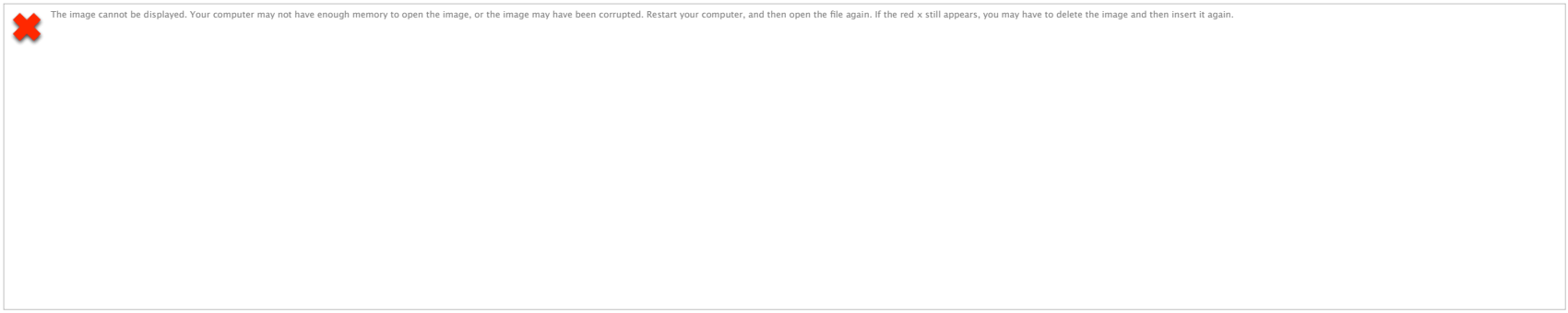


Transposable elements in humans

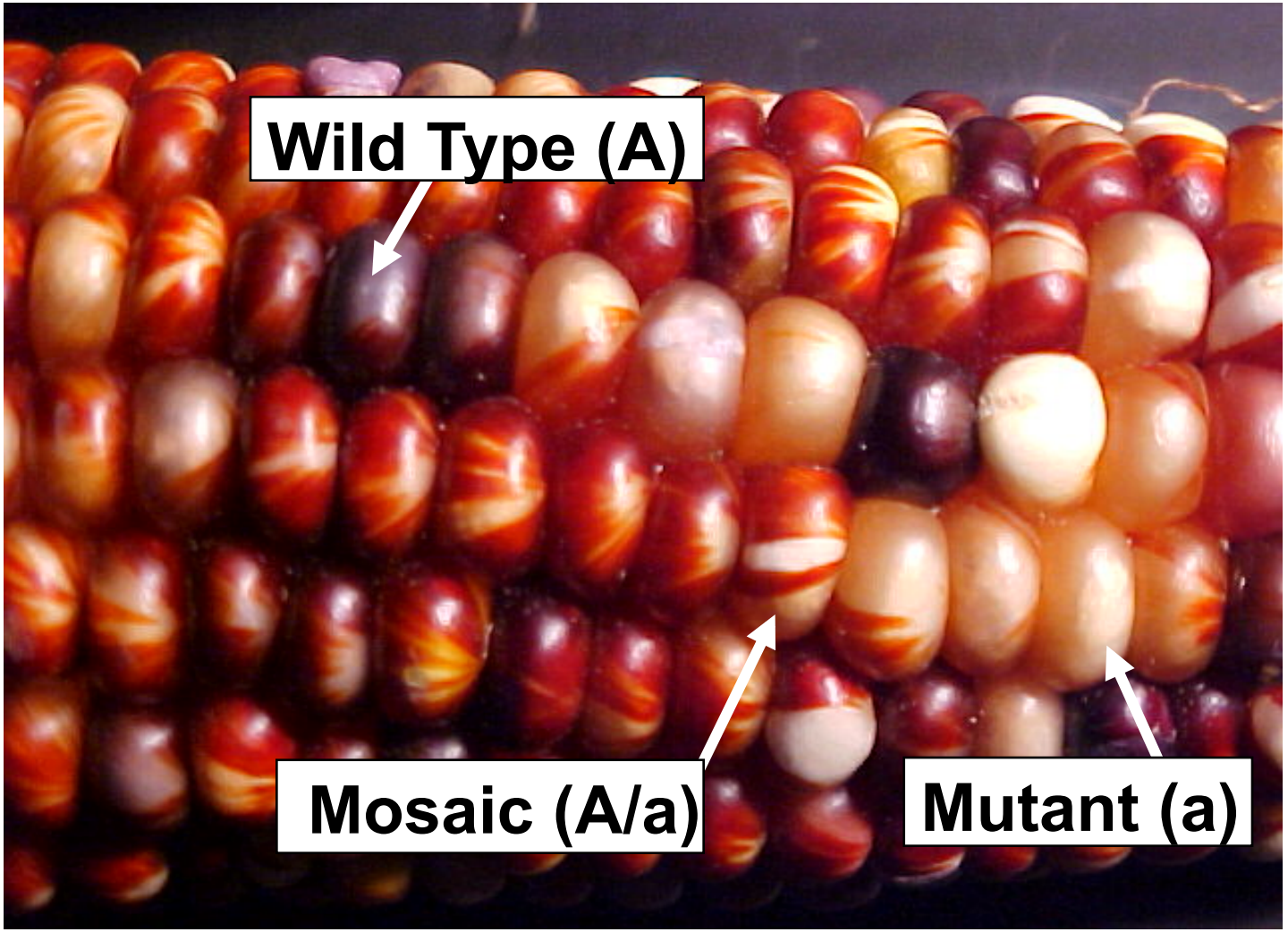
Scott E. Devine, Ph.D.
Associate Professor
Institute for Genome Sciences,
Department of Medicine,
and Greenebaum Cancer Center,
University of Maryland School of Medicine

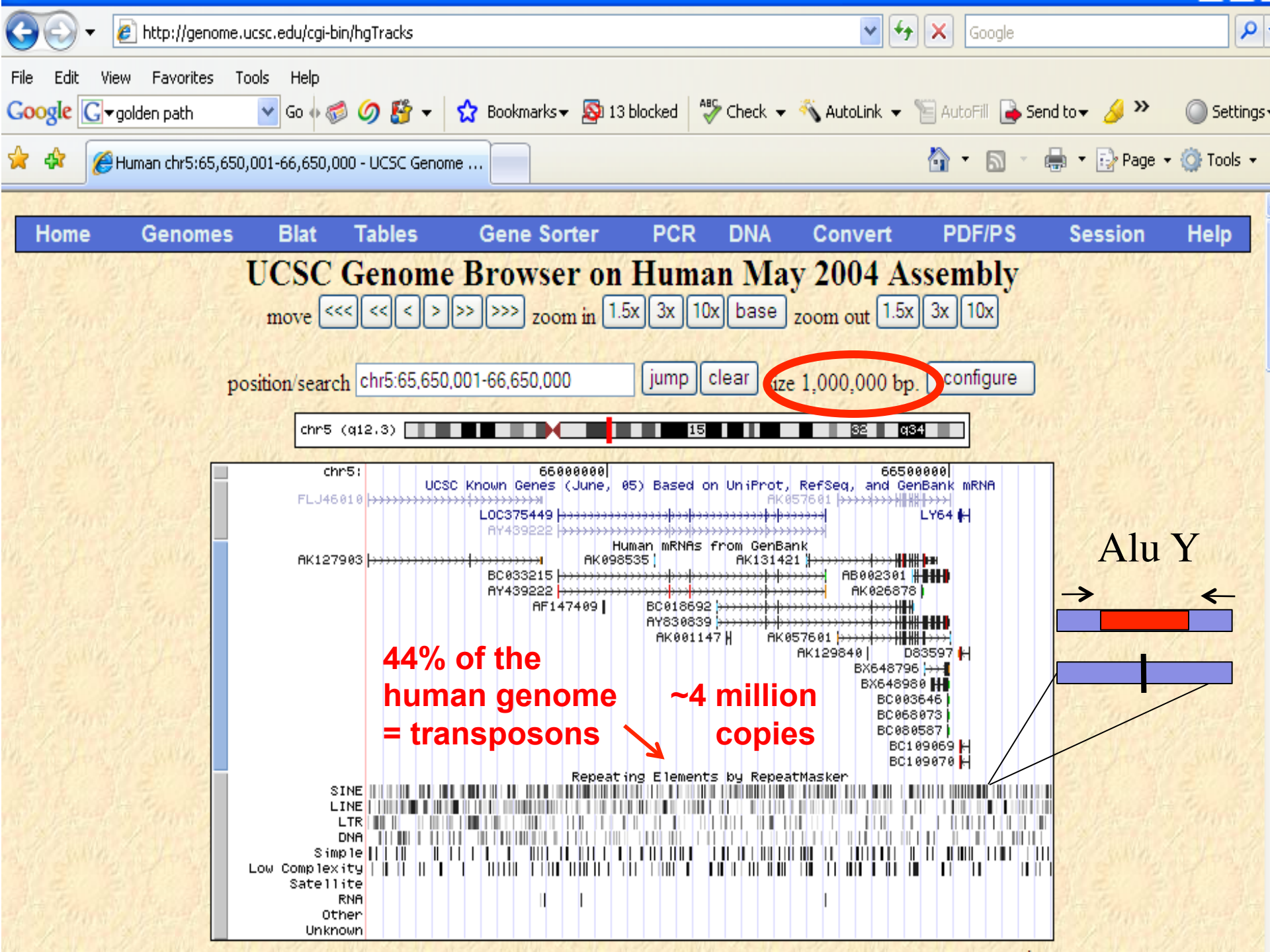
McClintock's transposons caused phenotypic changes

