High Throughput Detection of Insertion Sites

Christof von Kalle

NCT
NATIONALES CENTRUM
FÜR TUMORERKRANKUNGEN
HEIDELBERG

LAM + Integration Site Sequencing: –2006/07

2003 → 2006/07

	ADA-SCID	SCID-X1 France	SCID-X1 UK	CGD
vector	MLV	MLV	MLV	SFFV
transgene	ADA	IL2RG	IL2RG	CYBB
patients	2	9	6 (+ 1 nd)	2
mappable RIS	9	475	560	722
RIS post-transplantation	9	379	295	722
RIS pre-transplantation	---	96	265	---

# of Insertions

NCT

2006 ff: Next Generation Sequencing

- 2nd, 3rd, …, x^{th}? generation sequencing technologies:
- 454/Roche, Solexa/Illumina, ABI, Helicos, PacBios

- 454 pyrosequencing GSflx (>400000 reads, 250bp)
- 454 pyrosequencing Titanium (~1x10e6, 400bp)

- 2013: <15min human genome sequencing
- 1000 Dollar genome

NCT

LAM-PCR + Next Generation Sequencing (NGS) to Reach

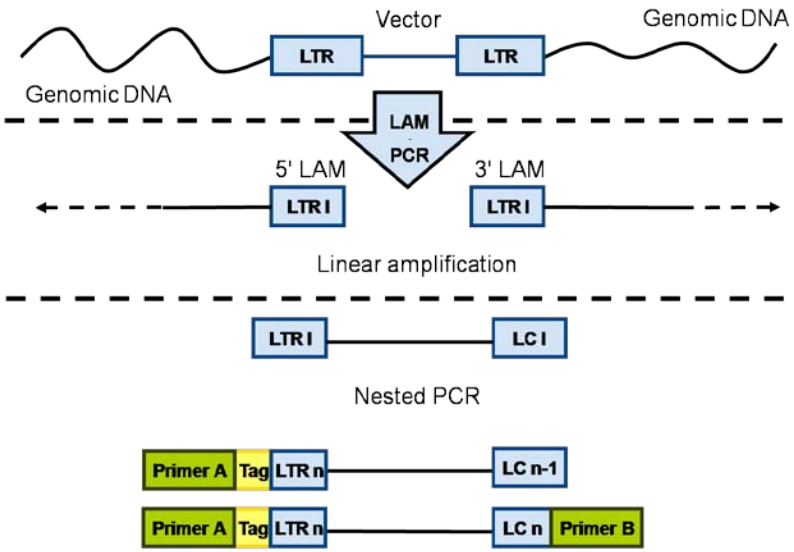
Comprehensive Integration Site Analysis

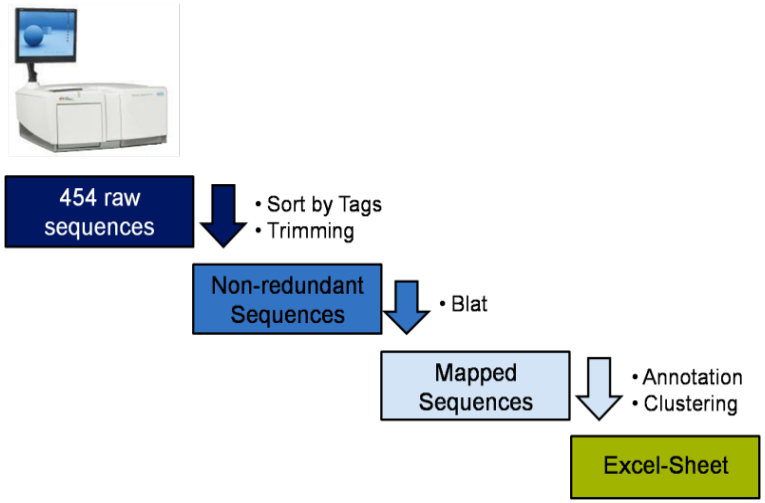
1. Combine LAM-PCR and next generation sequencing
2. Define the limitations introduced by the use of restriction enzymes in current integration site assays
3. Maximize the access to insertions *a priori*
4. Develop a non-restrictive integration site approach

