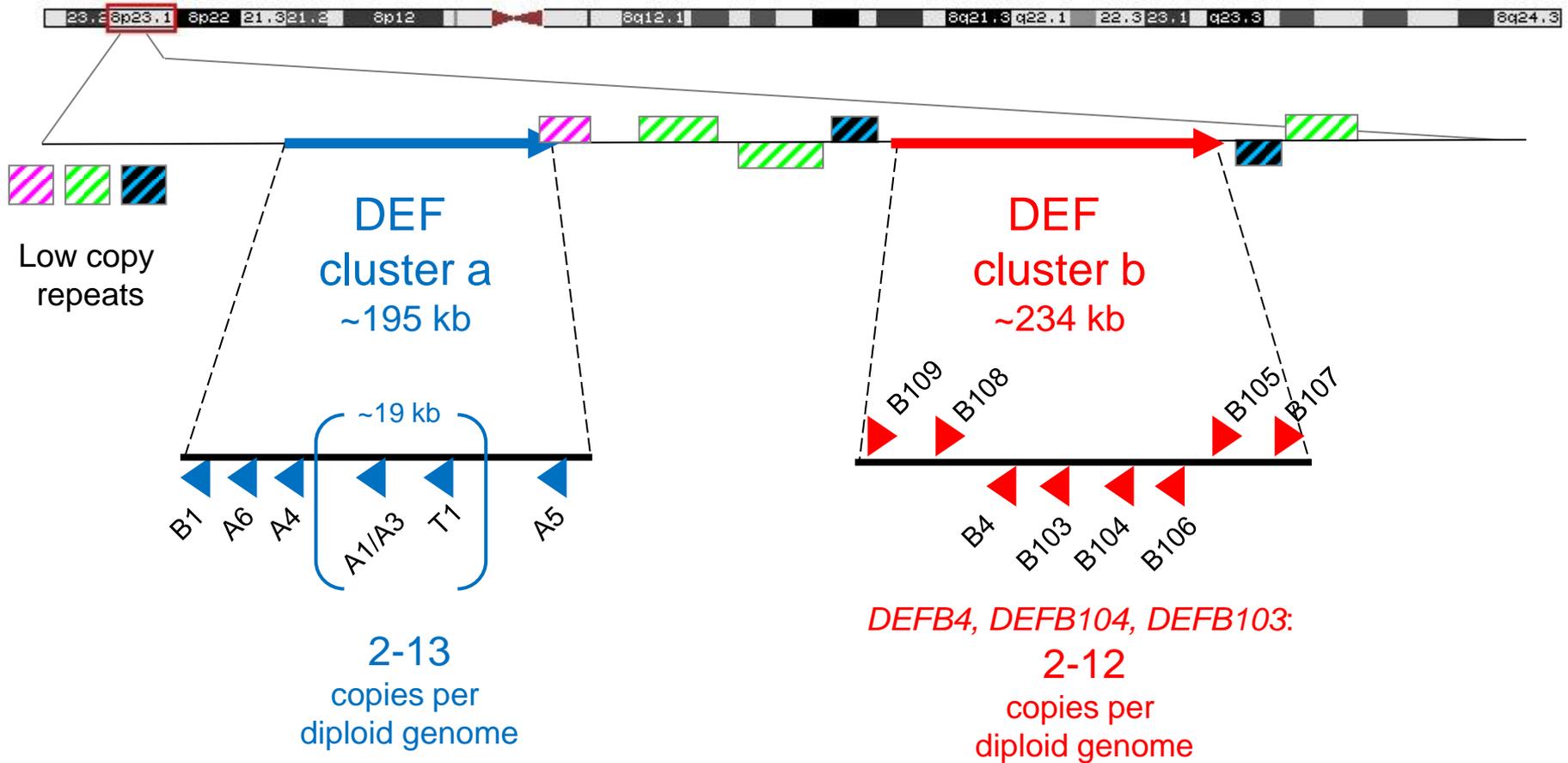
Genomanalyse

Paralog-Ratio-Test/Ortholog-Ratio-Test



Genomanalyse

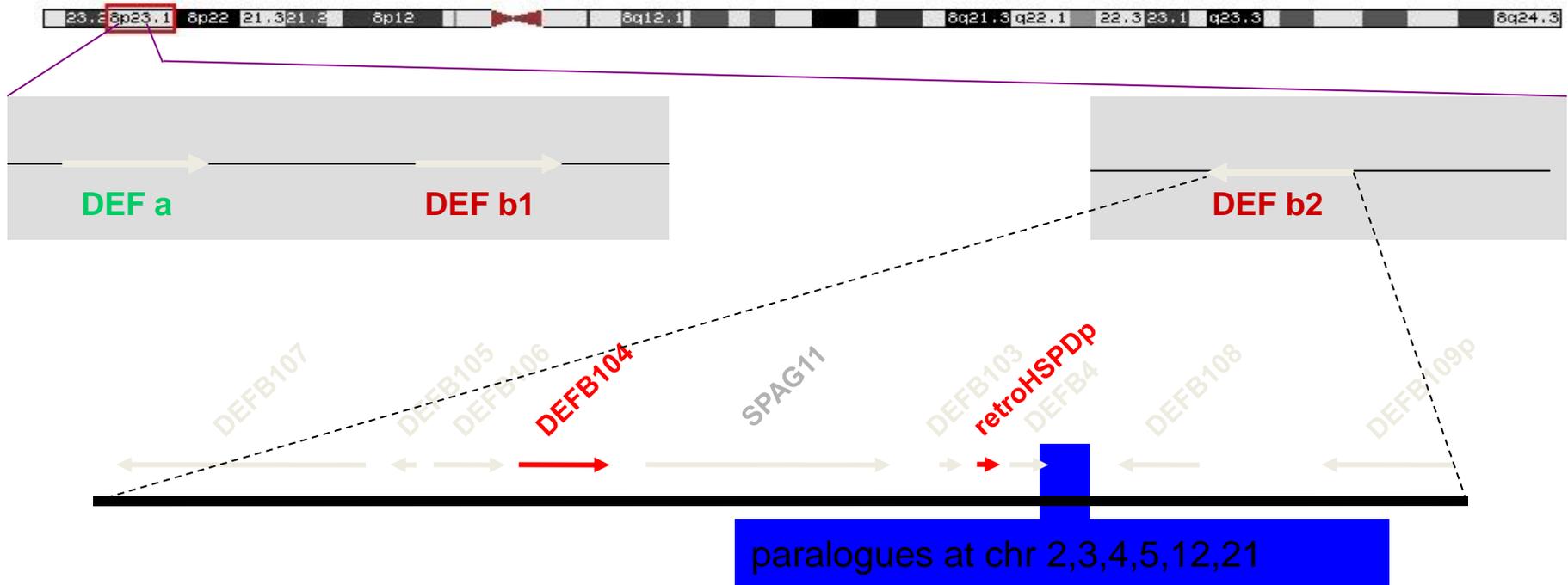
Defensin cluster auf 8p23



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

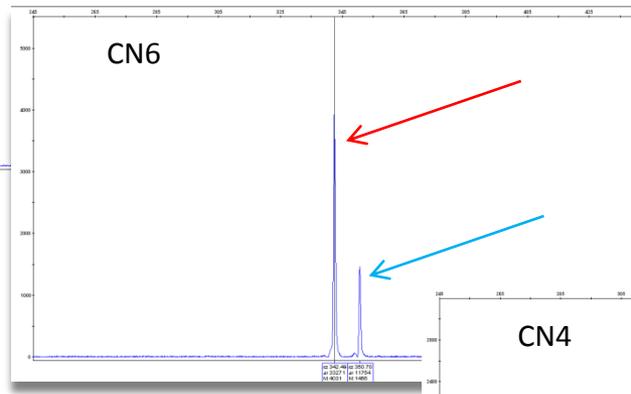
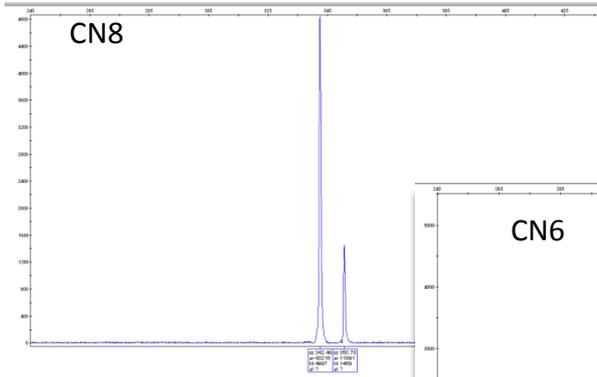
```
chr12 ACAT CACTTGTCCCACCAACCTTC AGCACAGCTACTACATCTGAAAGTTTTGCCAGCCGTTTCAT ... TTACGGT GAGCATTGGCAATTTCAA GAGCAG
chr5 ACAT CACTTGTCCCACCAA CTTC AGCACAACTACTCCATCTGAAAGTTTTGCCAGCCATTCAT ... TTATGGT GAGCATTGGCAATTTCAAGAGCAG
chr3 ACAT CACTTGTCCCACCAACCTT TGGCATAGCTATCCCATCTGAAAGTTTTGCCAG----TCAT ... TTACGGT GAGCATTGGCAGTTTCAA GAGCAG
chr4 GCAT CACTTGTCCCACCAACCTTC AGCACAGCTACTCCATCTGAAAGTTTTGCCAGT-GTTCAT ... TTACAGT AAGCATTGGCAATTTCAA GAGCAG
chr8 ACAT CACTTGTCCCACCAACCTTC AGCACAGCTACTCCATCTGAAAGTTTC ----- TCCAT ... TTACAGT AAGCATTGGCAATTTCAA GAGCAG
chr6 ACAT CACTTGTCCCACCA GCCTTC AGCACTGCTACTCCATCTGAAAGTTTTGCCAGACATACAT ... TTACAGT GAGCATTGGCAATTTCAA GAGCAG
chr11 ACAT TACCTGTCCCATGAAC TTTC AGTACAGGTATTCCATCTAAAAGTTTCCCAGATGCTCAT ... TTTCT GTGAGCATCA GCAATTG CAAGGGCAA
chr13 ACAT CACTTGTCCCACCAACCTTC AGCACAGCTACTCCATCTGAAAGTTTTCCAGCCATTCTA ... TTATGGT GAGCATTGGCAA CTTC AAGAGCAG
chr21 ACGT GACTTGTCCCAC TAACCTTC AGCACAGCTACTCCATCTGAAAGTTTTGCCAGCTGTTTCAT ... TTACGGT GAGCATTGGCAATTTCAA GATCAG
chr13 ACAT CACTTGTCTCACCAACCA TC AACACGGCTACTCCATCTGAAAGTTTTGCCAGATATTCAT ... TTAAG GTAAGCATTGGCAATTTCAA GAGCAG
chr5 ACAT CACTTGT TACACCAACCTTC AGCACAGCTACTCCATTTAAAGTTTTGCCA----TTCAT ... GTATA CTGAGCAC TGGCAATTTCAAGAGTAG
chr1 ACAT CATTTT TCCCATCAACCTTC AGCACAGCTATTGCATCC-AAAATTTTGCCGGATGTTCTT ... TTTCA GTGGCATTGGCAATTTT AAG-GTGG
```

fw primer

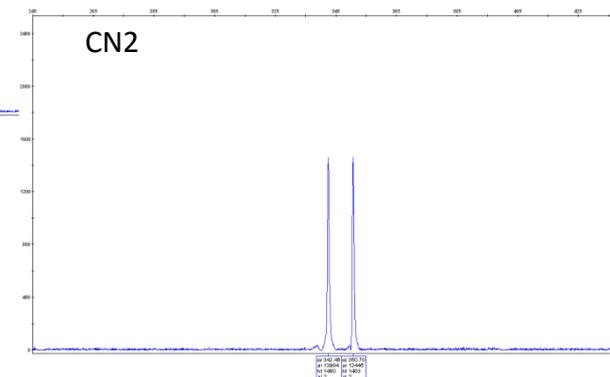
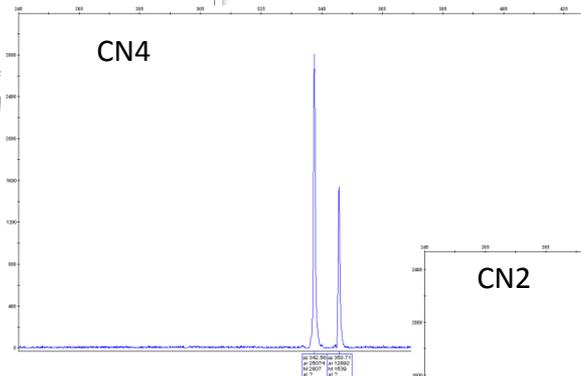
rev primer

Genomanalyse

Paralog-Ratio-Test für Defensincluster auf Chromosom 8



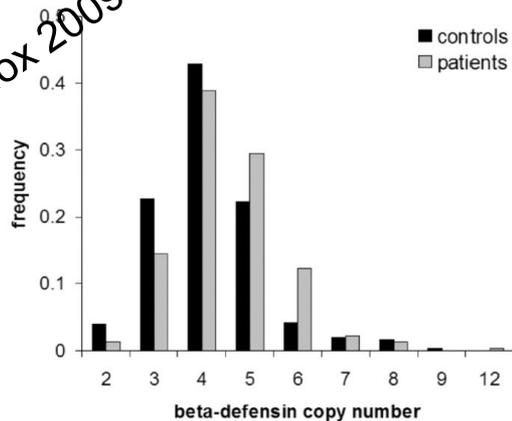
Chromosome 8 (DEF cluster)
Chromosome 4 (Referenz)



