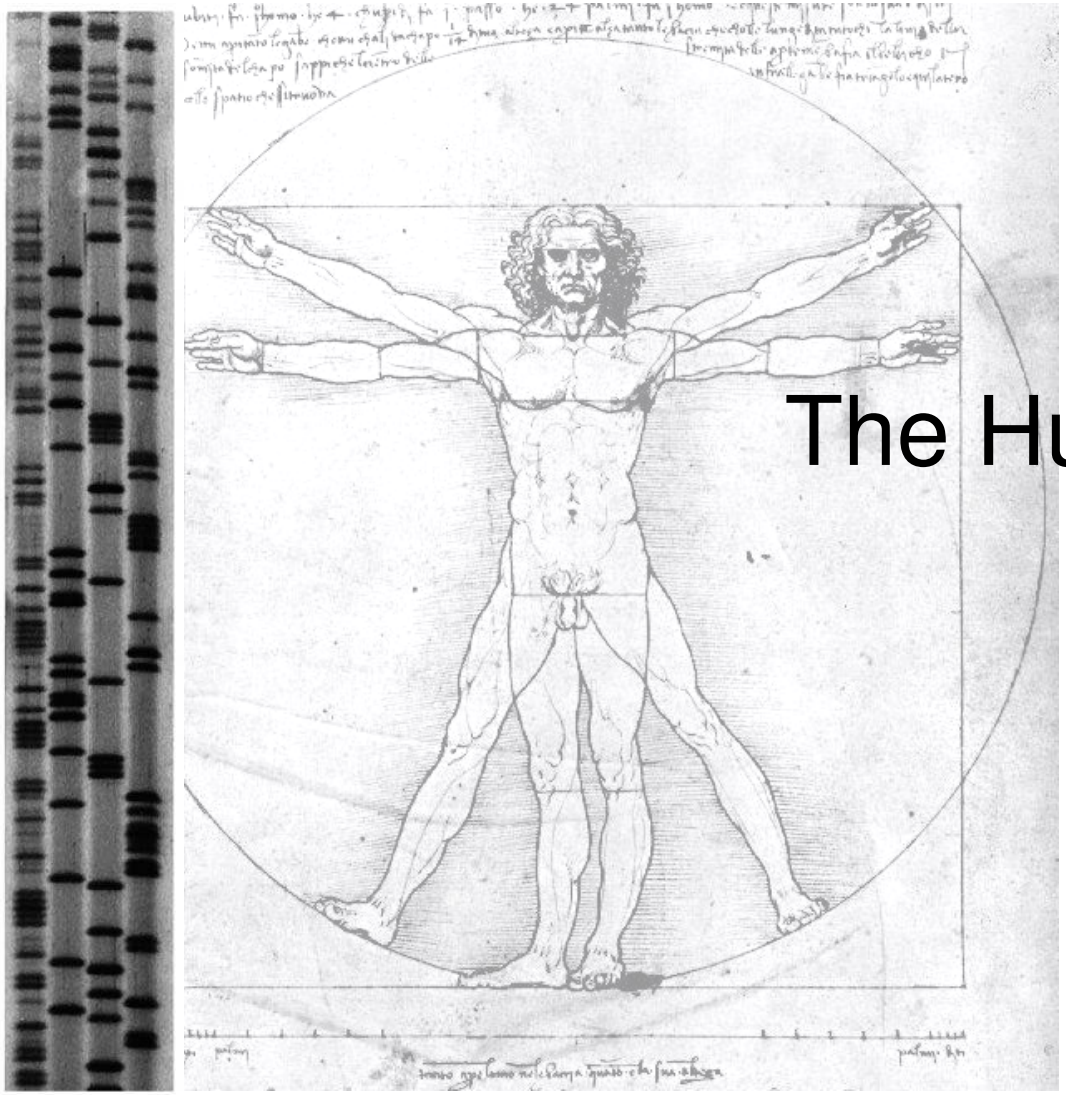


The Human Genome
and its upcoming Dynamics

Matthias Platzer

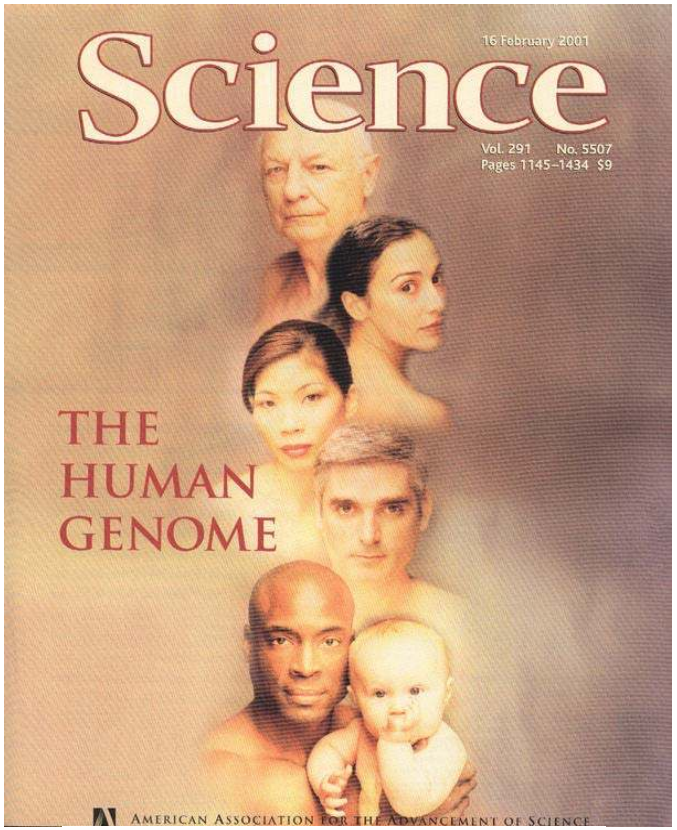
Genome Analysis
Leibniz Institute for Age Research
- Fritz-Lipmann Institute (FLI)