NCT

1. Combining LAM-PCR and NGS (454 Pyrosequencing)

LAM-PCR and 454 Pyrosequencing and Datamining



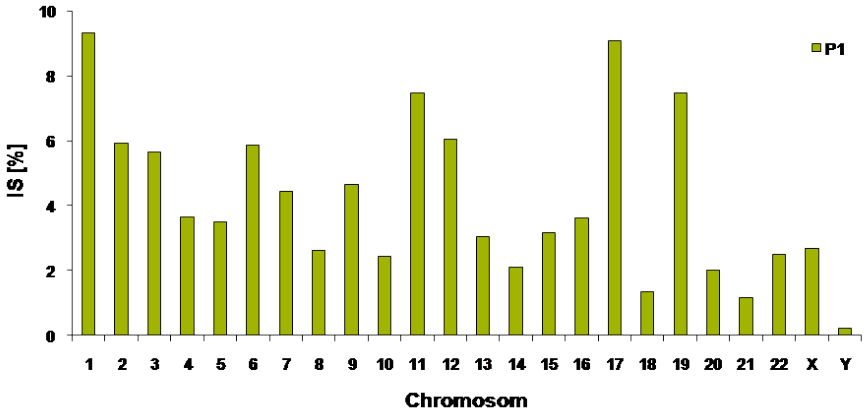


→>500.000 LAM amplicon sequences / 2-3 days instead of 65.000 LAM amplicons / 10 years

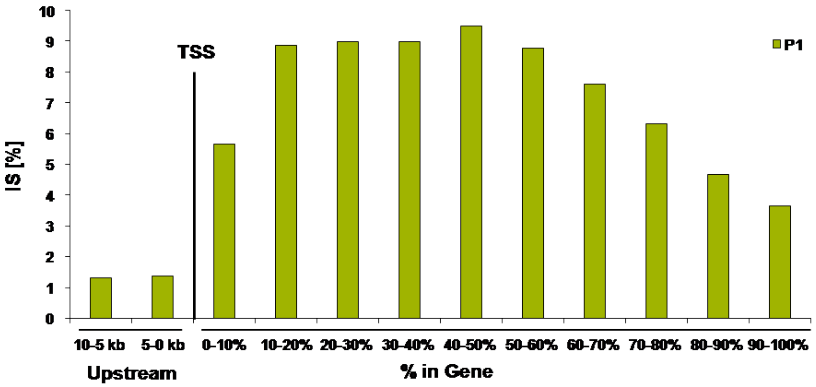
NCT

Large Scale Integration Site Profiling and Characteristics

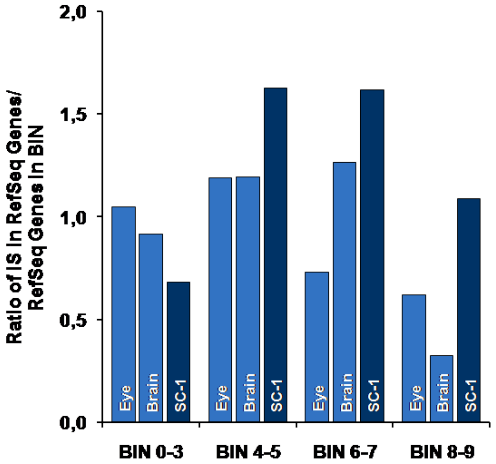
Chomosomal IS Distribution



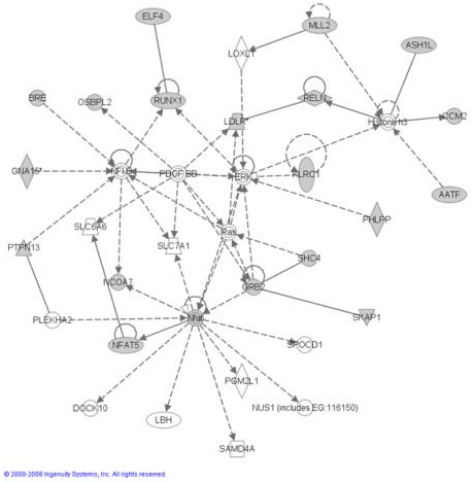
Gene Related IS Distribution



Functional IS Analysis



Ingenuity Pathway Analysis (IPA)