Born 1902, Hartford, CT.
B.A. 1923, Cornell University
Ph.D. 1927, Cornell University, Botany
1927-1931, Instructor in Botany, Cornell University
1931-1933, Fellow, National Research Council
1933-1934, Fellow, Guggenheim Foundation
1934-1936, Research Associate, Cornell University
1936-1941, Assistant Professor, University of Missouri
1942-1967, Staff member, Carnegie Institution of Washington's Department of Genetics, Cold Spring Harbor, NY
1967-1992, Distinguished Service Member, CIW Department of Genetics, Cold Spring Harbor
1944, Member, National Academy of Sciences
1945, President, Genetics Society of America
1967, Kimber Medal
1970, National Medal of Science
1981, Lasker Award
1983, Nobel Prize in Physiology or Medicine

Source: Cold Spring Harbor Laboratory (http://www.cshl.org/History/mcclintock.html)

McClintock's transposons caused phenotypes

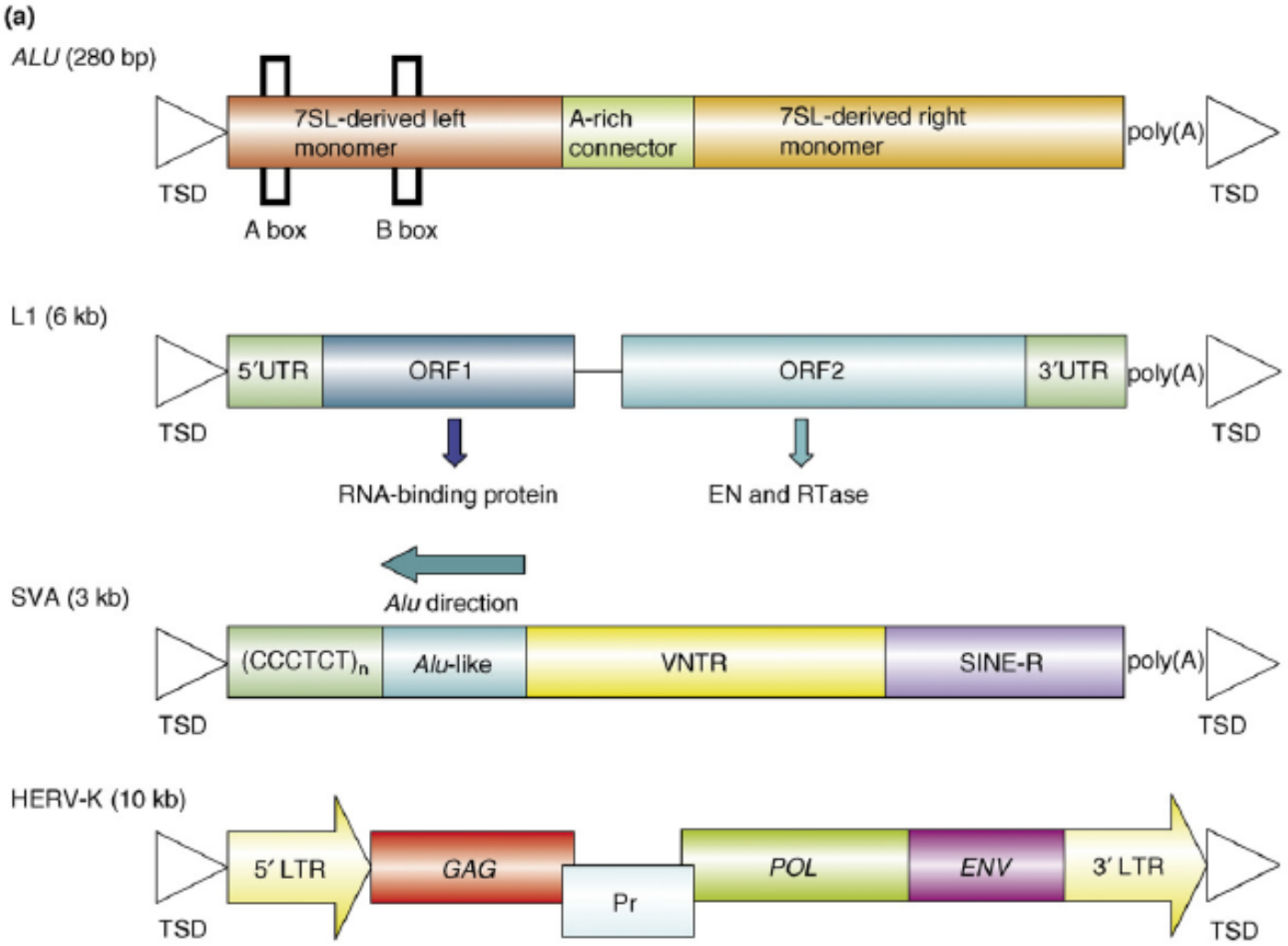



http://genome.ucsc.edu/cgi-bin/hgTracks
Human chr5:65,650,001-66,650,000 - UCSC Genome ...
Home Genomes Blat Tables Gene Sorter PCR DNA Convert PDF/PS Session Help
UCSC Genome Browser on Human May 2004 Assembly
move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x
position/search chr5:65,650,001-66,650,000 jump clear ize 1,000,000 bp. configure
chr5 (q12.3)
chr5: 66000000 66500000
UCSC Known Genes (June, 05) Based on UniProt, RefSeq, and GenBank mRNA
Human mRNAs from GenBank
Repeating Elements by RepeatMasker
SINE
LINE
LTR
DNA
Simple
Low Complexity
Satellite
RNA
Other
Unknown
44% of the human genome = transposons
~4 million copies
Alu Y

Human transposons cause phenotypes (diseases) too

	Gene	Disorder	Element	Mechanism
Alu	NF1	Neurofibromatosis	Alu Ya5	Intron/skipping
	BCHE	Acholinesterasemia	Alu Yb8	Exon insertion
	F9	Hemophilia B	Alu Ya5	Exon insertion
	CASR	Familial hypocalciuric hypercalemia	Alu Ya4	Exon insertion
	ADD1	Huntington disease	Alu	Intron
L1	Factor VIII	Hemophilia A	L1	Exon insertions
	APC	FAP	L1	Exon insertion
	Dystrophin	Muscular Dystrophy	L1	Exon insertions
	Globin	beta thalassemia	L1	Intron
	RP2	Retinitis Pigmentosis	L1	Intron
	Fukutin	Muscular Dystrophy	L1	Intron/skipping

Mills, R.E., Bennett, E.A., Iskow, R.C., and Devine, S.E. (2007) *Trends Genet.* 23:183-91.



Mills, R.E., Bennett, E.A., Iskow, R.C., and Devine, S.E. (2007) *Trends Genet.* 23:183-91.