Sequencing of the Human Genome

Publications

2004

2001

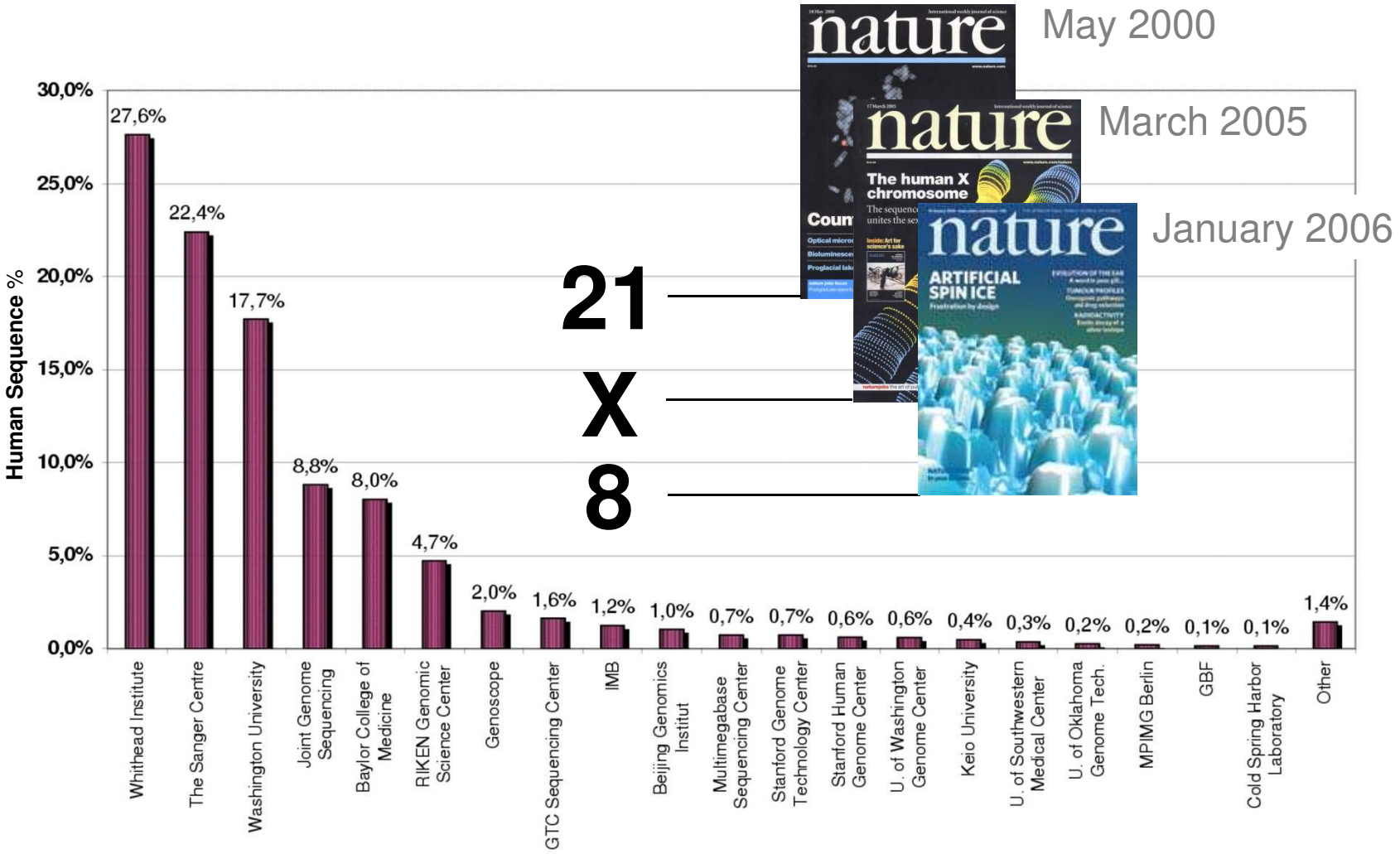


2001

International academic consortium

International Human Genome Project

Contributions of the Members



IHGSC (2001), Nature 409: 860

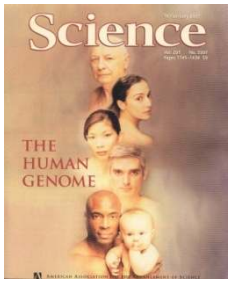
Human Genome

Working Draft versions February 2001

Academic

Initial Sequencing & Analysis…

Private



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%**

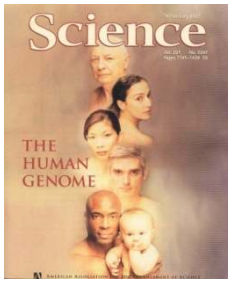
human population heterogeneity: **1 SNP/kb**
variation between two individuals: **1 SNP/10kb**

Human Genome

Final version October 2004

Academic

Private



Initial Seq… Finishing the euchromatic sequence… 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**