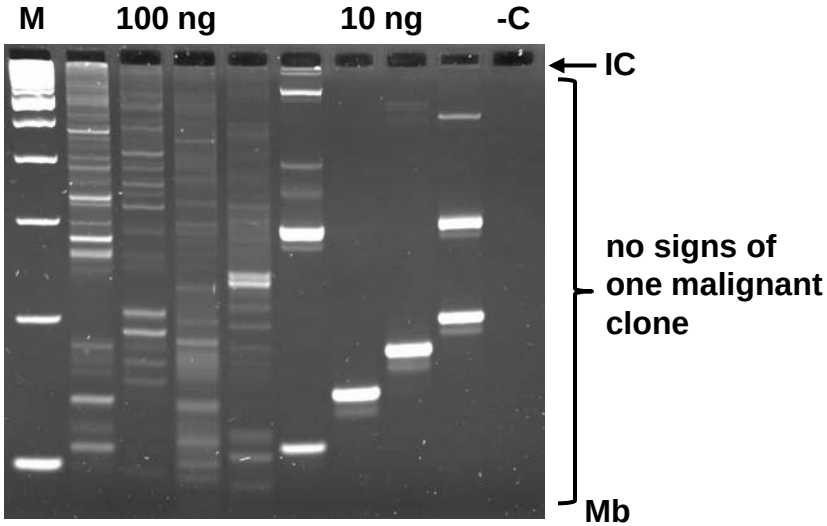


NCT

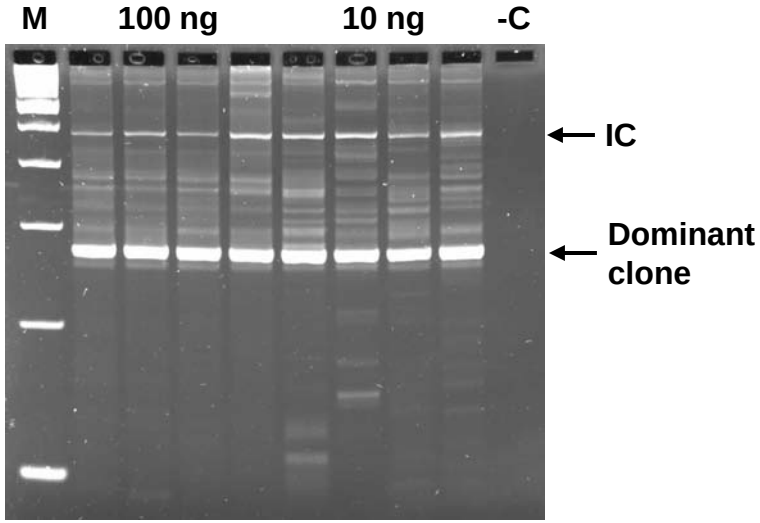
2. Limitations in Integration Site Retrieval

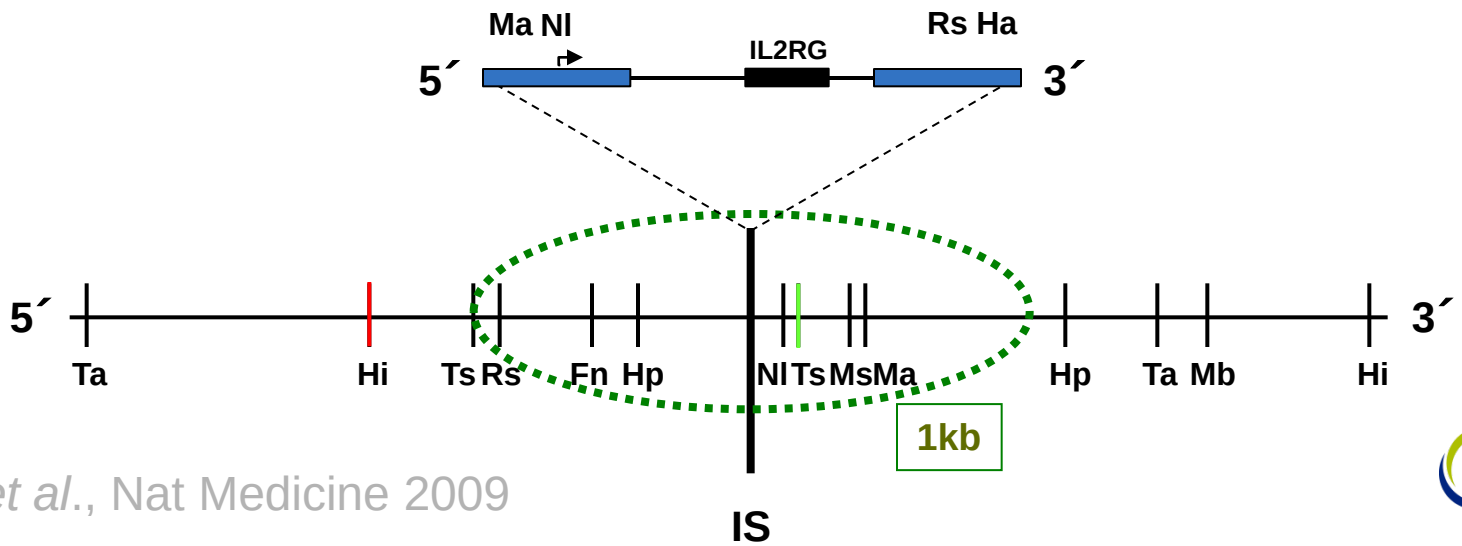
(by Restriction Motif and Amplification Length)

5‘LAM HinP1I (GCGC)



3‘LAM Tsp509I (AATT)