Human TEs produce a large amount of genetic variation

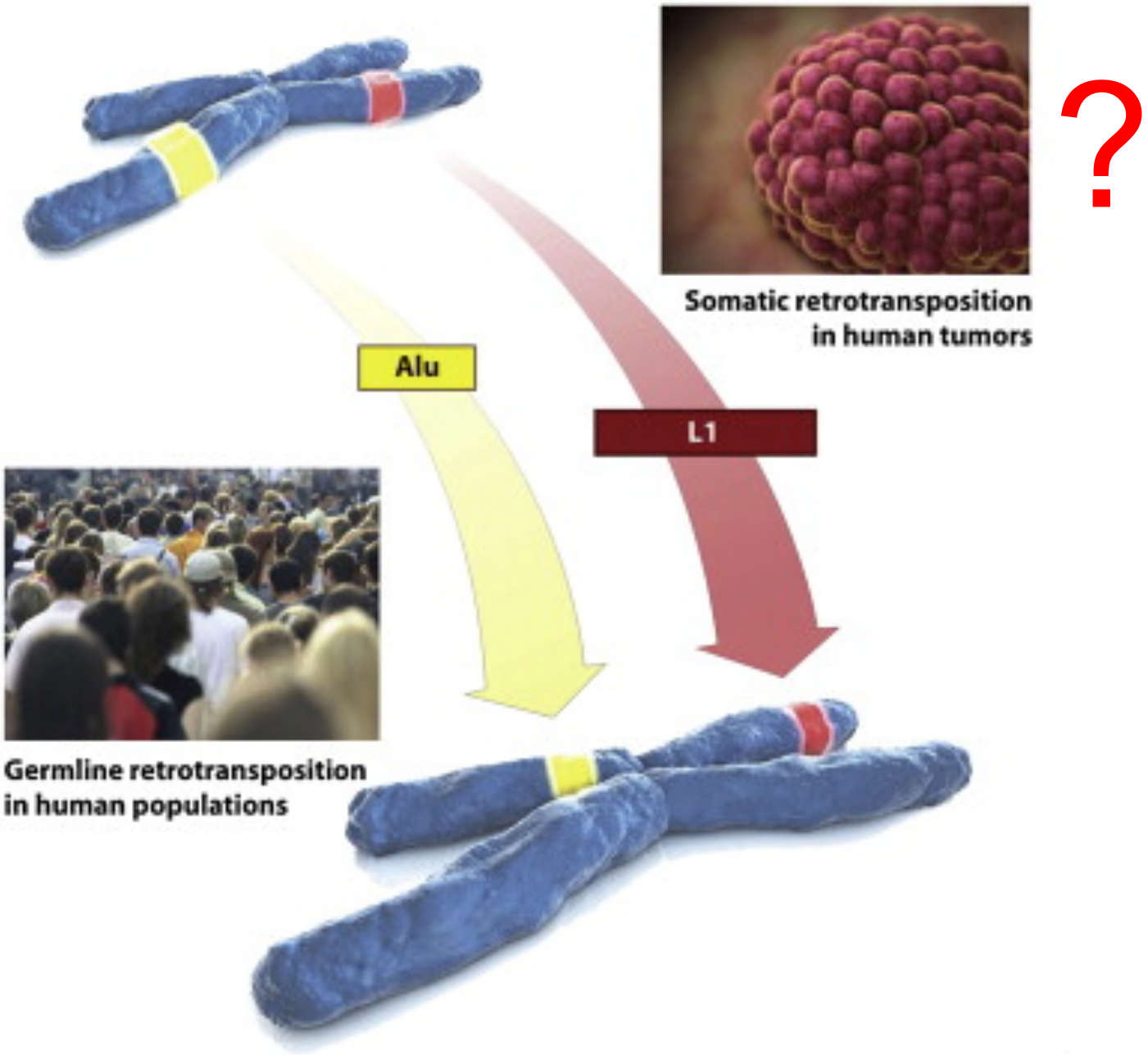
~1 in 20 to 200 live births is predicted to have a new germline TE insertion

Kazazian, H.H. (1999) *Nat. Genet.* **22**, 130.
Li, X. et al. (2001) *Hum. Mutat.* **34**, 511-519.
Cordaux, R. et al. (2006) *Gene* **373**, 134-137.

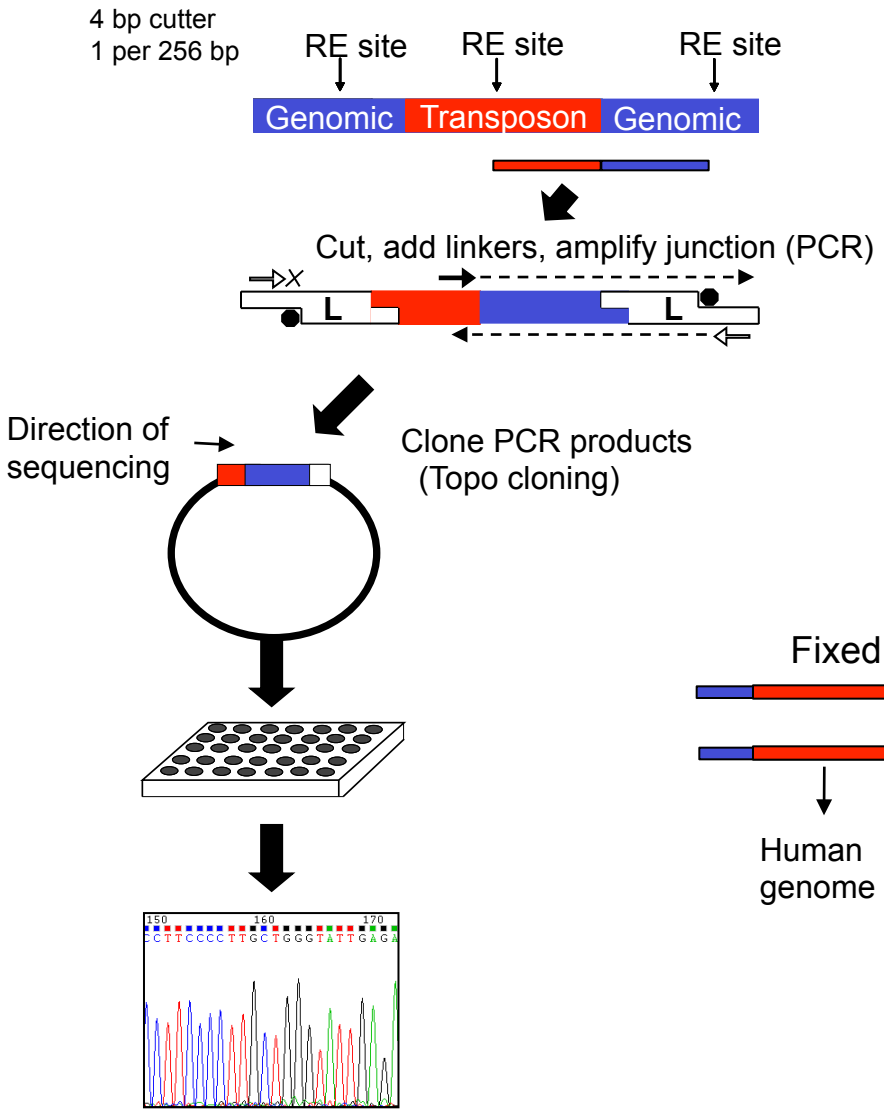
6 billion people X 0.05 = 300 million germline TE insertions

Equivalent to 1 new TE insertion for every 10 bp of genome (an impressive mutagenesis experiment!!)

Needed a better method to find new insertions

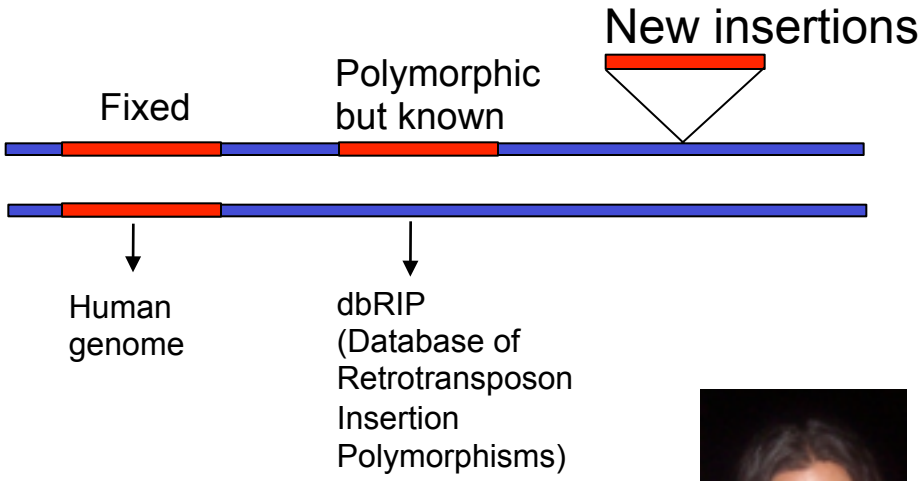


"Transposon-seq" technologies to identify new L1 and Alu insertions



ABI capillary sequencing:

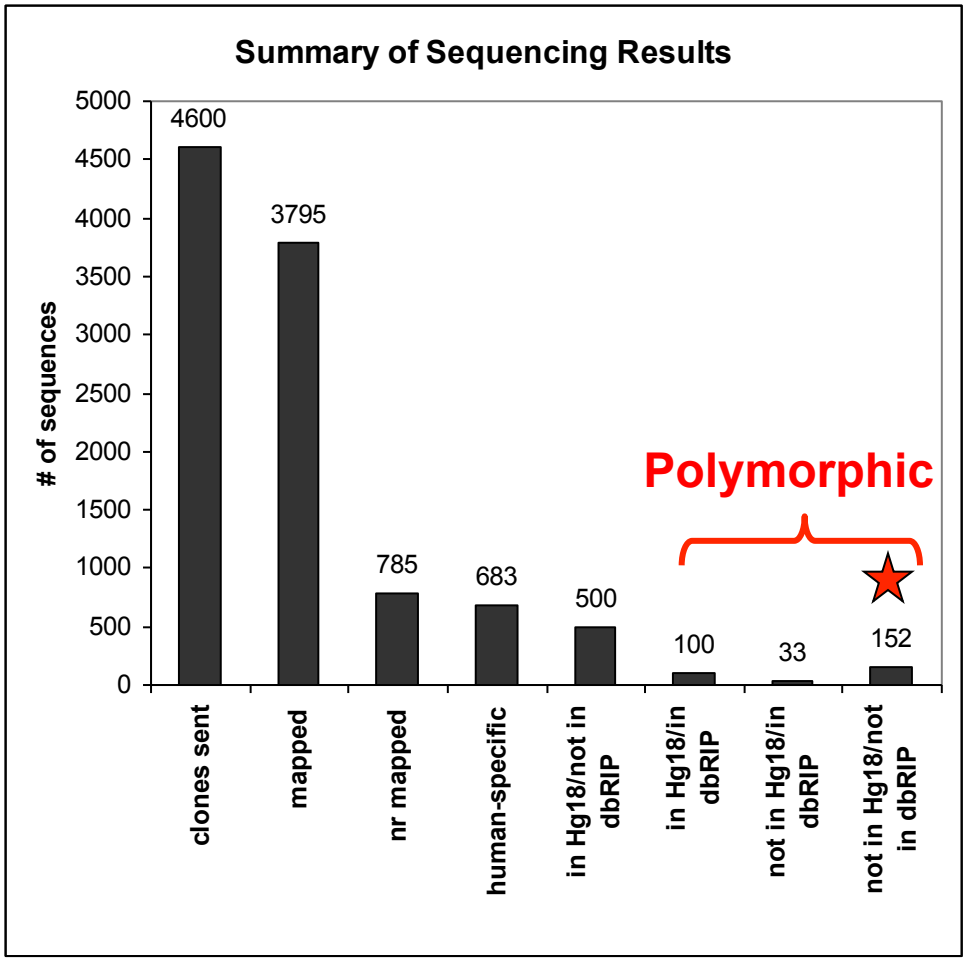
- 4,600 cloned PCR products
- 24 diverse humans-polymorphism discovery resource (pdr) panel (Coriell)
- 14 diverse humans of known geographic origin
- 8 tumor-derived cell lines

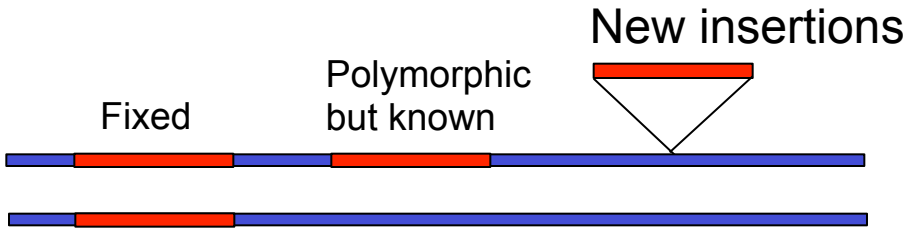


Iskow et al. (2010) Cell 141, 1253-1261

Rebecca Iskow

We found 152 New L1 insertions





Can we detect new germline and somatic TE insertions in cancer patients?

Somatic retrotransposition
in human tumors

?

Alu

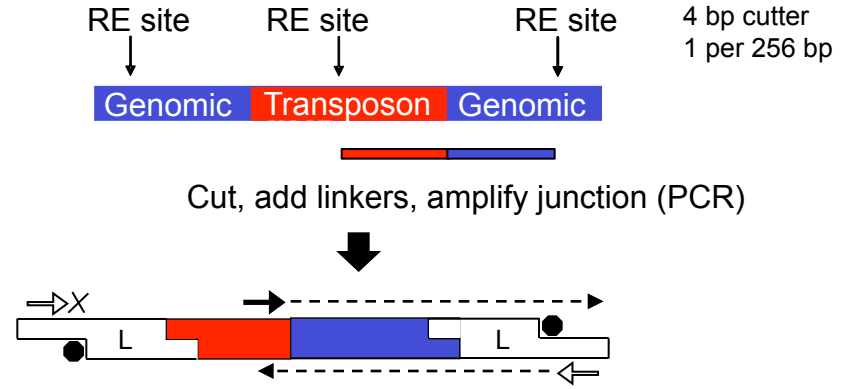
L1

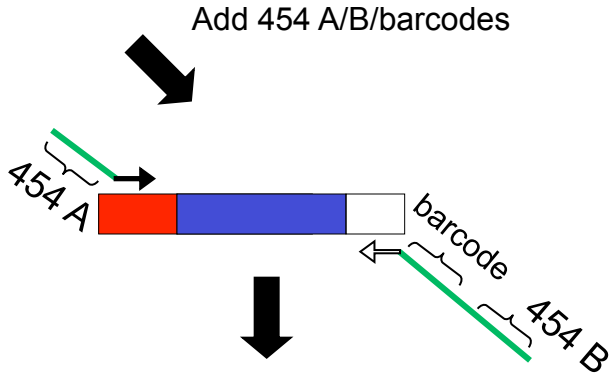
?

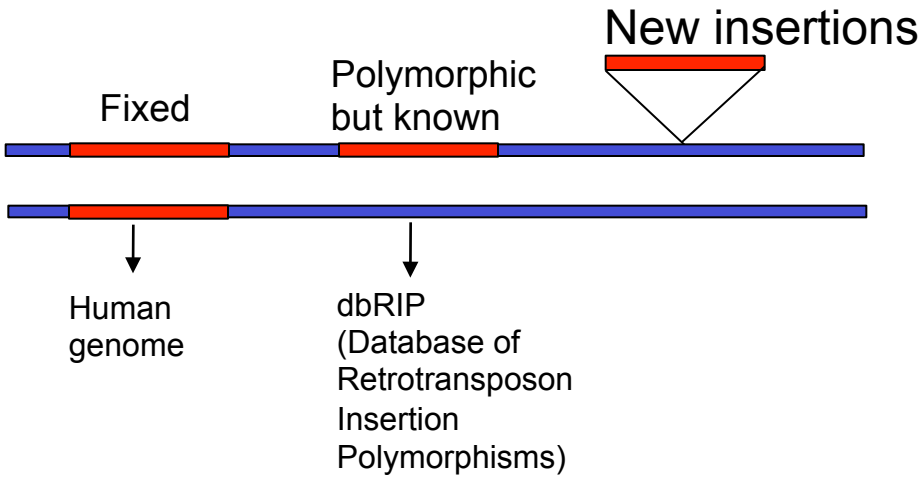
Germline retrotransposition
in human populations

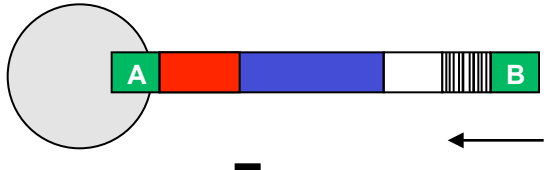
454 pyrosequencing:

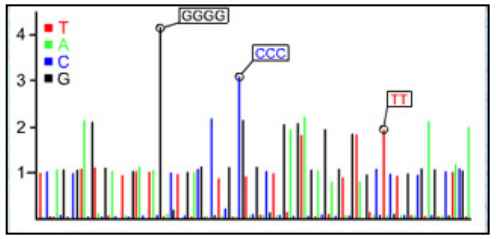
20 lung tumor tissues
20 matched adjacent (normal) tissue
10 brain tumor tissues (5 GMB, 5 MBM)
10 matched blood leukocytes
60 specimens total







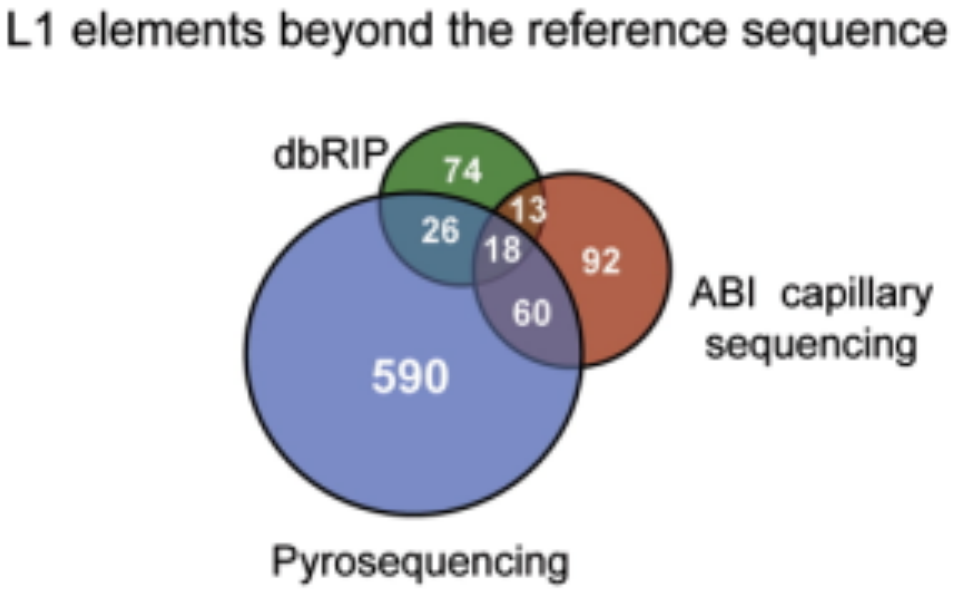




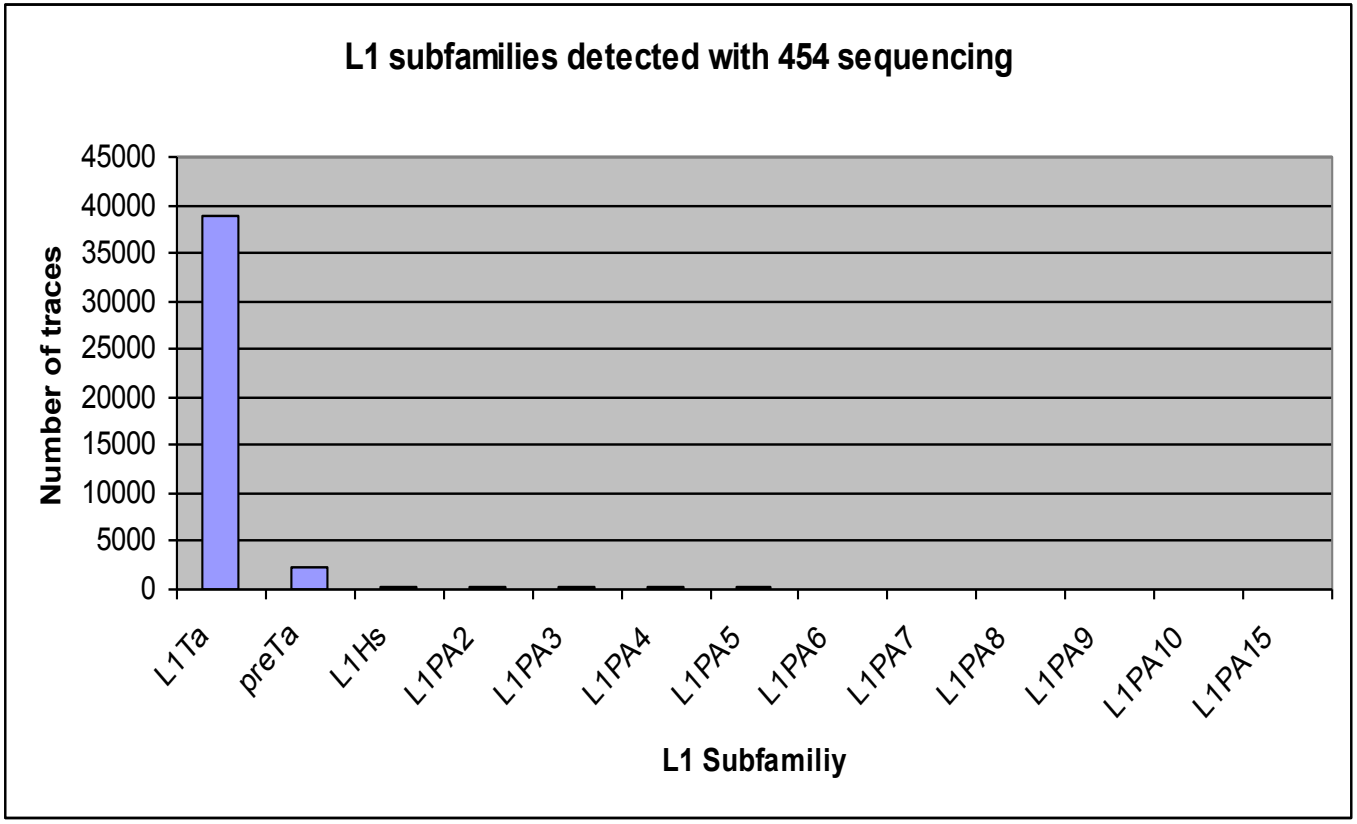
Summary of transposon-seq data

Sequencing Strategy	Sample Description	Retrotransposon	Reads	Mapped	Distinct Retrotransposons	Previously Unknown	PCR Validated
ABI capillary sequencing	Pools of diverse human DNA and tumor-derived cell line DNA.	L1	4,600	3,795	785	152	64/66 (97%)
Pyrosequencing	Lung tumor and adjacent normal lung DNA. Brain tumor and matched normal blood DNA.	L1	266,126	50,532	1,389	650	162/182 (89%)
Pyrosequencing	Brain tumor and matched normal blood DNA.	Alu	35,022	22,338	3,799	403	53/56 (95%)

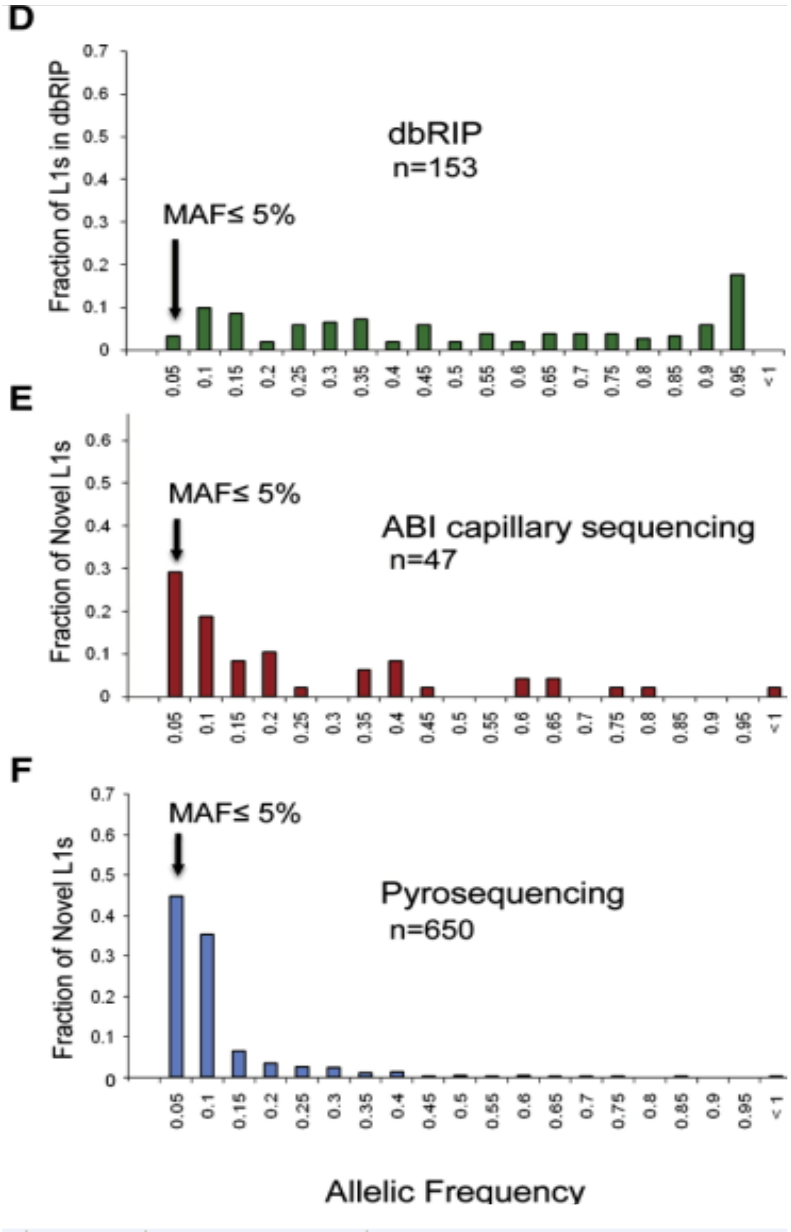
We identified 742 novel L1's
and 403 novel Alus (1145 total)