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%**

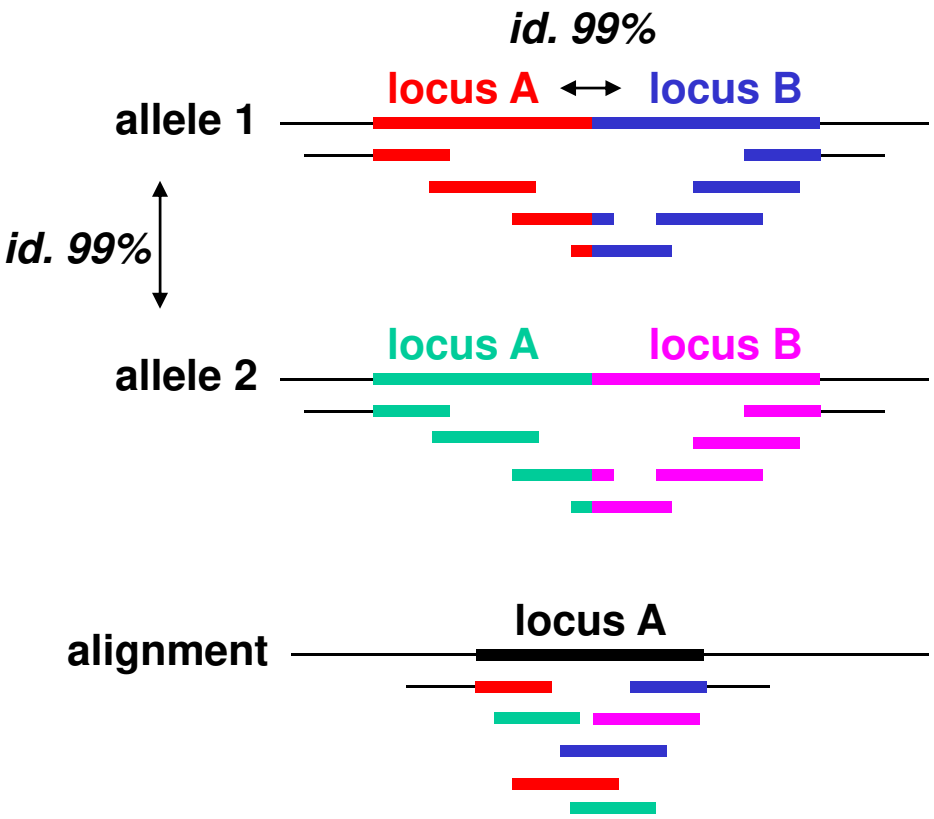
Human genome

5.3% segmentally duplicated

87% of all segmental duplications >**50 kb**

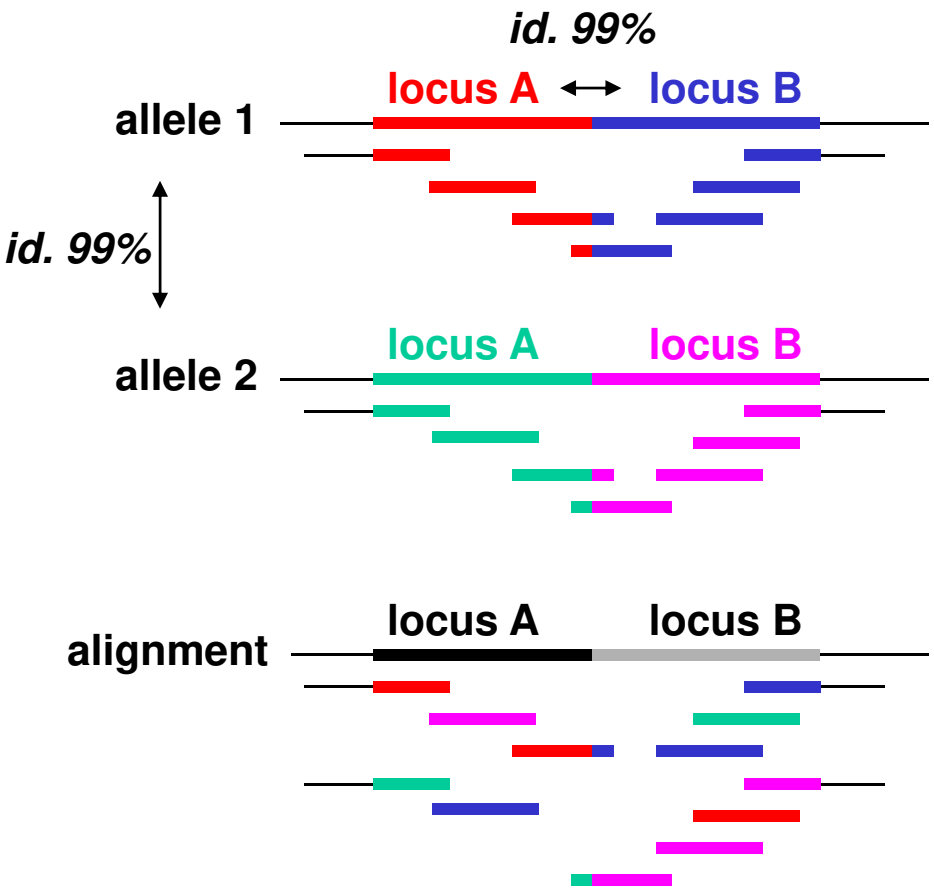
Segmental Tandem Duplications

Assembly problems



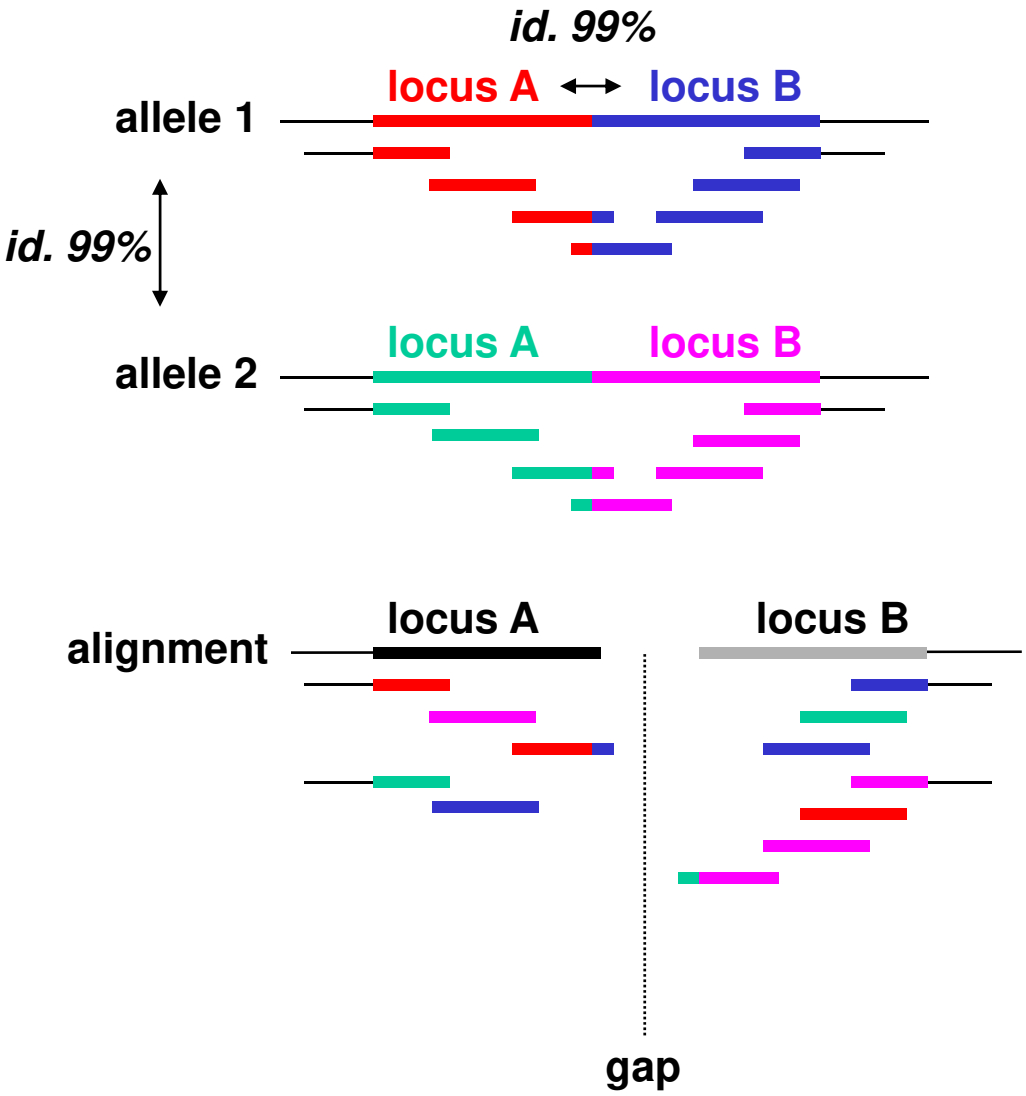
Segmental Tandem Duplications

Assembly problems



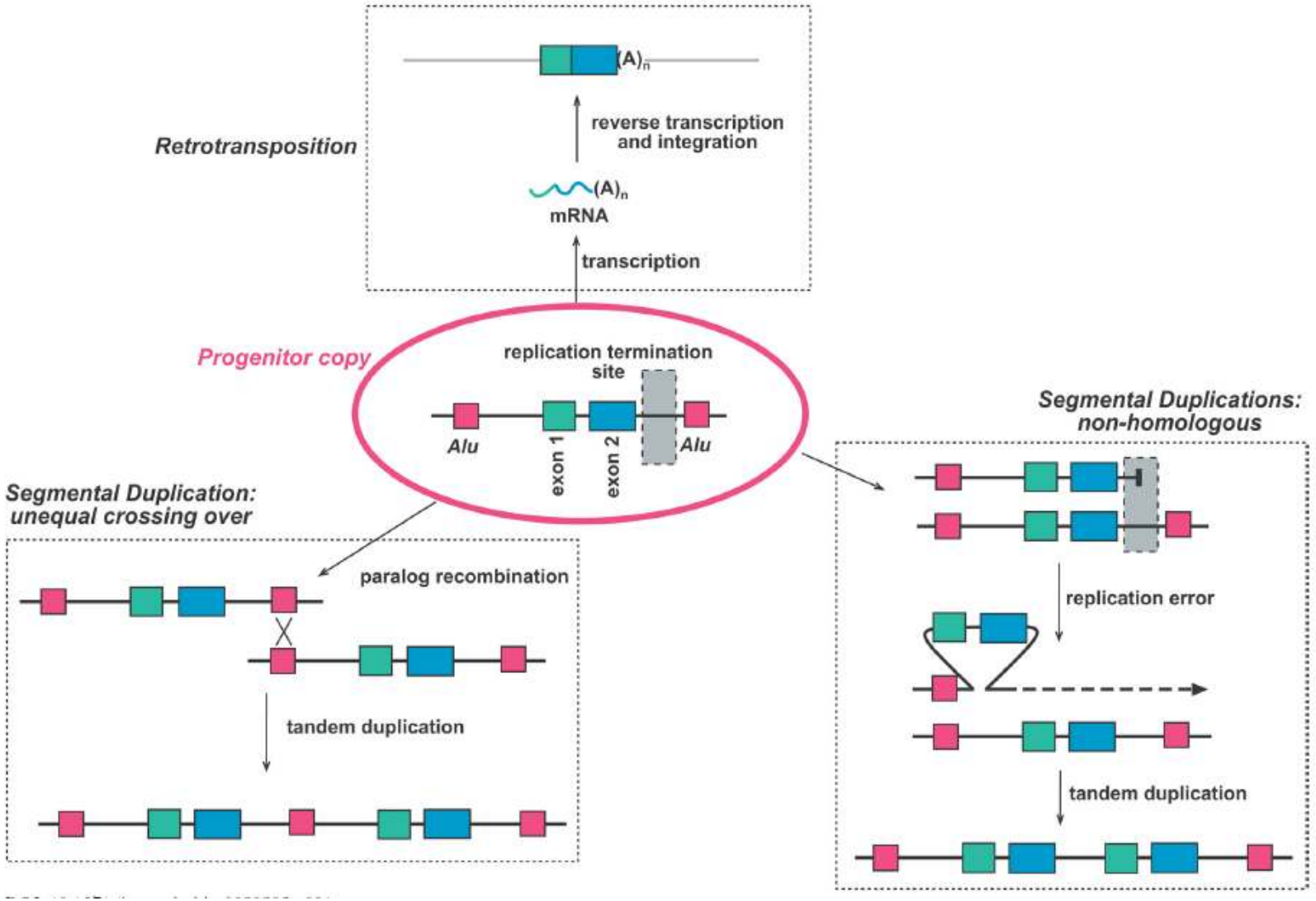
Segmental Tandem Duplications

Assembly problems



Segmental duplications

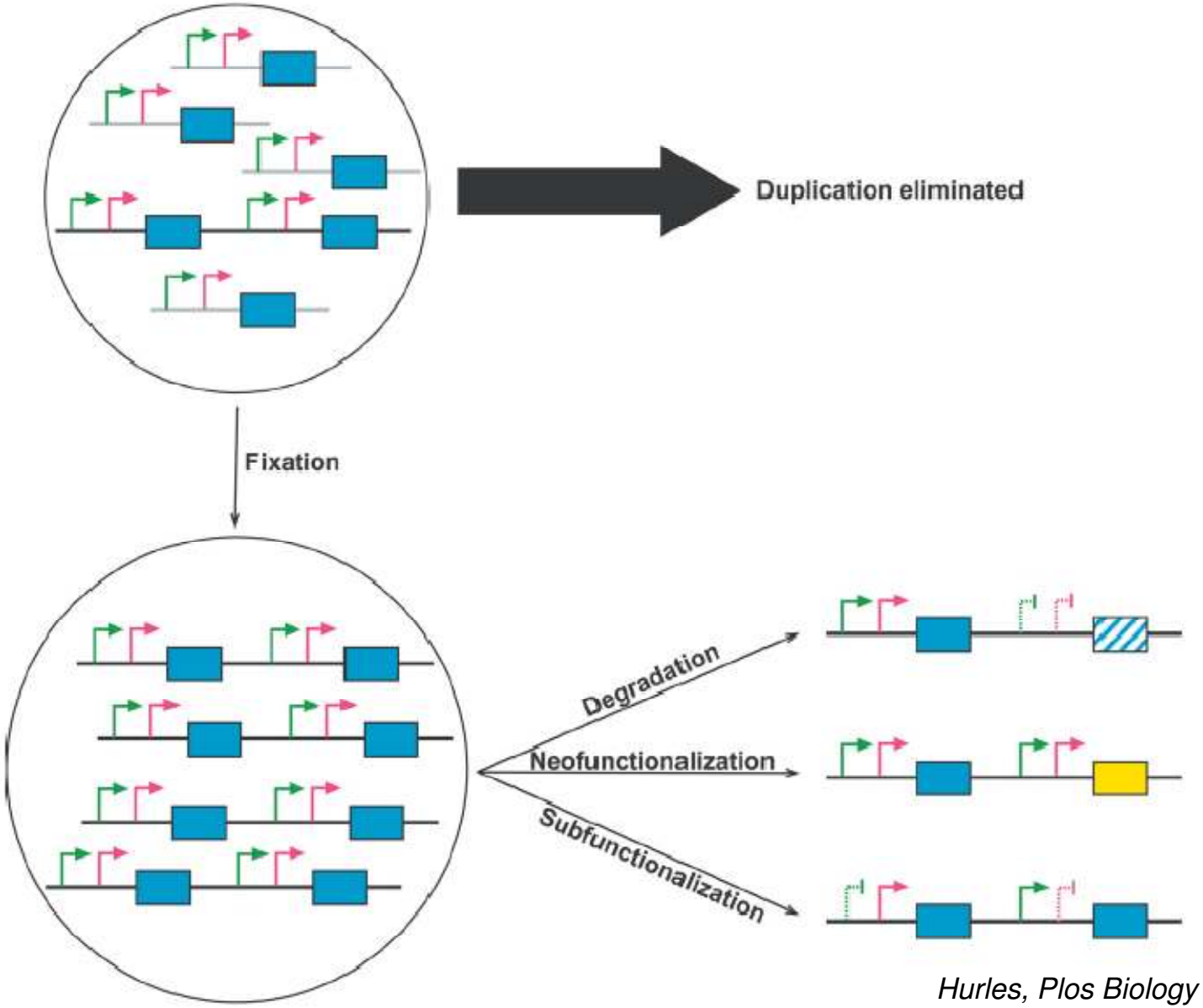
Mechanisms



Hurles, Plos Biology 2:900 (2004)

Segmental duplications

Fate of duplicated genes



Hurles, Plos Biology 2:900 (2004)

Structural Variations / Polymorphisms