Genomanalyse

Assoziation von Kopienzahl und Psoriasis

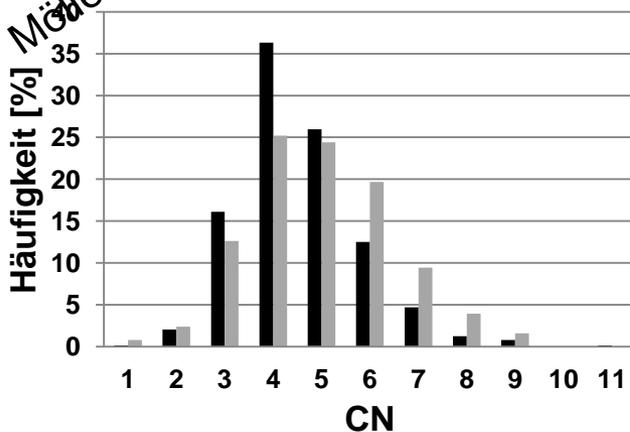
Hollox 2009



Patienten mit Psoriasis weisen höhere Kopienzahlen auf

Morbus Crohn
Prostatakarzinom
Pankreaskarzinom
Arthritis

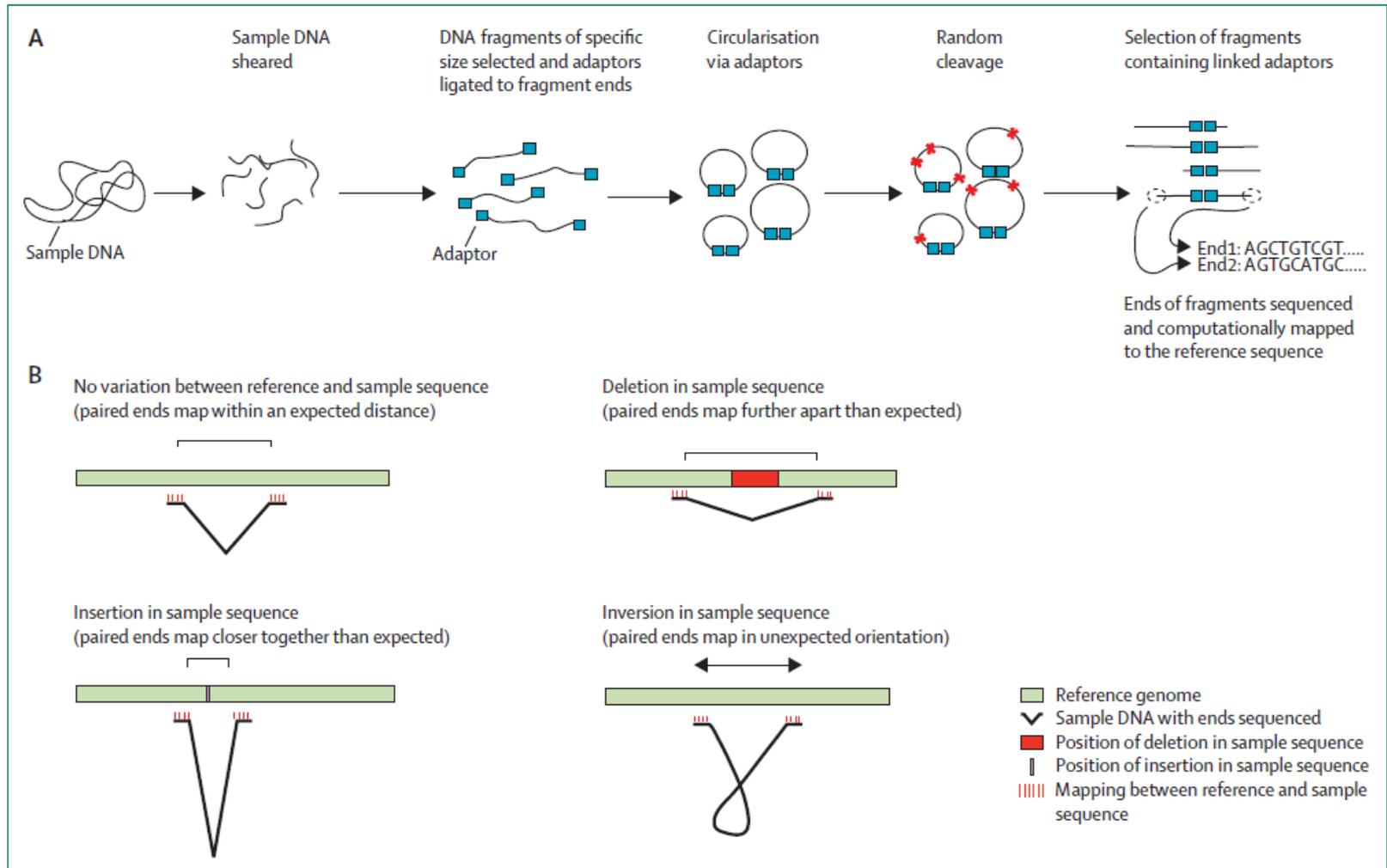
Möller 2009



Defensine sind an der Regulation der angeborenen Immunantwort beteiligt!

Genomanalyse

NextGen Sequencing und Strukturvarianten



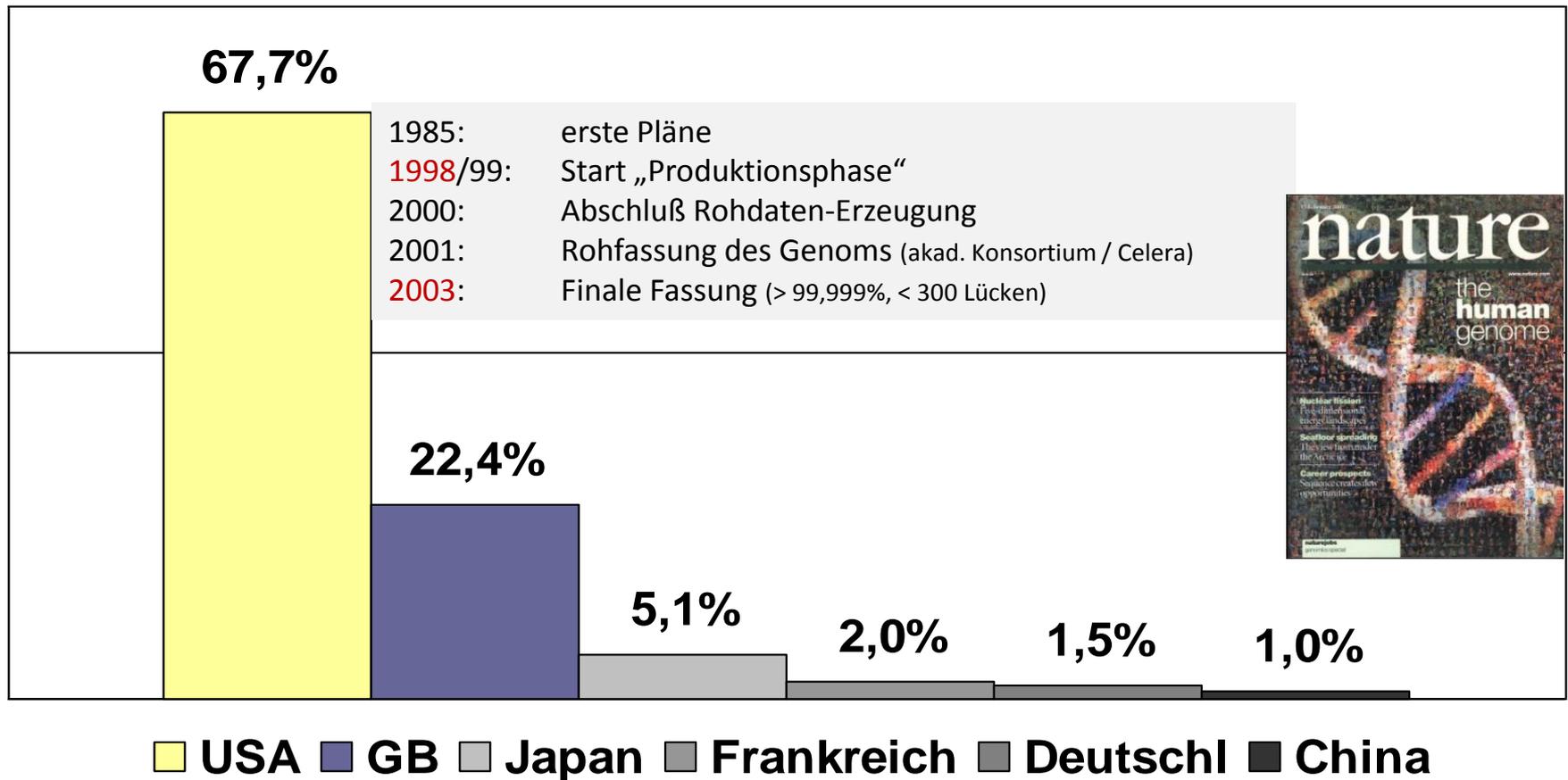
Genomanalyse

Eine „ultimative“ Genomstruktur: Die Nukleotidsequenz

AAGAGCCAGTCACACGAGGACTGGCTGTGTCTACCTTGCCCCGCGCCGTGGGCATCCCCAGCACCAAGCATGGTGCCT
GGCATGCAGTGAACCTTCAGTGTACCTGACCGAGTAAACACAGGACTGAAGGAAGGGAAGGCGAGATGGATGCAAAR
GGCCCAGTGTGCCTCAGTCTTCTTTGGGAGGCAGCGGGGACCAGGGACATTTTTGAGGTGGTTTCCCTCTCCAGGGATT
AACCGAGTTTATCCCTCCTCAGTTCTCCAGGCTATGGGATGGATGCCAAGGGACCTTCCAGACAAACTATTCTGAAG
TCTGGGCCTGTGAGAAGCATCTTGAGAGATCTTTATGCTGCTGTGAGGATTCTAGGACCTGGAGGTCTGTCAGGGCAA
TCAGGAATGCTCCCTGGAGGAAGTGGCAGTGAGGCAGAGCTTTGTGGAGGGGCAGCAGGGGAGAGAGTGCTGTGGCCA
GAGGAAGGGAAGGGCCATGGCAGGGAAGCAGGCCAGACACTGTTAGGCCATTAGGAGAGTGTGAAGGGTAATTGGA
TCTGAGTGGGAGGAGCGGCTGTGGTAACTCCTTGTCTCAACGCCTCTGACTAAAGTGGCCAGGCACACCACCCAGGCA
CCTGCCGCCCTTGCCCCAGGGGCTGCTGCAACACCCAACCTCGGCTTTAGCCAGAGGGCACTGGGCCAGGCAGATATGG
GGTCCCTTCTCCCTGAAAGACAGGGAAGGAAGGTTCCCTGCCCAACCCCAGGCTCTGCCCAGGCTCGTGGAGAAGCCTG
TCCCCGAGGCAGTGGGCGGAACCAGGATGCTCTCAGTCACAGGGCAGCCAAGCGGCTGGGGCGGCAGCTTTCATCAT
CTTCAGTTGCATTTGACAGCTGTGGCTGACCAGGAGGGGGCAGCTAGGCAGAAGTGTGCATCCCGGTCACACCAGTGTC
ACCTGAGGCTCATCAGGGCAGTGACTTGCCCCGAGGTCTCCCGGAGCAGCTCAGTAAGGATTGAGCATTTGTTTACTCA
TCTCCATCGCTGACAGCTGAAACTGGATTTCATCCTGCCCTGCCCAGGGGGCTGGGAATCAGGGATCGGGGCCAGCTCA
GCTGAGATGTGAGGGTCCCAAATAGTATTTCCACCCCTGACCCCAAATAGGCCCAAATCCCCTGTCCCAAAGACTTG
ACGTGCCCATTTGATGTGTCTGTCGTTCCCTGACAGGGTGTCTCTTGGTTCCTCAGGTCAGAAACCCAAGGGTCATCTC
TGTCCCTGGGCTCTTCCCTCACTGCCACATCCAGTCCATCAGCGAAGTGTGTTAGTTGGCTACCTGCATGCGTCTGA
ATCCGTCTCCTTCTCCTCCTCCATCTCTCTGCTGGACCACAGCCGTGAACTCATGCCCCAGCTTCCAGGCCTTCCTGG
TCCAGTCCATCTTCCCTCTCAAAGCCCATTGATCTCCAAAACACAAGTCTGAGCAAGGCTCTCCCCTCTTGAAATCCT
TCAATGCCTCCCCGTTACTCTCAGGACGAAGCCACATTTCTGCCTACTTCTCCAGGCTCAACTTTTTCTCCTCCAGCTC
CATCACCCACTATATCATTTTGTCCCAGCCACATGGACCTTCTTTCAGTTCCTTGAAAATGCCGGGCCTTTACTGATG
CCCAAGCCTTCGTACAAGCTGTTCTTCCCTATCTGGAACATTTCTTCCCTGTTCCCCGTCTCACTTCCCTTCACTCCTA
CTYTTATTTTAGAGTGCAGCTTGAACCCTTACTCCTTCCCTCCTCTATGTCCCCAGAGTTCAGTACTTCCCTCATCT