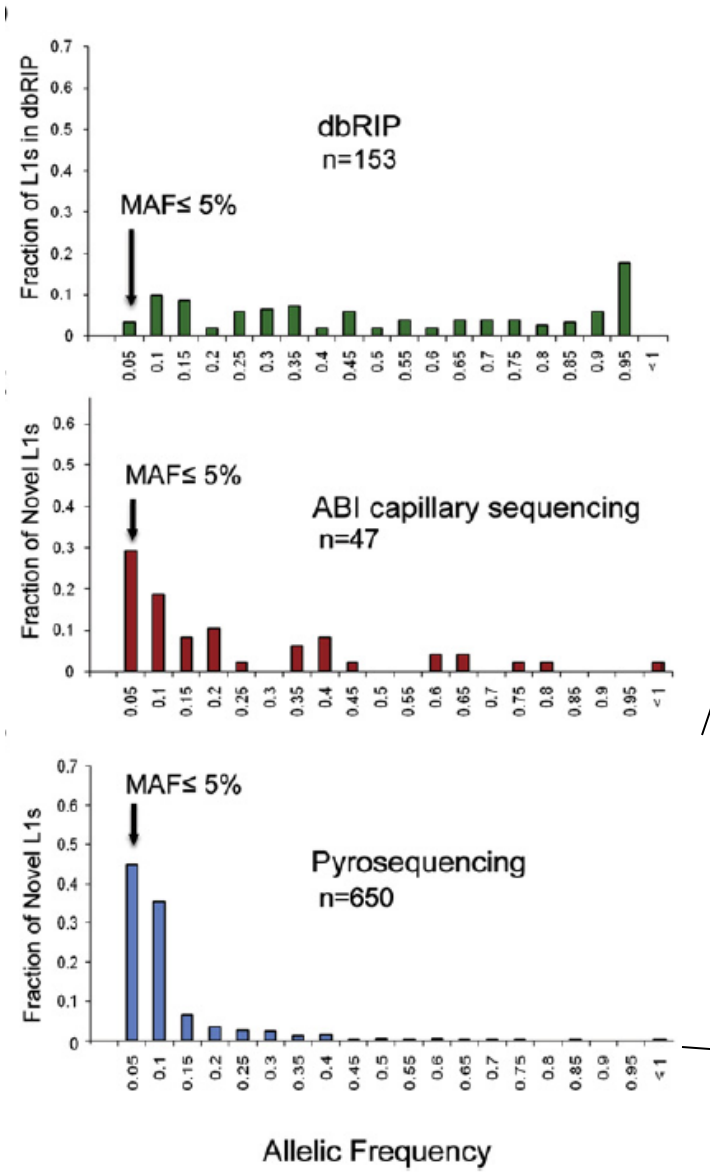
Most of new insertions detected belong to the youngest, most active subfamilies in the genome



Most of the novel insertions also had low allelic frequencies, further indicating that they were inserted recently



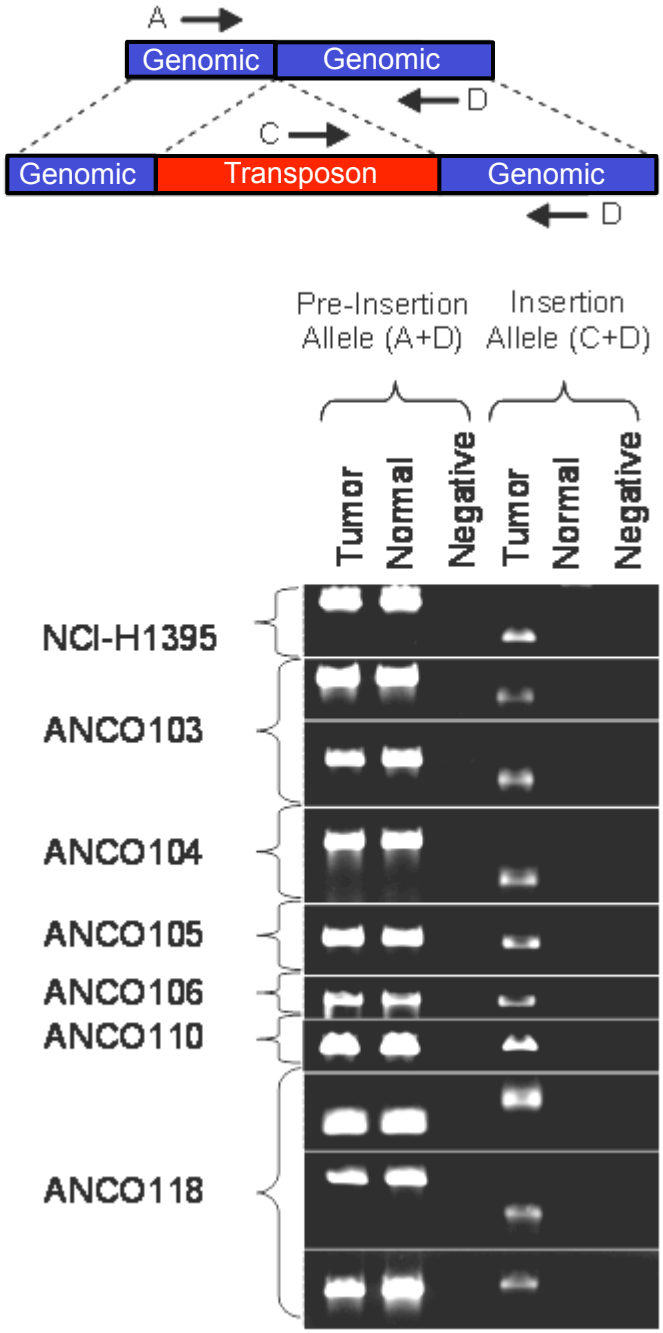
Singletons (insertions found only in a single specimen) Include some of the most recently-inserted L1's



339 "singleton" L1's

Glioblastoma (GBM):	57
Normal leukocytes, GBM:	56
Medulloblastoma (MBM):	18
Normal leukocytes, MBM:	32
Lung tumor:	101
Normal lung:	75

Follow up PCRs on lung tumor singletons:
9 were somatic insertions;
Most of the rest were in matched normal too, so rare germline insertions
~10% no PCR (some likely subclonal)



9 somatic insertions
In lung cancers validated by PCR

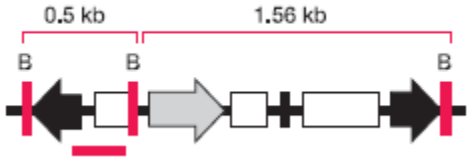
Cloned and sequenced Insertion junctions to Verify—flanked by tsd's