Types

255 large-scale **copy number variations** (>100 kb).

Nat Genet 36:949 (2004)

Intermediate size variations >8 kb: 139 insertions, 102 deletions, 56 inversion break points

Nat Genet 37:727 (2005)

76 large-scale **copy number polymorphisms**; on average of 11 variations between two individuals with a median length of 222 kb.

Science 305:525 (2004)

586 **deletion polymorphisms** covering 276 genes; typical individual is hemizygous for 30-50 deletions >5 kb, totaling 550-750 kb.

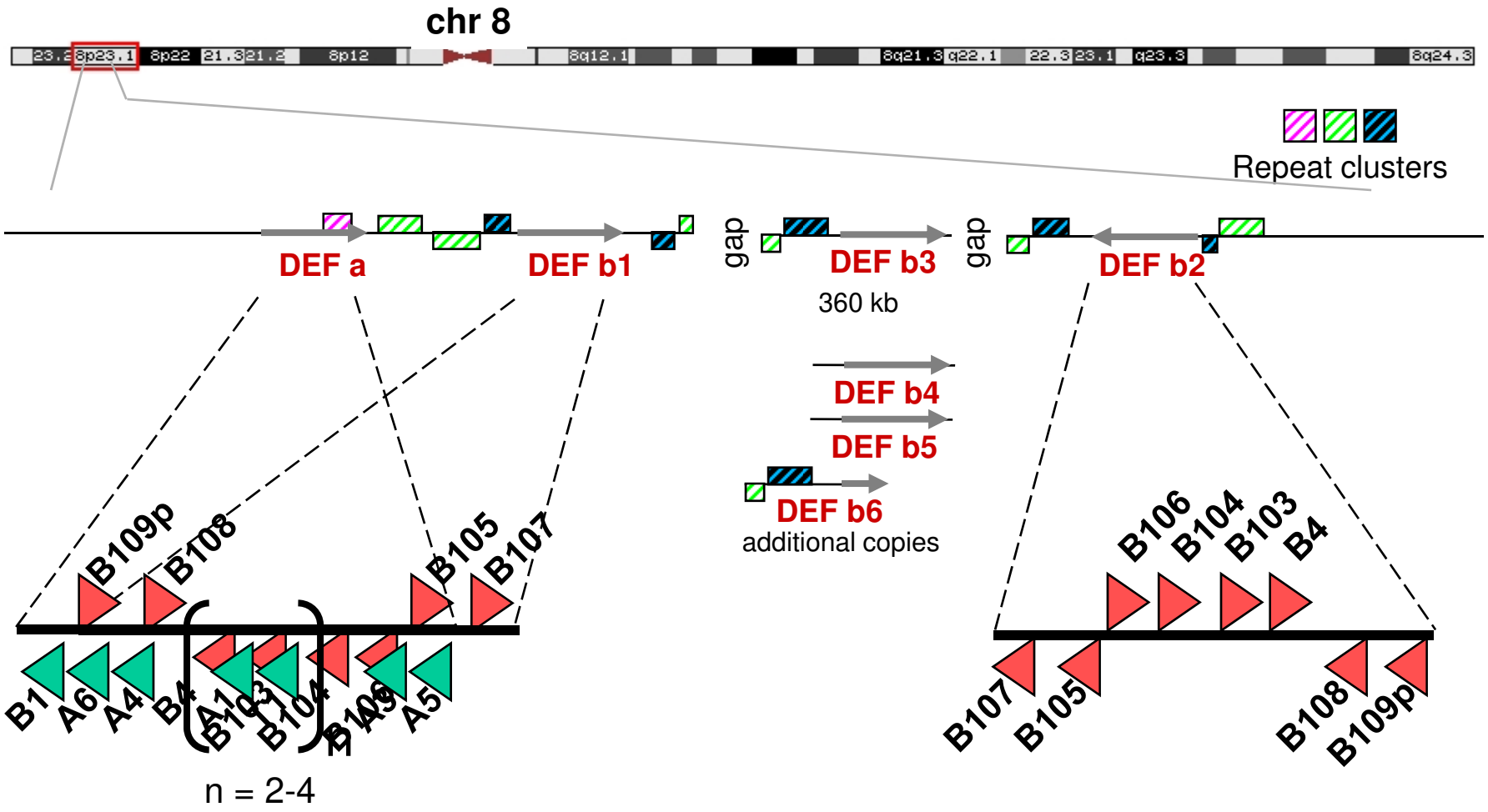
Nat Genet 38:75 (2006)

541 **deletion polymorphisms** covering 1-745 kb; 278 in multiple, unrelated individuals; 120 in homozygous state.

Nat Genet 38:86 (2006)