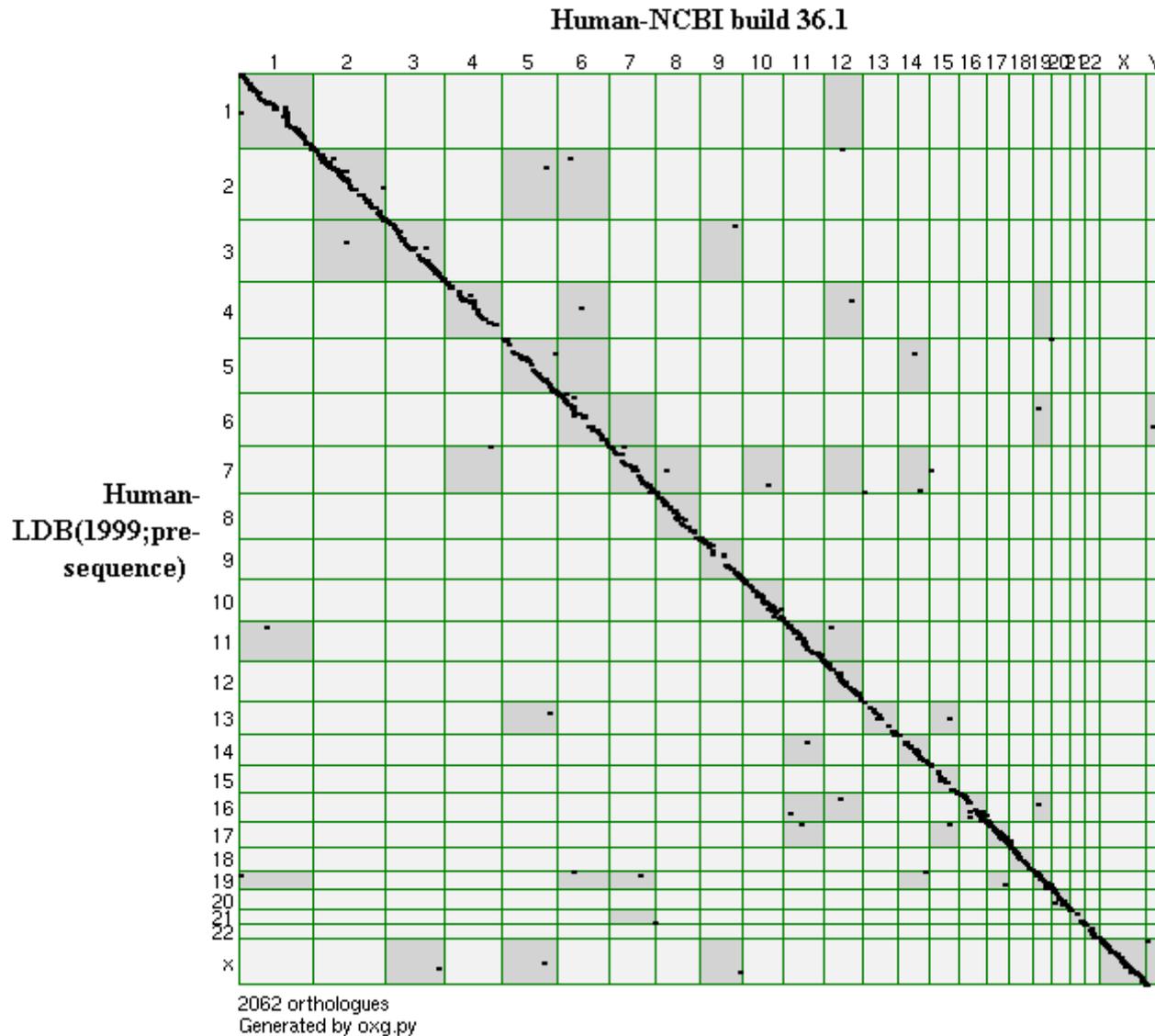
Genomanalyse

Eine „ultimative“ Genomstruktur: Die Nukleotidsequenz



Genomanalyse

Klassische Kopplungskarten und Sequenzen differieren (ein wenig)



Genomanalyse

Eine „ultimative“ Genomstruktur: Die Nukleotidsequenz

Contig Editor: 1 AF178930.1

< C: 1 > < Q: 5 > Insert Edit Modes >> Cutoffs Undo Next Search Commands >> Settings >> Quit Help >>

<< < > >>

	4330	4340	4350	4360	4370	4380	4390	4400	4410	4420	4430
-1061 5. No. r27	CATGCTGTGC-C-TGTT										
-168 o404. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
131 o404. NOD2f23	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-166 1035. 2. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-166 1035. 2. NOD2f23	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-156 169. s1. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-161 2103. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGATCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
118 2103. NOD2f23	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGATCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-153 o826. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGATCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
138 o826. NOD2f23	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGATCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-164 1035. 1. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTG										
137 1035. 1. NOD2f23	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-160 o775. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATG										
117 o775. NOD2f23	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-157 138. s1. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGATCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
140 138. s1. NOD2f23	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGATCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-152 367. s1. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
123 367. s1. NOD2f23	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-145 311. s2. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
126 311. s2. NOD2f23	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-144 o589. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
119 o589. NOD2f23	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-142 341. s1. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-151 292. s2. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGATCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
122 292. s2. NOD2f23	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGATCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-167 o288. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
134 o288. NOD2f23	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
-165 2040. NOD2r27	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
130 2040. NOD2f23	GATGCTGTGCAAAATGTTAATTATTTTAAACATTATGATGTGTGAAAACTGGTTAATAATTTTATAGGTCACTTTGTTTTACTGCTTTAAGTTTATACTCTTATAGACAACATGGCC										
< >	CONSENSUS ---										

Base confidence:10 (Probability 0.900000) Position 4394

Genomanalyse

Eine „ultimative“ Genomstruktur: Die Nukleotidsequenz

The screenshot displays the 'Contig Editor' interface for file '1 AF1 78930.1'. The main window shows a sequence alignment with various contigs listed on the left and their corresponding nucleotide sequences on the right. The contigs listed include identifiers like 'o404.NOD2r27', '1035.2.NOD2f23', and '169.s1.NOD2r27'. The sequence alignment shows dots representing gaps or matches between the contigs. At the bottom, the 'CONSENSUS' sequence is displayed as 'TGCAAAATGTTATTATTTTAAACATTATGATGCTGCGAAAAC TGGTTAATATTTATAGCTCACITTTGTTTTACTGTCTTAAGTTTATACTCTTATAGACAACATGCCCGTGAAC T'. A status bar at the very bottom provides details for the selected contig: 'Reading:2040.NOD2r27(#165) Length:409(1394) Vector:unknown Clone:unknown Chemistry:terminator Primer:forward universal'.