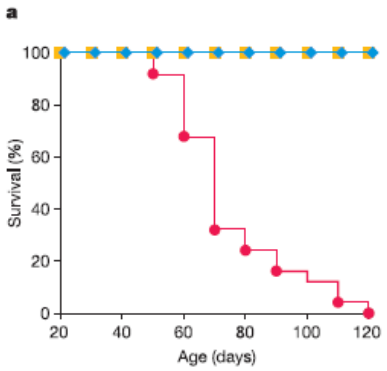
6/20 (30%) of lung tumors positive for at least one new L1 Insertion

ARTICLES

Mammalian mutagenesis using a highly mobile somatic *Sleeping Beauty* transposon system

Adam J. Dupuy[1], Keiko Akagi[1], David A. Largaespada[2], Neal G. Copeland[1] & Nancy A. Jenkins[1]





b

Mouse ID	Age (days)	Tumour type
TG6057-16315	114	T-cell
TG6057-17106	101	Medulloblastoma
TG6057-17306	84	T-cell, non-T/non-B
TG6057-18137	45	B-cell
TG6070-16392	57	B-cell
TG6070-16515	65	T-cell (thymic), non-T/non-B (disseminated)
TG6070-16881	58	Non-T/non-B
TG6070-16887	62	T-cell
TG6070-17221	60	T-cell
TG6070-17223	87	B-cell, hyperplasia in duodenum
TG6070-17503	102	Non-T/non-B
TG6070-17900	68	T-cell (thymic), non-T/non-B (disseminated)
TG6070-17902	68	T-cell
TG6070-18533	73	T-cell
TG6070-18665	53	T-cell
TG6070-18732	47	T-cell, hyperplasia in duodenum
TG6070-18804	58	Non-T/non-B, medulloblastoma, intestinal neoplasia
TG6113-16835	67	T-cell
TG6113-16927	96	T-cell, pituitary hyperplasia
TG6113-16930	71	T-cell
TG6113-17548	52	T-cell, non-T/non-B
TG6113-18019	59	Non-T/non-B
TG6113-18022	62	T-cell, non-T/non-B
TG6113-18521	62	T-cell

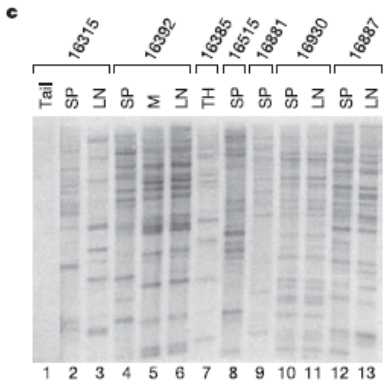
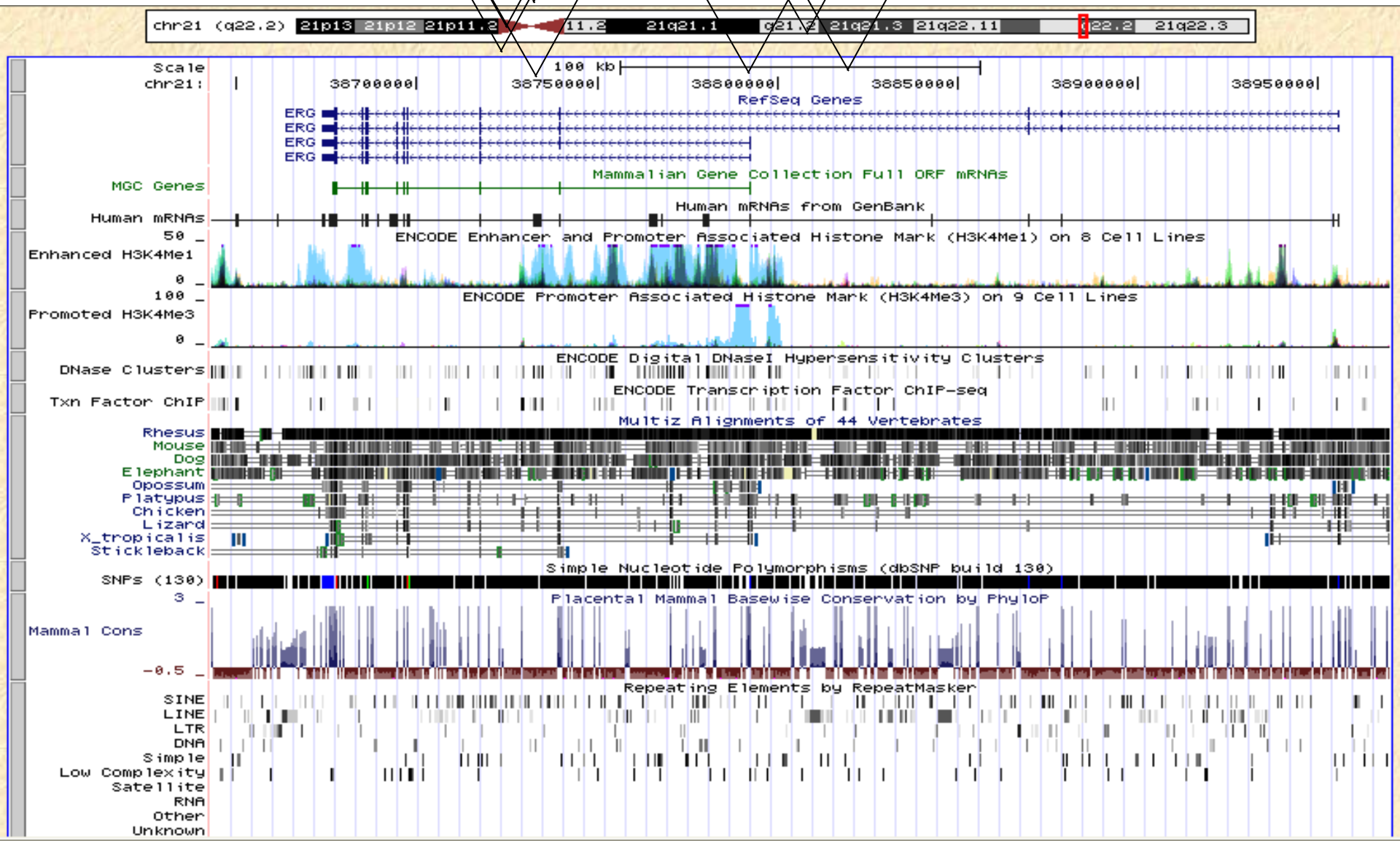


Figure 2 | Adult double-transgenic mice die from cancer. **a**, Survival curves showing decreased viability of double-transgenic mice: yellow, RosaSB; blue, T2/Onc2; red, double-transgenic. **b**, Age at death and tumour type of double-transgenic mice. **c**, Southern analysis of *Bam*HI-digested tumour DNA. Each band represents a separate SB transposon integration. LN, lymph node; M, mass; SP, spleen; TH, thymus.

ERG (frequently mutated in cancers)

Dupuy et al. 2005
SB-induced mouse tumors

Iskow et al. 2010
(BL-normal control for MBM)