DEF cluster at 8p23.1

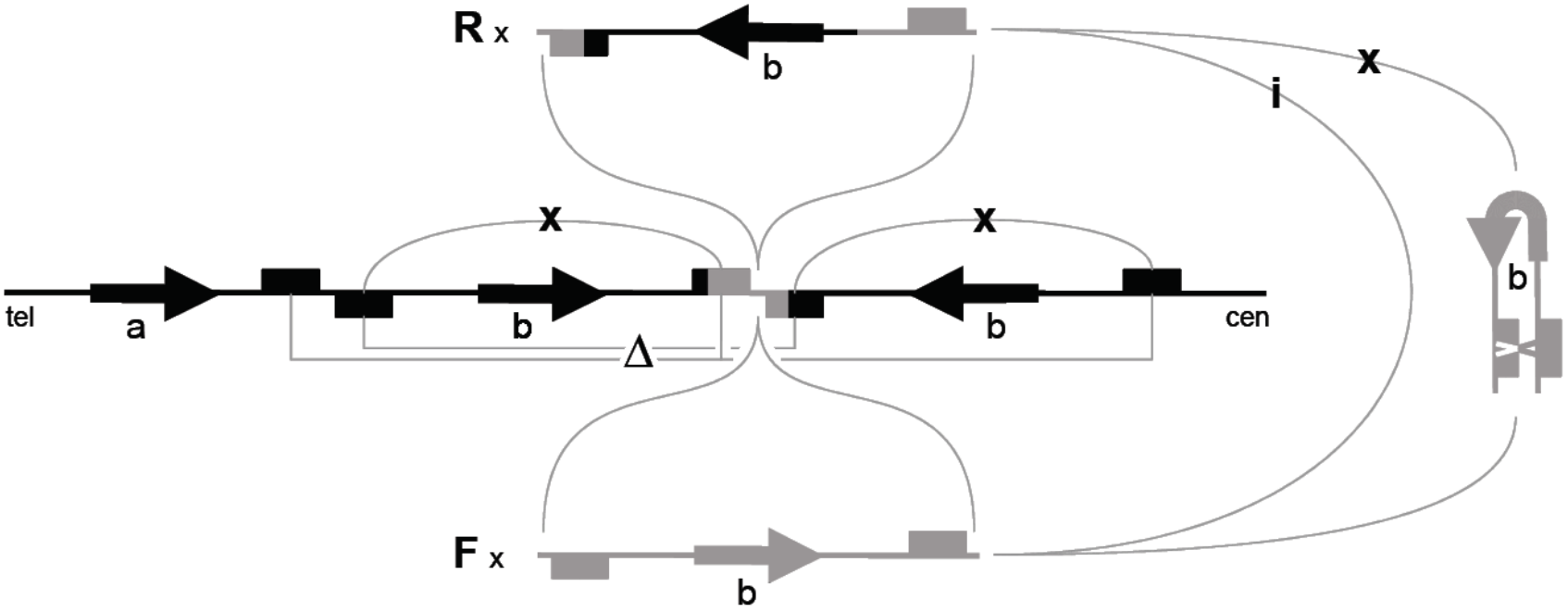
hg16: 6.3-8.3 Mb



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

Genomic variability of 8p23.1 DEF locus

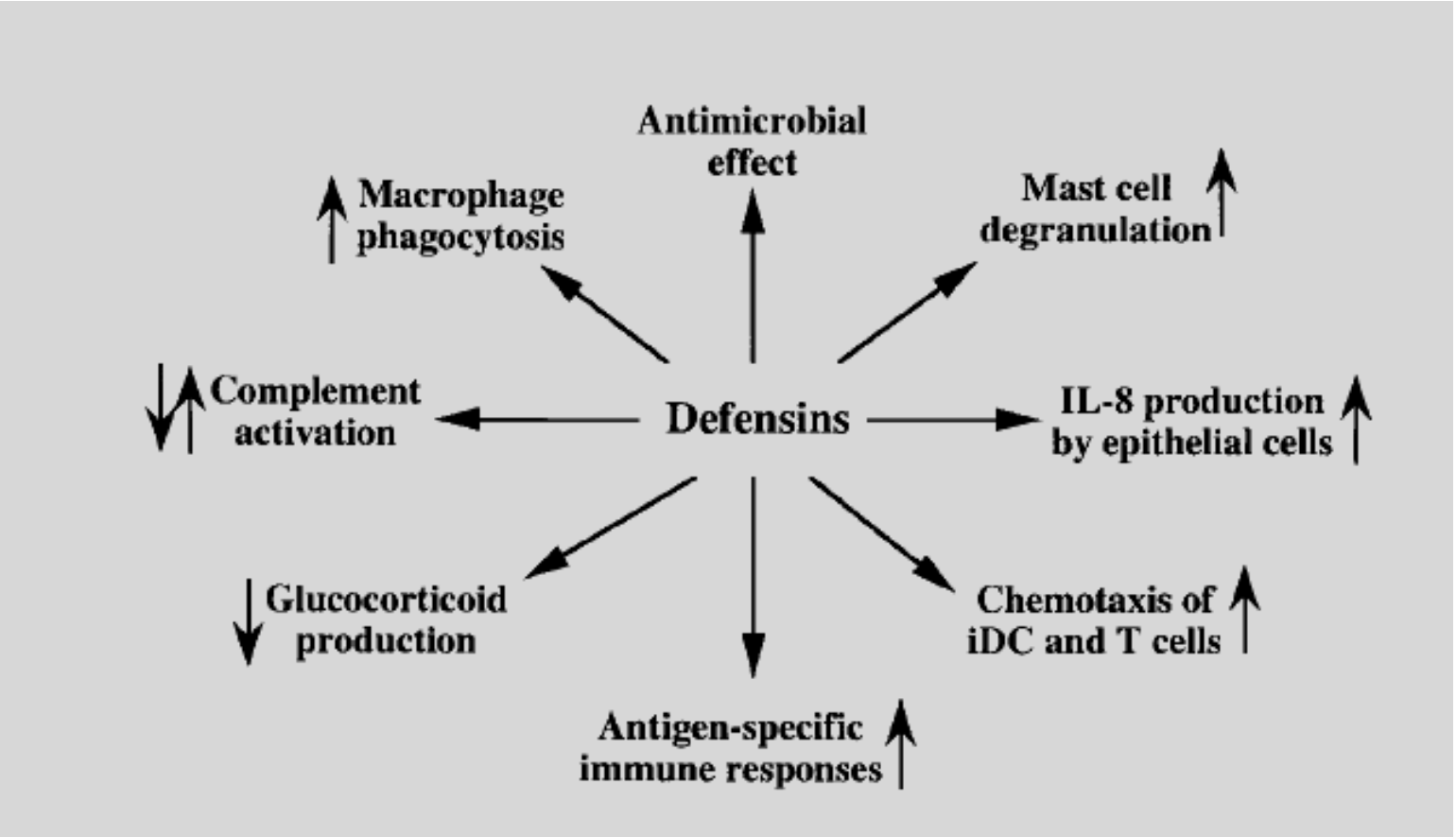
Hypothetical organisation



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

Defensins (DEF)

Multiple roles



Immunity & Cancer

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

Complex phenotypes / diseases

Structural variations

Gene	Type	Locus	Size (kb)	Phenotype	Copy number
UGT2B17	Deletion	4q13	150	Variable testosterone levels, risk of prostate cancer	0–2
DEFB4	VNTR	8p23.1	20	Colonic Crohn's disease	2–10
FCGR3	Deletion	1q23.3	>5	Glomerulonephritis, systemic lupus erythematosus	0–14
OPN1LW/OPN1MW	VNTR	Xq28	13-15	Red/green colour blindness	0–4/0–7
LPA	VNTR	6q25.3	5.5	Altered coronary heart disease risk	2–38
CCL3L1/CCL4L1	VNTR	17q12	Not known*	Reduced HIV infection; reduced AIDS susceptibilty	0–14
RHD	Deletion	1p36.11	60	Rhesus blood group sensitivity	0–2
CYP2A6	Deletion	19q13.2	7	Altered nicotine metabolism	2–3

*Precise boundaries of the copy-number variant are not known.
VNTR, variable number tandem repeats.

Eichler et al., Nature 447:161 (2007)

Complex phenotypes / diseases

Structural variations

FCGR3 copy number & glomerulonephritis in humans and rats

Nature 439:851 (2006)

Strong association of *de novo* copy number mutations with autism

Science 316: 445 (2007)