Contig Editor: 1 AF1 78930.1

< C: 1 > < Q: 10 > Insert Edit Modes >> Cutoffs Undo Next Search Commands >> Settings >> Quit Help >>

30 4340 4350 4360 4370 4380 4390 4400 4410 4420 4430 4440

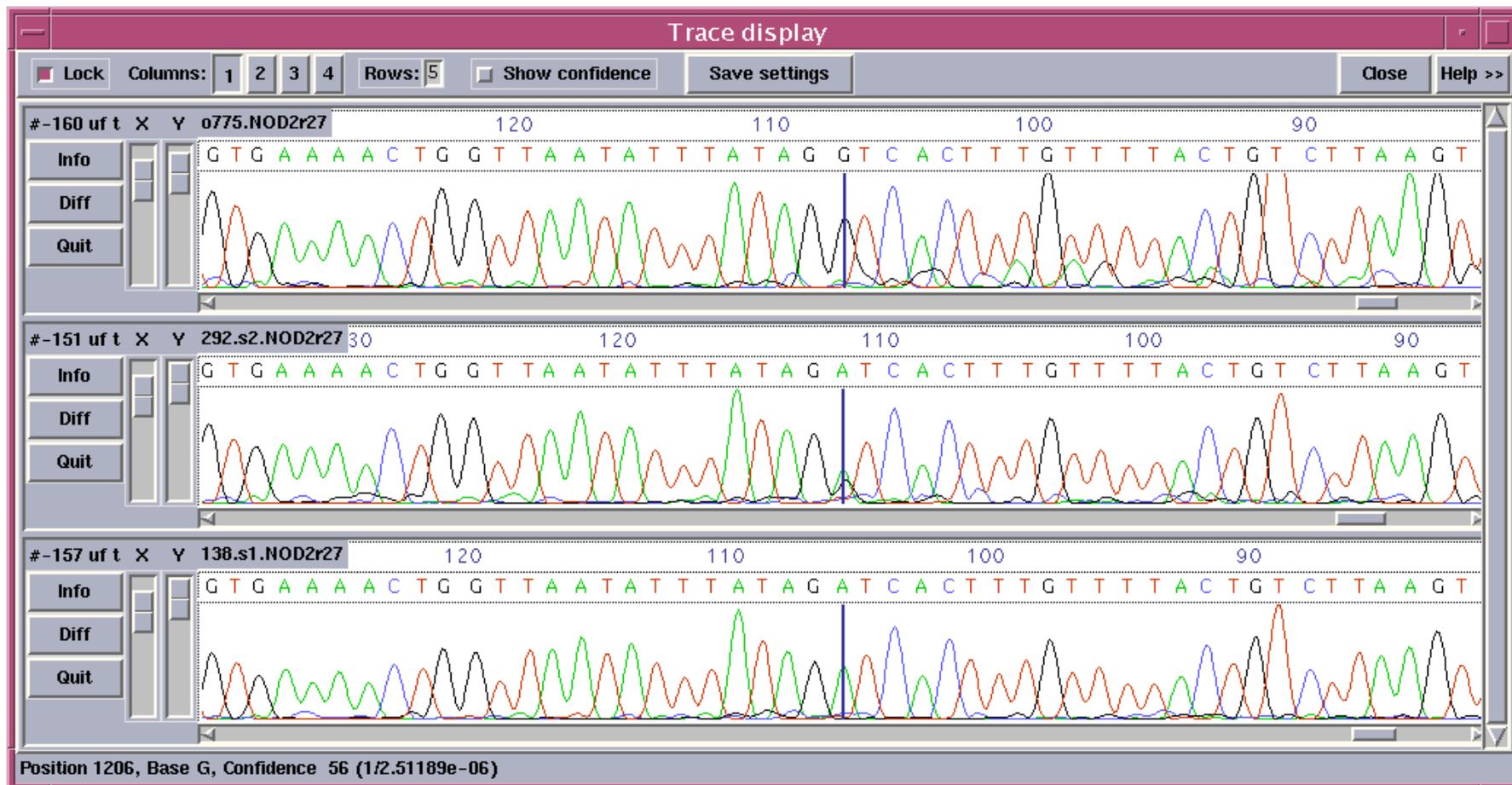
-1061 5. NO. r27
-168 o404. NOD2r27
131 o404. NOD2f23
-166 1035. 2. NOD2r27
127 1035. 2. NOD2f23
-156 169. s1. NOD2r27
-161 2103. NOD2r27
118 2103. NOD2f23
-153 o826. NOD2r27
138 o826. NOD2f23
-164 1035. 1. NOD2r27
137 1035. 1. NOD2f23
-160 o775. NOD2r27
117 o775. NOD2f23
-157 138. s1. NOD2r27
140 138. s1. NOD2f23
-152 367. s1. NOD2r27
123 367. s1. NOD2f23
-145 311. s2. NOD2r27
126 311. s2. NOD2f23
-144 o589. NOD2r27
119 o589. NOD2f23
-142 341. s1. NOD2r27
-151 292. s2. NOD2r27
122 292. s2. NOD2f23
-167 o288. NOD2r27
134 o288. NOD2f23
-165 2040. NOD2r27
130 2040. NOD2f23

< > CONSENSUS -----TGCAAAATGTTATTATTTTAAACATTATGATGCTGCGAAAAC TGGTTAATATTTATAGCTCACITTTGTTTTACTGTCTTAAGTTTATACTCTTATAGACAACATGCCCGTGAAC T

Reading:2040.NOD2r27(#165) Length:409(1394) Vector:unknown Clone:unknown Chemistry:terminator Primer:forward universal