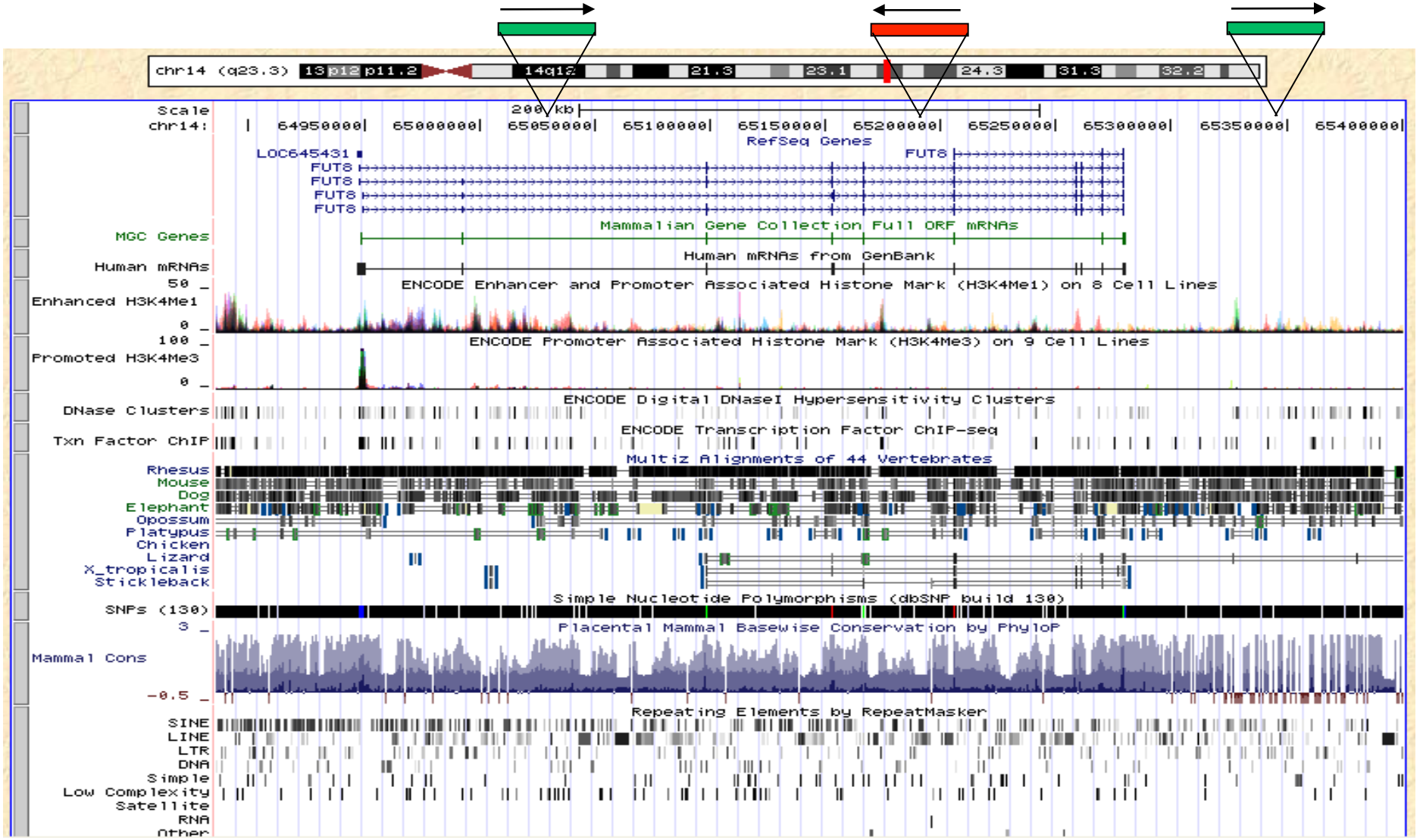


FUT8 (Fucosyltransferase 8—linked to tumor invasiveness)

Dupuy et al. 2005 (SB-induced mouse tumor)

Iskow et al. 2010 (human lung tumor-somatic)

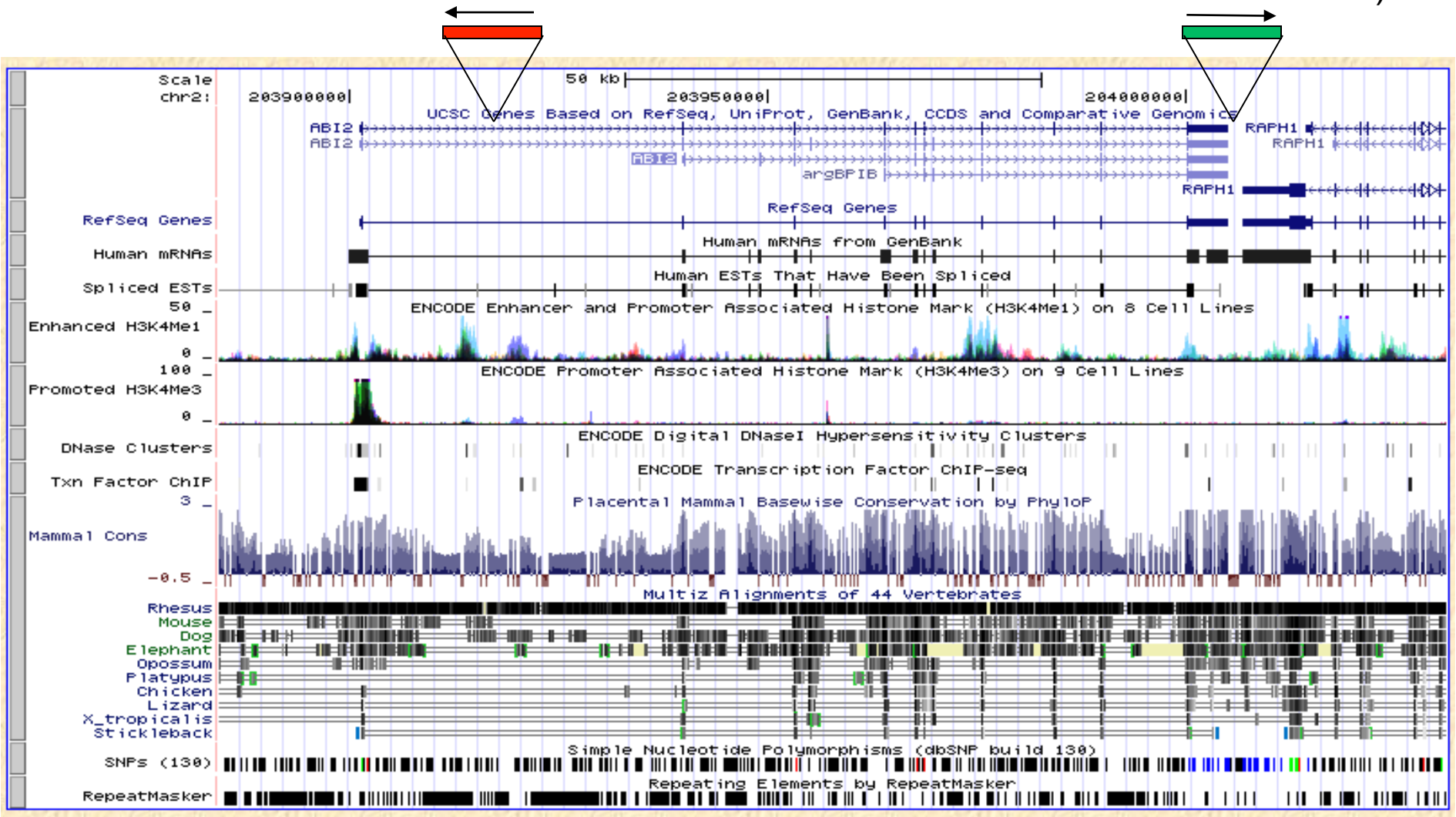
Dupuy et al. 2005 (SB-induced mouse tumor)

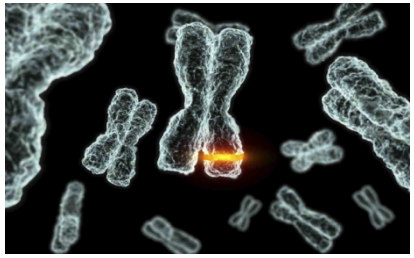


ABI-2 (ABL-Interacting-2; antagonizes oncogenic potential of ABL)

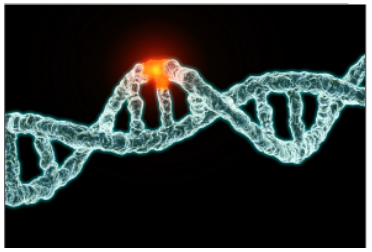
Iskow et al. 2010
(human lung tumor)

Dupuy et al. 2005
(SB-induced mouse
medulloblastoma)





Conclusions



1. Transposon-seq revealed that rare Alu, L1 germline insertions are highly abundant in human genomes. Likely to have greatest Impact on humans because selection has not yet fully acted on them.

2. Somatic L1 insertions also occur at high frequencies in lung tumor genomes

3. Both germline and somatic insertions may work together to initiate and then drive tumorigenesis by hitting specific genes

4. Provides a new mechanism for mutagenizing cancer genomes in addition to point mutations, DNA repair, large-scale rearrangements

5. Once integrated, L1 might also promote large scale chromosomal rearrangements that are commonly found in human tumors

Emory University

Steve Pittard

Ryan Mills

R01HG002898
2T32GM00849
SUN Microsystems

Paula Vertino
Erwin Van Meir
Mike McCabe

University of Maryland Institute for Genome Sciences

Julienne Mullaney	Research Associate
Luke Tallon	Genomics Resource Center
Jerry Liu	Senior bioinformatics engineer
Ankit Maroo	Bioinformatics analyst
Brandi Cantarel	Research associate
Ellen McRae	Postdoctoral fellow
Andy Neuwald	IGS faculty member
Steve Pittard	Senior informatics engineer

R01HG002898-07, R01CA166661-01