Segmental duplications

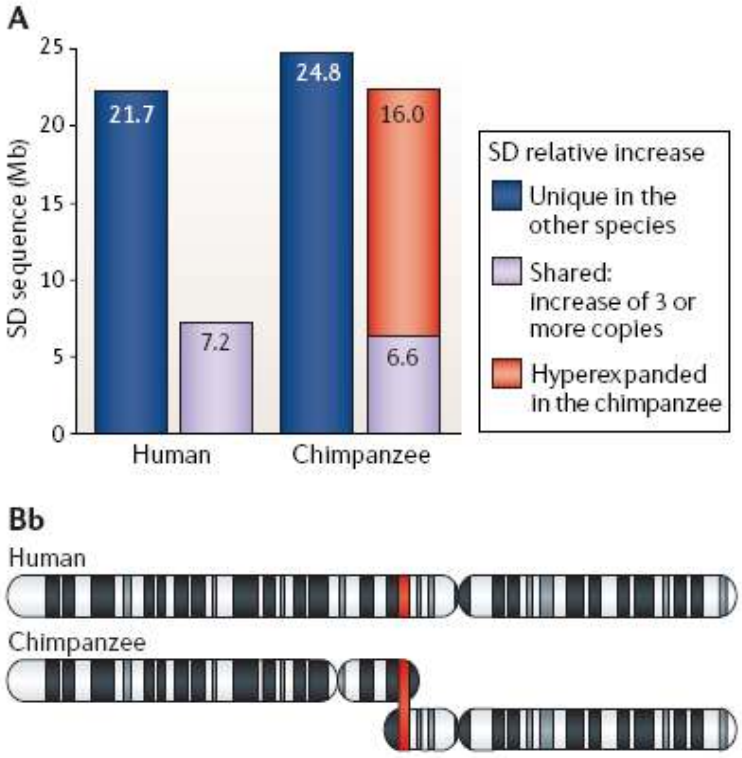
Content of sequenced animal genomes

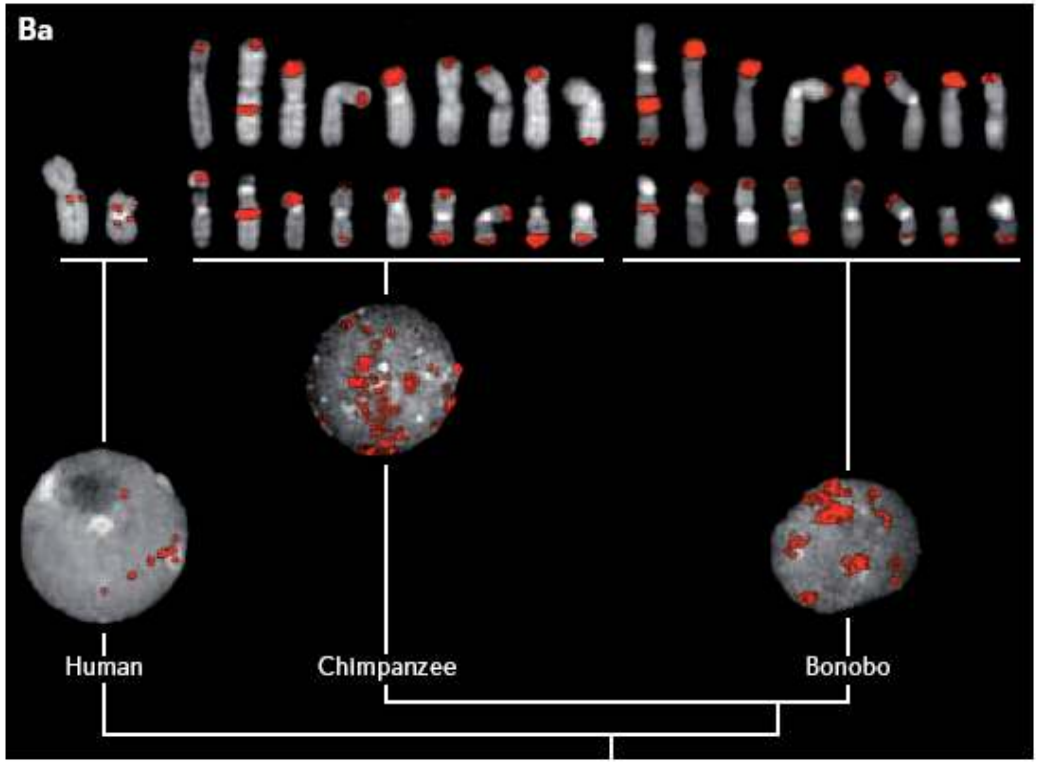
	Caenorhabditis elegans	*Drosophila melanogaster*	Human	Mouse	Rat	Chicken	Chimpanzee*
SDs of >1 kb	4.3%	1.2%	5.2%	2.7%	1.6%	2.7%	N.D.
SDs of >10 kb	0.7%	0.1%	4.5%	2.2%	1.5%	0.3%	N.D.
SDs of >20 kb	N.D.	N.D.	4.0%	1.7%	0.9%	0.0%	~4.8%
Genome size	97	123	2,866	2,506	2,566	1,040	2,866

Bailey & Eichler, Nat Rev Genet 7:552 (2006)

Segmental duplication content of hominoids

Hyperexpansions in chimpanzee

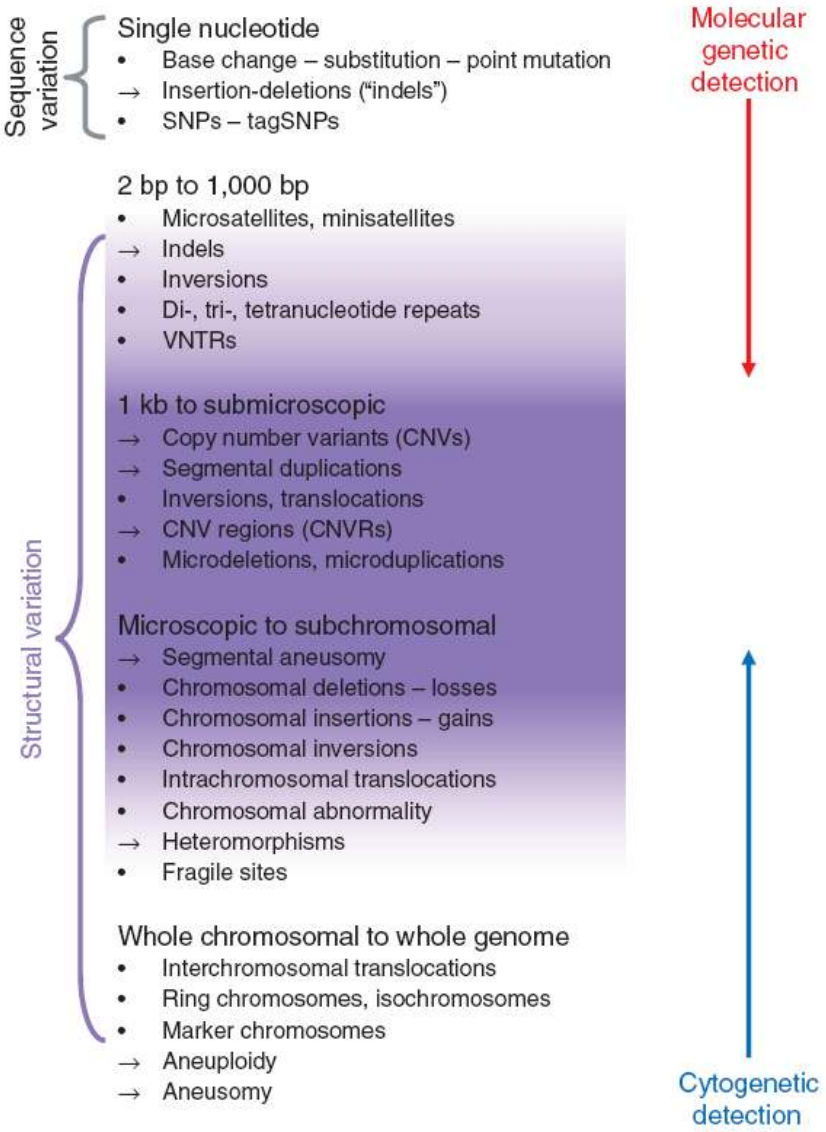




Bailey & Eichler, Nat Rev Genet 7:552 (2006)