Genomanalyse

Genomsequenzen sind individuell



Genomanalyse

Einzelnukleotidvariationen: SNPs

ATTTCGAC**C**GTATTG

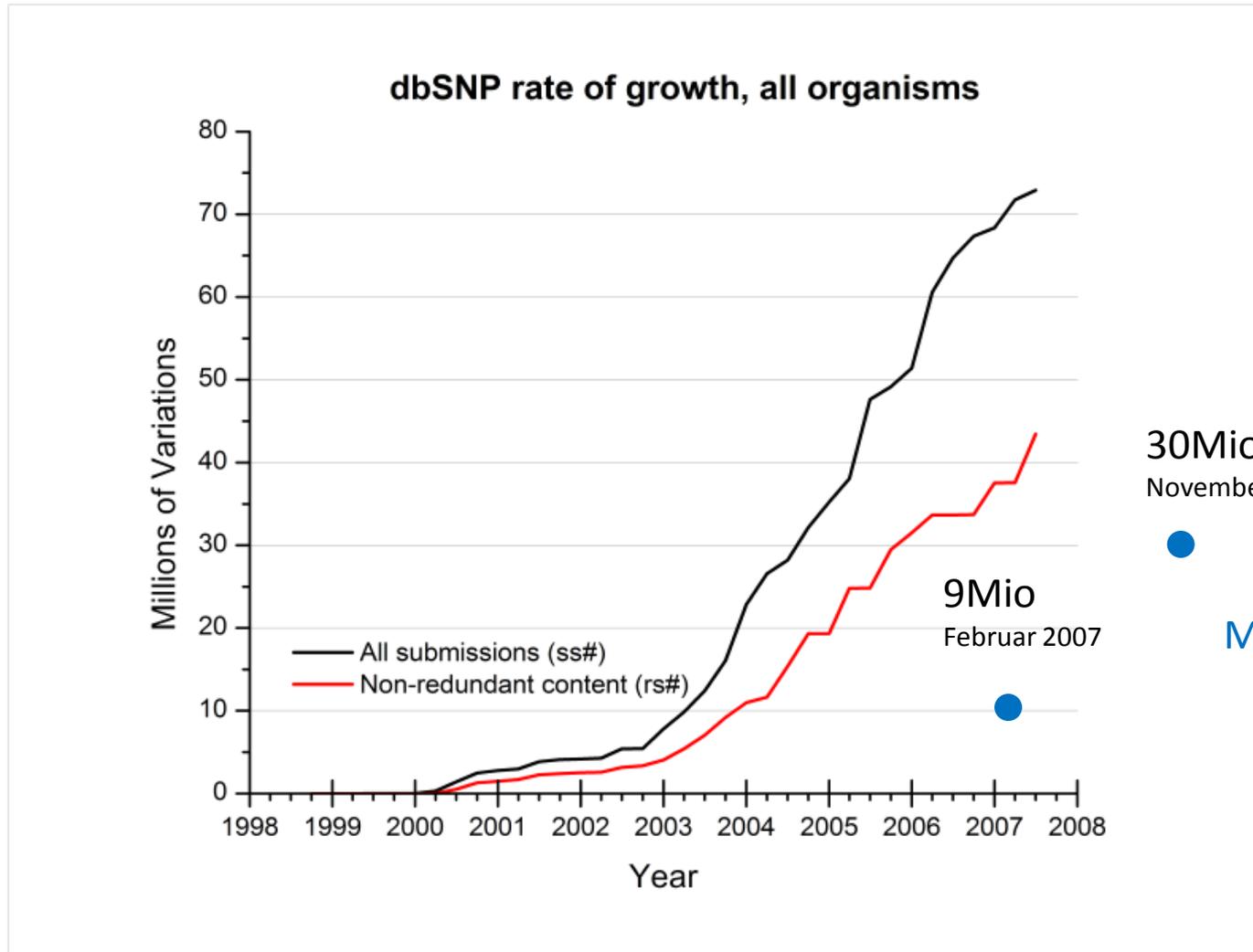
ATTTCGAT**T**GTATTG



SNP

Genomanalyse

SNP Datenbanken



Genomanalyse

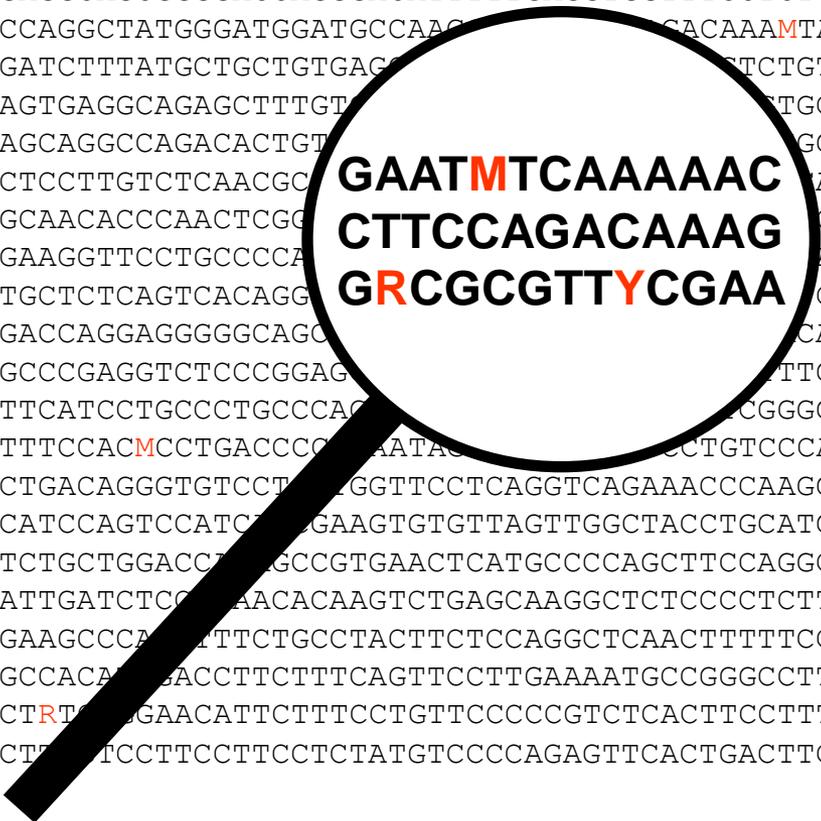
Die „ultimative“ Genomstruktur: Sequenz *und* SNP-Katalog

AAGAGCCAGTCACACGAGGACTGGCTGTGTCTACCTTGCCCCGCGCCGTGGGCATCCCCAGCACCAAGCATGGTGCCT
GGCATGCAGTGAACCTTCAGTGTACCTGACCGAGTAAACACAGGACTGAAGGAAGGGAAGGCGAGATGGATGCAAAR
GGCCCAGTGTGCCTCAGTCTTCTTTGGGAGGCAGCGGGGACCAGGGACATTTTTGAGGTGGTTTCCCTCTCCAGGGATT
AACCGAGTTTATCCCTCCTCAGTTCTCCAGGCTATGGGATGGATGCCAAGGGA^MCTTCCAGACAAAM^TTATTCTGAAG
TCTGGGCCTGTGAGAAGCATCTTGAGAGATCTTTATGCTGCTGTGAGGATTCTAGGACCTGGAGGTCTGTCAGGGCAA
TCAGGAATGCTCCCTGGAGGAAGTGGCAGTGAGGCAGAGCTTTGTGGAGGGGCAGCAGGGGAGAGAGTGTGTGGCCA
^RAGGAAGGGAAGGGCCATGGCAGGGAAGCAGGCCAGACACTGTTAGGCCATTAGGAGAGTGTGAAGGGTAATTGGA
TCTGAGTGGGAGGAGCGGCTGTGGTAACTCCTTGTCTCAACGCCTCTGACTAAAGTGGCCAGGCACACCACCCAGGCA
CCTGCCGCCCTTGCCCCAGGGGCTGCTGCAACACCCAACCTCGGCTTTAGCCAGAGGGCACTGGGCCAGGCAGATATGG
GGTCCCTTCTCCCTGAAAGACAGGGAAGGAAGGTTCCCTGCCCAACCCCAGGCTCTGCCCAGGCTCGTGGAGAAGCCTG
TCCCCGAGGCAGTGGGCGGAACCAGGATGCTCTCAGTCACAGGGCAGCCAAGCGGCTGGGGCGGCAGCTTTCATCAT
CTTCAGTTGCATTTGACAGCTGTGGCTGACCAGGAGGGGGCAGCTAGGCAGAAGTGTGCATCCCGGTCACACCAGTGTC
ACCTGAGGCTCATCAGGGCAGTGACTTGCCCAGGTCTCCCGGAGCAGCTCAGTAAGGATTGAGCATTTGTTTACTCA
TCTCCATCGCTGACAGCTGAAACTGGATTTCATCCTGCCCTGCCCAGGGGGCTGGGAATCAGGGATCGGGGCCAGCTCA
GCTGAGATGTGAGGGTCCCAAATAGTATTTCCAC^MCCTGACCCCAAATAGGCCCAAATCCCCTGTCCCAAAGACTTG
ACGTGCCCATTTGATGTGTCTGTCGTTCCCTGACAGGGTGTCTCTTGGTTCCTCAGGTCAGAAACCCAAGGGTCATCTC
TGTCCCTGGGCTCTTCCTCACTGCCACATCCAGTCCATCAGCGAAGTGTGTTAGTTGGCTACCTGCATGCGTCTCTGA
ATCCGTCTCCTTCTCCTCCTCCATCTCTCTGCTGGACCACAGCCGTGAACTCATGCCCCAGCTTCCAGGCCTTCCTGG
TCCAGTCCATCTTCCTCTCAAAGCCCATTGATCTCCAAAACACAAGTCTGAGCAAGGCTCTCCCCTCTTGAAATCCT
TCAATGCCTCCCCGTTACTCTCAGGACGAAGCCACATTTCTGCCTACTTCTCCAGGCTCAACTTTTTCTCCTCCAGCTC
CATCACCCACTATATCATTTTGTCCCAGCCACATGGACCTTCTTTCAGTTCCTTGAAAATGCCGGGCCTTTACTGATG
CCCAAGCCTTCGTACAAGCTGTTCTTCCT^RTCTGGAACATTCCTTTCCTGTTCCCCCGTCTCACTTCCCTTCACTCCTA
CTYTTATTTTAGAGTGCAGCTTGAACCCTTACTCCTTCCCTCCTCTATGTCCCCAGAGTTCAGTACTTCCCTCATCT

Genomanalyse

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

AAGAGCCAGTCACACGAGGACTGGCTGTGTCTACCTTGCCCCGCGCCGTGGGCATCCCCAGCACCAAGCATGGTGCCT
GGCATGCAGTGAACCTTCAGTGTACCTGACCGAGTAAACACAGGACTGAAGGAAGGGAAGGCGAGATGGATGCAAAR
GGCCCAGTGTGCCTCAGTCTTCTTTGGGAGGCAGCGGGGACCAGGGACATTTTTGAGGTGGTTTCTCTCCAGGGATT
AACCGAGTTTATCCCTCCTCAGTTCCTCCAGGCTATGGGATGGATGCCAAGCAGACAAAMTATTCTGAAG
TCTGGGCCTGTGAGAAGCATCTTGAGAGATCTTTATGCTGCTGTGAGGCTCTGTTCAGGGCAA
TCAGGAATGCTCCCTGGAGGAAGTGGCAGTGAGGCAGAGCTTTGTGCTGTGGCCA
RAGGAAGGGAAAGGGCCATGGCAGGGAAGCAGGCCAGACACTGTGGTAATTGGA
TCTGAGTGGGAGGAGCGGCTGTGGTAACTCCTTGTCTCAACGCCTACCCAGGCA
CCTGCCGCCCTTGCCCCAGGGGCTGCTGCAACACCCAACCTCGGCAGATATGG
GGTCCCTTCTCCCTGAAAGACAGGGAAGGAAGGTTCCCTGCCCAAGAAGCCTG
TCCCCGAGGCAGTGGGCGGAACCAGGATGCTCTCAGTCACAGGCTTTCATCAT
CTTCAGTTGCATTTGACAGCTGTGGCTGACCAGGAGGGGGCAGCCACCAGTGTC
ACCTGAGGCTCATCAGGGCAGTGACTTGCCCCGAGGTCTCCCGGAGTTGTTTACTCA
TCTCCATCGCTGACAGCTGAAACTGGATTTCATCCTGCCCTGCCCAACCGGGCCAGCTCA
GCTGAGATGTGAGGGTCCCAAATAGTATTTCCACMCCTGACCCCAATACCTGTCCCAAAGACTTG
ACGTGCCCATTTGATGTGTCTGTCTGTTCCCTGACAGGGTGTCTTGGTTCCTCAGGTCAGAAACCAAGGGTCATCTC
TGTCCCTGGGCTCTTCCCTCACTGCCACATCCAGTCCATCGAAGTGTGTTAGTTGGCTACCTGCATGCGTCTGA
ATCCGTCTCCTTCTCCTCCTCCATCTCTCTGCTGGACCTGCGGTGAACTCATGCCCCAGCTTCCAGGCCTTCTGG
TCCAGTCCATCTTCTCTCAAAGCCCATTGATCTCCCAACACAAGTCTGAGCAAGGCTCTCCCCTCTTCAAATCCT
TCAATGCCTCCCCGTTACTCTCAGGACGAAGCCCACTTCTGCCTACTTCTCCAGGCTCAACTTTTTCTCCAGCTC
CATCACCCACTATATCATTTTGTCCAGCCACACACCTTCTTTCAGTTCCTTGAAAATGCCGGGCCTTTACTGATG
CCCAAGCCTTCGTACAAGCTGTTCTTCCCTRTGGAACATTTCTTCCCTGTTCCCCCGTCTCACTTCCCTTCACTCCTA
CTYTTATTTTAGAGTGCAGCTTGAACCCTTCTTCCCTTCTCTATGTCCCCAGAGTTCAGTACTTCCCTCATCT