Genomic variations

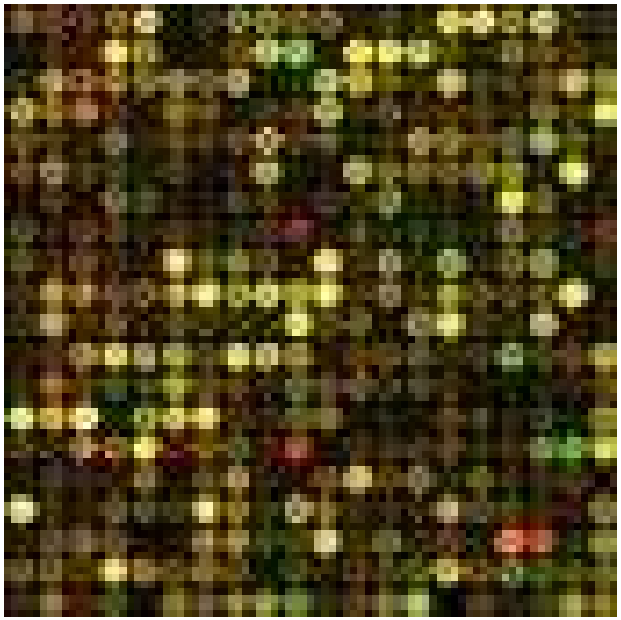
Lexicon

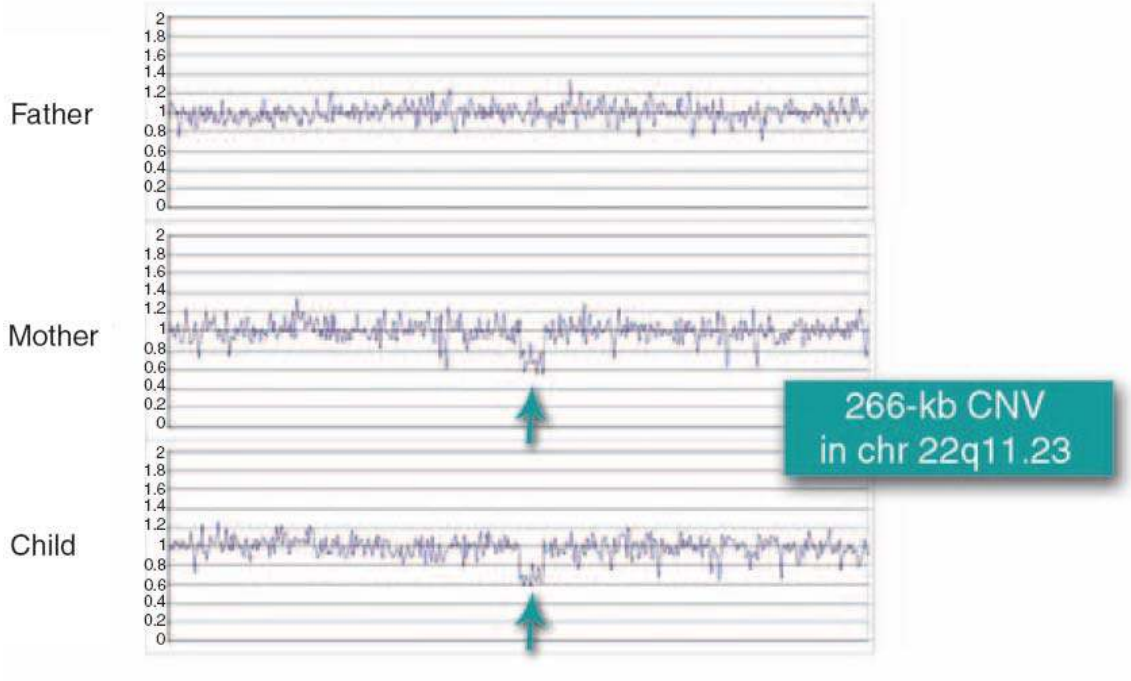


Scherer et al., Nat Genet Suppl 39:s7 (2007)

Copy number variation (CNV)

Detection by DNA microarrays



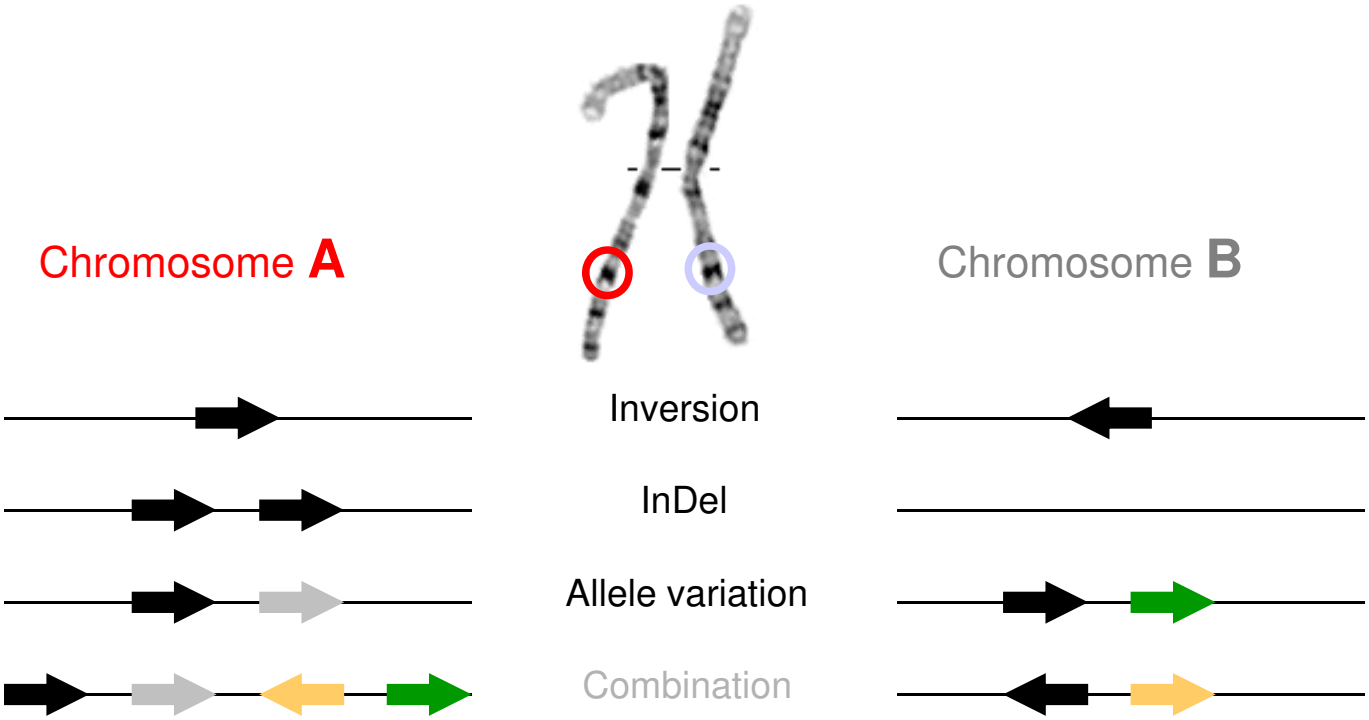


- 0.5-2 Mio data points
- comparative hybridization vs. a reference

Lee et al., Nat Genet Suppl 39:s48 (2007)

Genetic Variability

Structural variations



Conclusions

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)

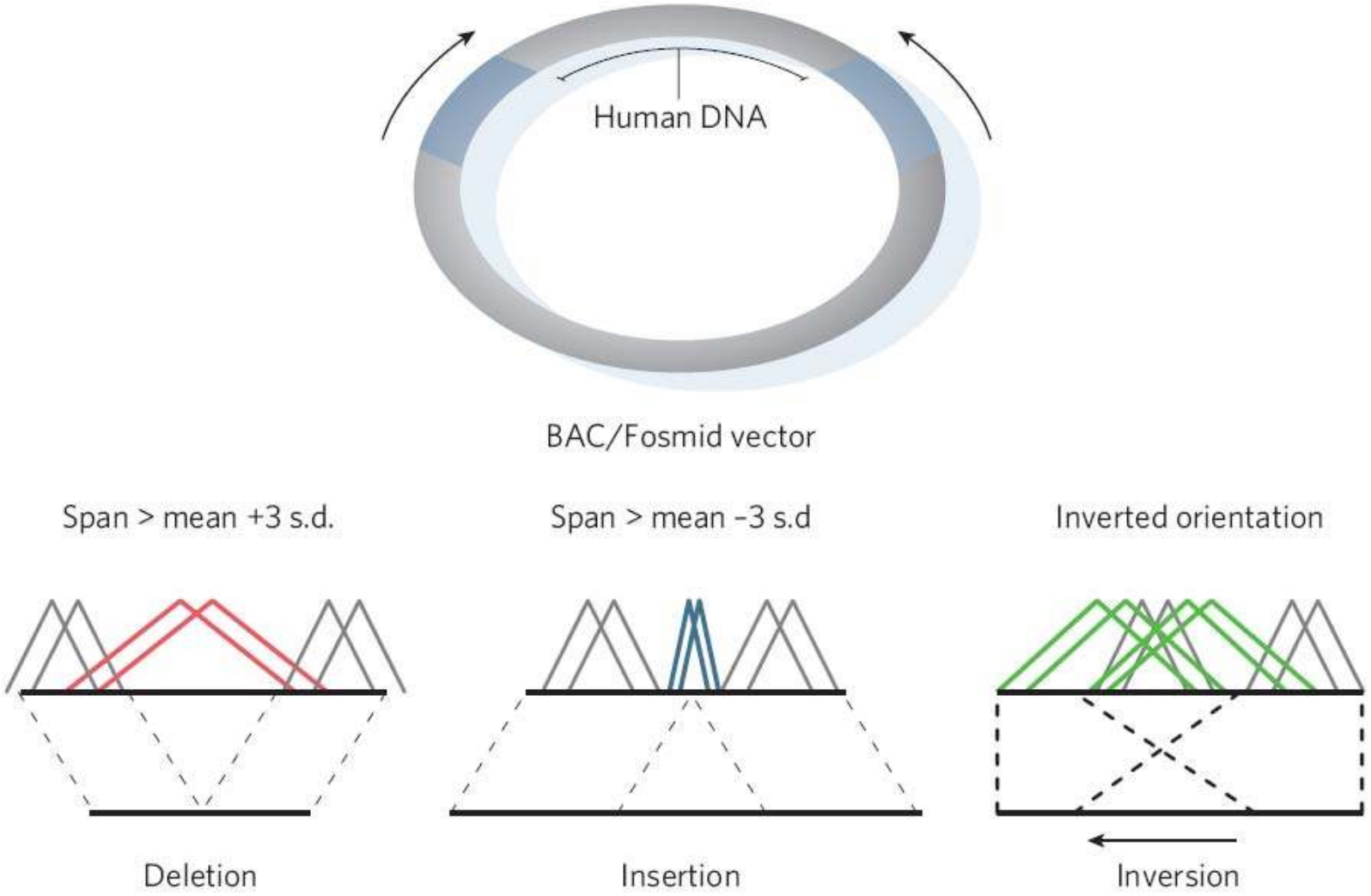
Conclusions

Copy number polymorphism at orthologous regions of divers genomes … suggests that genome plasticity… is a more common cause of genetically complex phenotypes than has hitherto been observed.