GAATMTCAAAAAC
CTTCCAGACAAAG
GRCGCGTTYCGAA

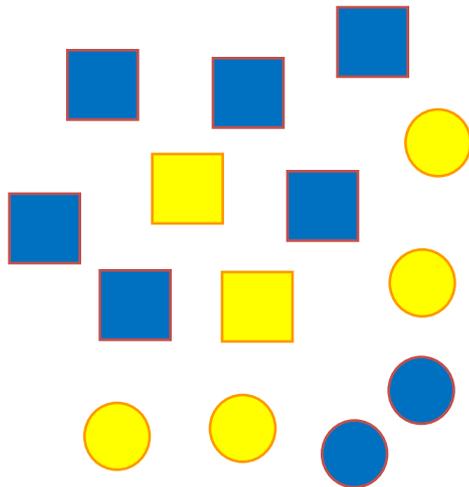
IUB code: M = aMino (A/C); R = puRin (G/A); Y = pYrimidin (C/T); ...

Genomanalyse

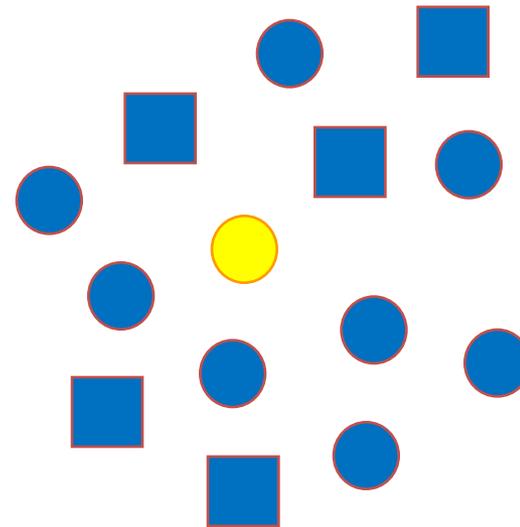
SNPs: wertvolle Marker im Genom

Typisierung von SNPs in Assoziations- und Kopplungsstudien

Patienten



Kontrollen



Genomanalyse

SNPs: wertvolle Marker im Genom



Journal home > Archive > Article > Full text > Figure 2

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

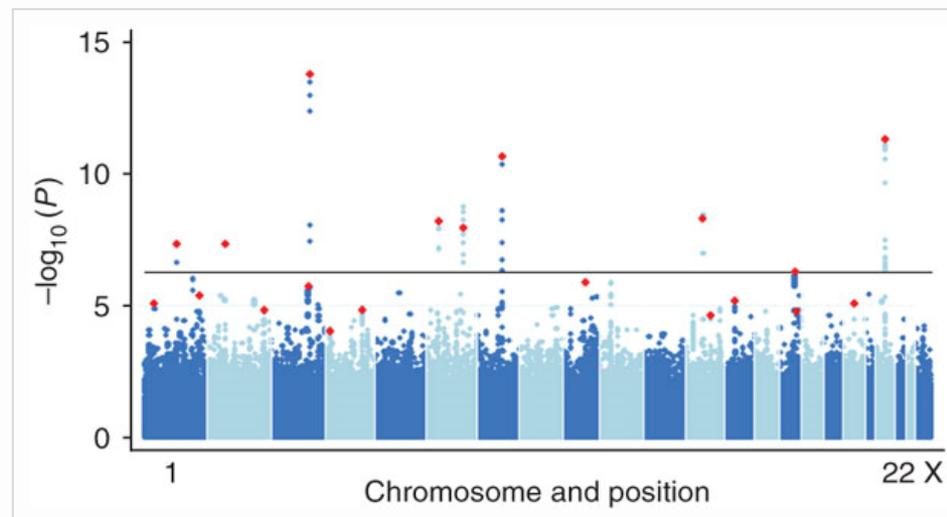
[Genome-wide association analysis identifies 20 loci that influence adult height](#)

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

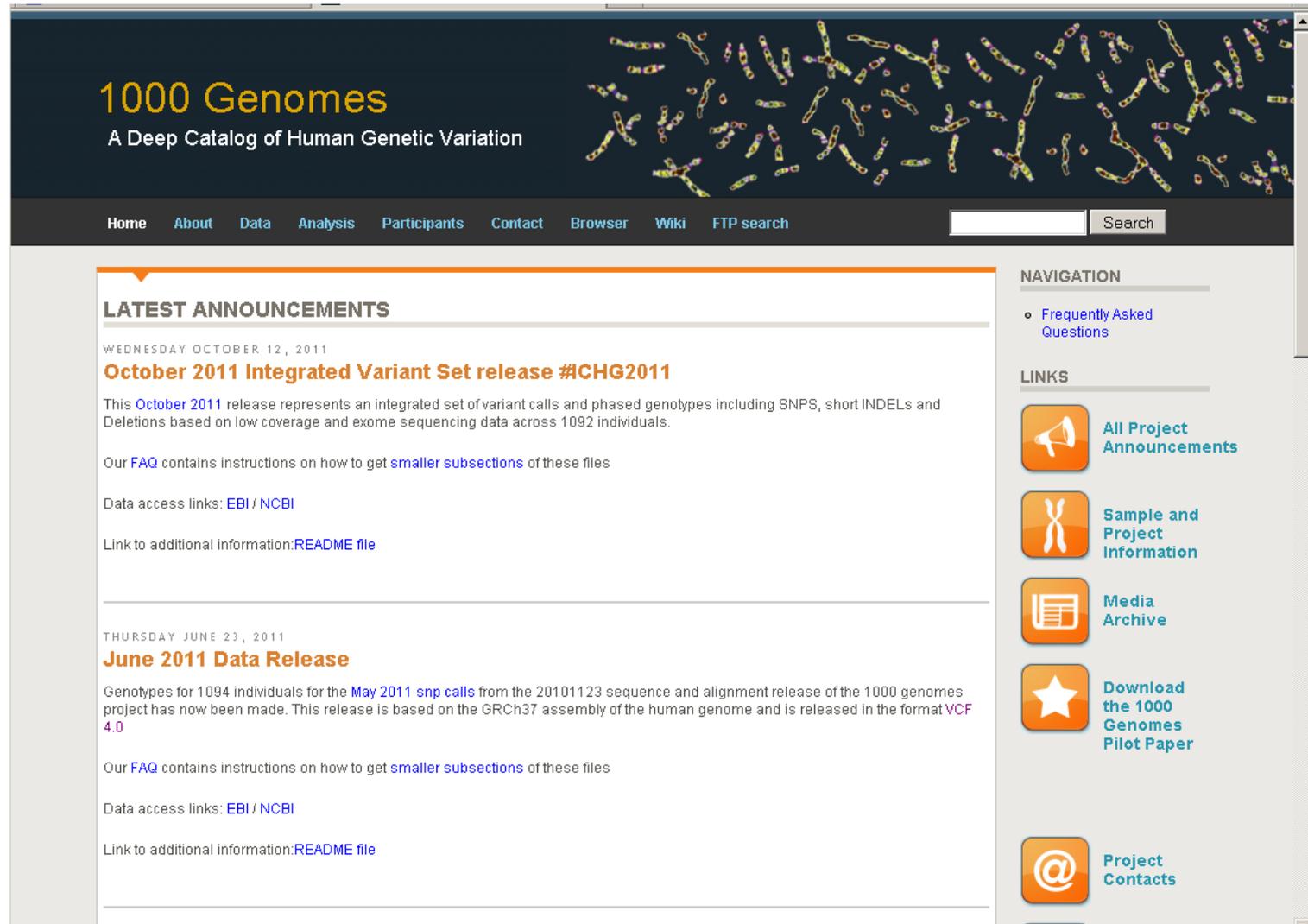
[back to article](#)



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/>



1000 Genomes
A Deep Catalog of Human Genetic Variation

Home About Data Analysis Participants Contact Browser Wiki FTP search Search

LATEST ANNOUNCEMENTS

WEDNESDAY OCTOBER 12, 2011

October 2011 Integrated Variant Set release #CHG2011

This [October 2011](#) release represents an integrated set of variant calls and phased genotypes including SNPs, short INDELs and Deletions based on low coverage and exome sequencing data across 1092 individuals.