Aitman et al., Nature 439:851 (2006)

Completing the map of human genetic variation

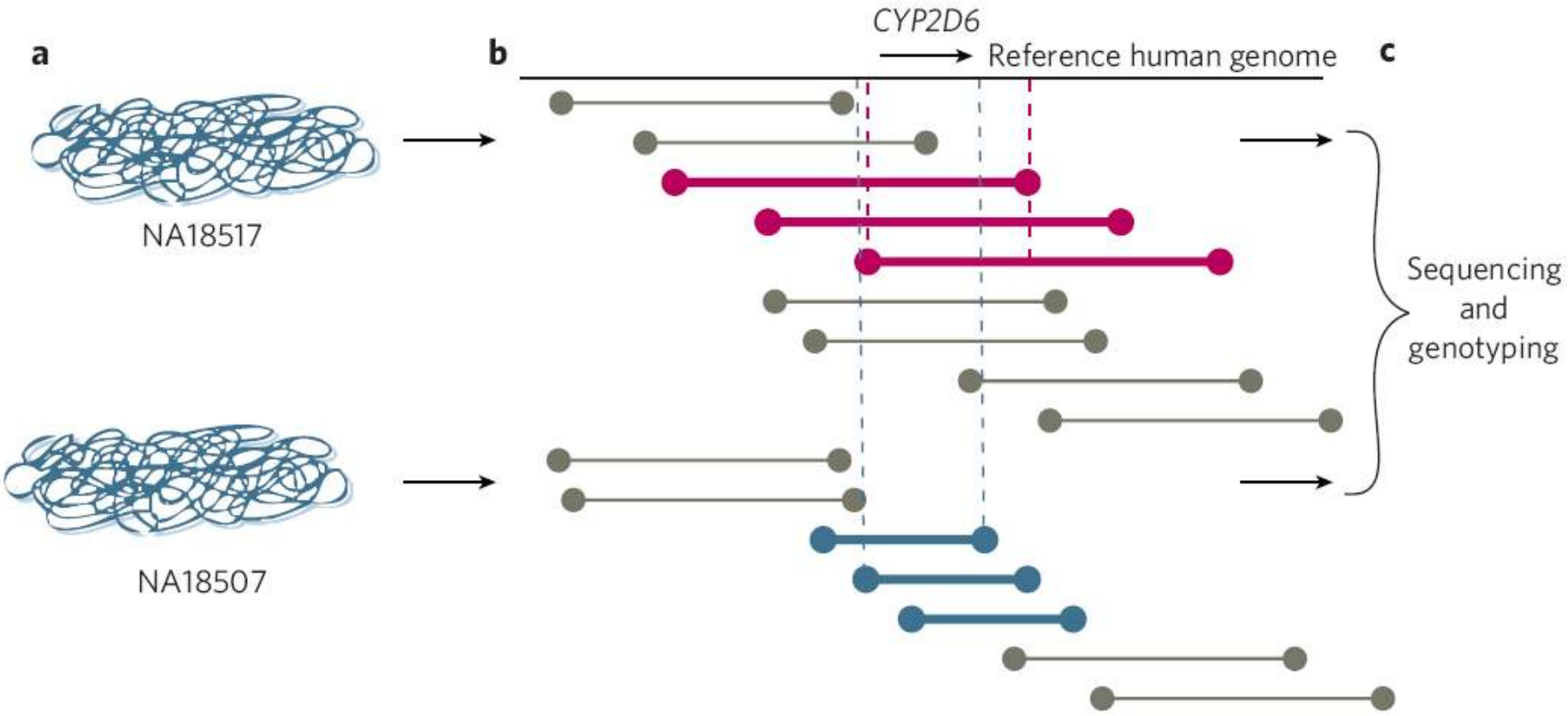
Mapping structural variations



Eichler et al., Nature 447:161 (2007)

Completing the map of human genetic variation

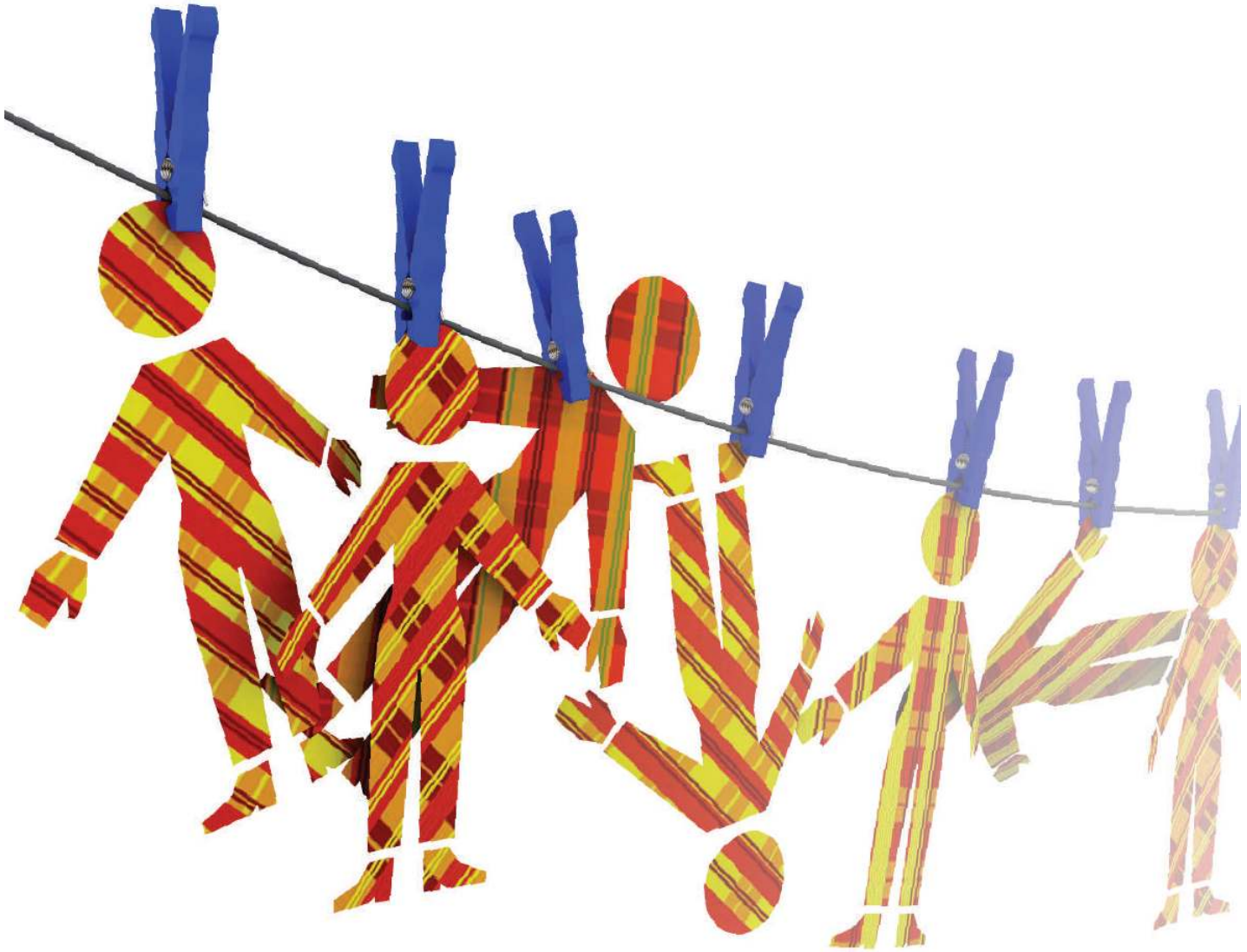
Sequencing structural variations



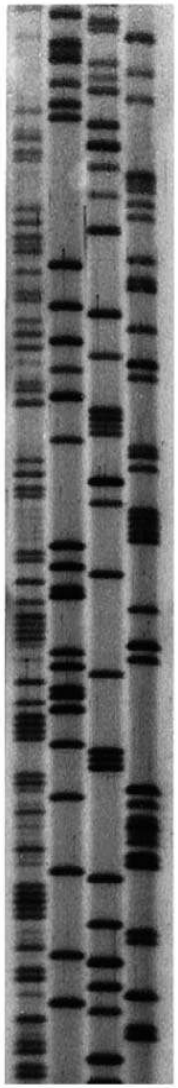
Eichler et al., Nature 447:161 (2007)

Genome Dynamics

Patchwork people ?



News Feature, Nature 437:1084 (2005)



genome.fli-leibniz.de
Lectures