Our [FAQ](#) contains instructions on how to get [smaller subsections](#) of these files

Data access links: [EBI / NCBI](#)

Link to additional information:[README file](#)

THURSDAY JUNE 23, 2011

June 2011 Data Release

Genotypes for 1094 individuals for the [May 2011 snp calls](#) from the 20101123 sequence and alignment release of the 1000 genomes project has now been made. This release is based on the GRCh37 assembly of the human genome and is released in the format [VCF 4.0](#)

Our [FAQ](#) contains instructions on how to get [smaller subsections](#) of these files

Data access links: [EBI / NCBI](#)

Link to additional information:[README file](#)

NAVIGATION

- [Frequently Asked Questions](#)

LINKS

-  [All Project Announcements](#)
-  [Sample and Project Information](#)
-  [Media Archive](#)
-  [Download the 1000 Genomes Pilot Paper](#)
-  [Project Contacts](#)

Genomanalyse

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

GENOME 10K®

Database & Species Lists News Events Publications Participants For G10K Organizers (restricted)

Search: Go

GENOME 10K®

Unveiling animal diversity

Genome 10K Project

To understand how complex animal life evolved through changes in DNA and use this knowledge to become better stewards of the planet.

The Genome 10K project aims to assemble a genomic zoo—a collection of DNA sequences representing the genomes of 10,000 vertebrate species, approximately one for every vertebrate genus. The trajectory of cost reduction in DNA sequencing suggests that this project will be feasible within a few years. Capturing the genetic diversity of vertebrate species would create an unprecedented resource for the life sciences and for worldwide conservation efforts.

The growing Genome 10K Community of Scientists (G10KCOS), made up of leading scientists representing major zoos, museums, research centers, and universities around the world, is dedicated to coordinating efforts in tissue specimen collection that will lay the groundwork for a large-scale sequencing and analysis project.

Accomplishments

- ▶ Inspired partly by the Genome 10K project, the i5K initiative to sequence 5,000 insect genomes began in March 2011.

Join us

Become a G10K affiliate

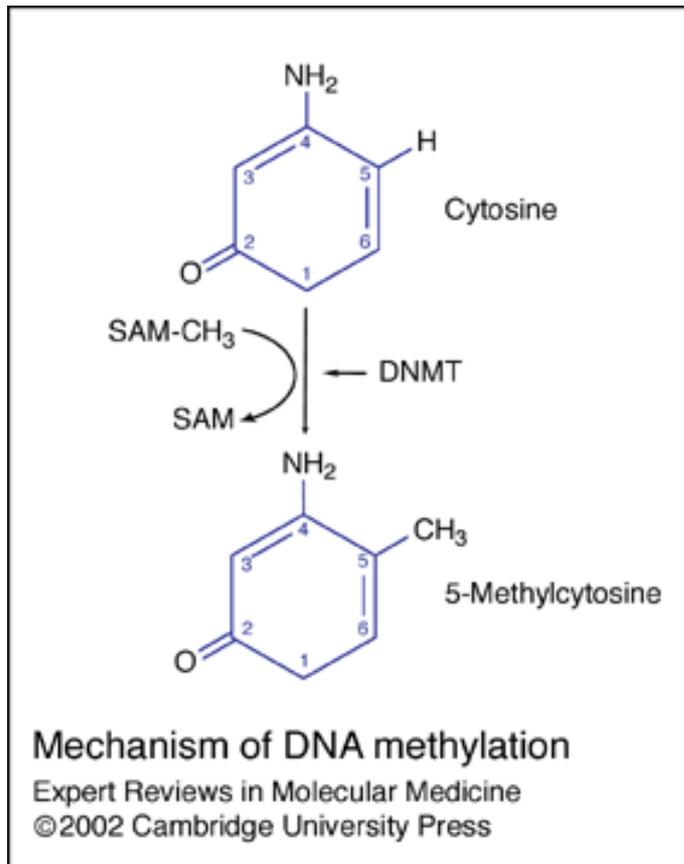
Co-directors

David Haussler,
Howard Hughes Medical
Institute Investigator
Professor of Biomolecular
Engineering
UC Santa Cruz

Oliver A. Ryder
Director of Genetics
Kleberg Chair
San Diego Zoo Institute for
Conservation Research

Genomanalyse

Sequenz ist nicht alles: Die fünfte Base



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. Tumorsuppressorgene betrifft. Genau dies passiert in der Cancerogenese.

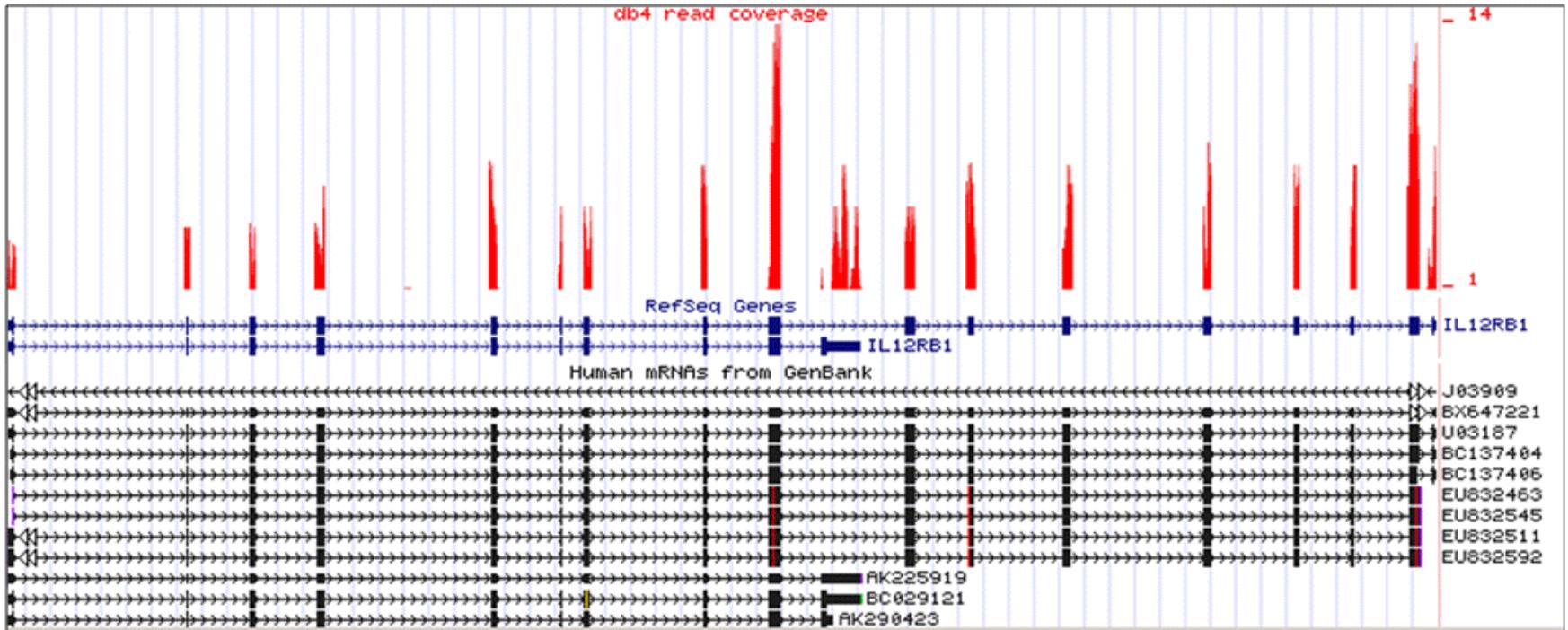
Die Kenntnis des Methylierungsmusters eines Genoms (des Methyloms bzw. Epigenoms) wird immer wichtiger!

Genomanalyse

Was nun?

Nach der Strukturanalyse die Funktionsanalyse

Genstrukturen – Genexpression - Genregulation



Vielen Dank!