Gabriel *et al*., Nat Medicine 2009

NCT

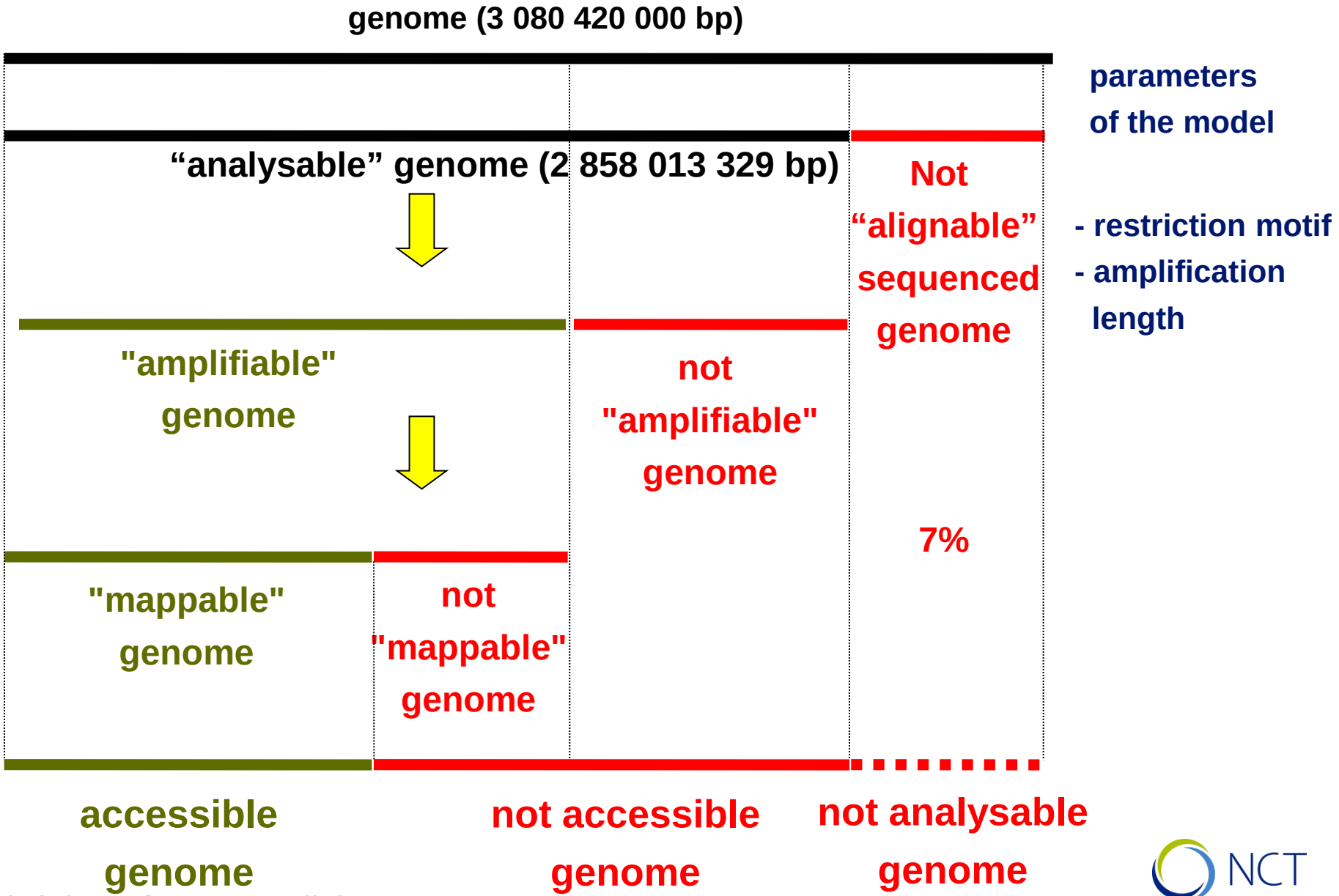
Distribution of 4-bp Restriction Motifs in the Human Genome

Restriction enzyme	Recognition motif	Average distance between two consecutive motifs [bp]	Standard deviation [bp]
Tsp509I	AATT	133,32	181,11
MseI	TTAA	150,17	213,8
NlaIII	CATG	208,93	224,91
MaeI	CTAG	368,08	389,59
MboI	GATC	401,35	419,68
RsaI	GTAC	566,75	606,12
HpaII	CCGG	1246,19	2199,15
HpyCH4IV	ACGT	1330,49	1441,09
HinP1I	GCGC	1728,91	3148,72
TaqI	TCGA	1893,08	2088,65
FnuDII	CGCG	3941,91	8721,09

Gabriel *et al*., Nat Medicine 2009

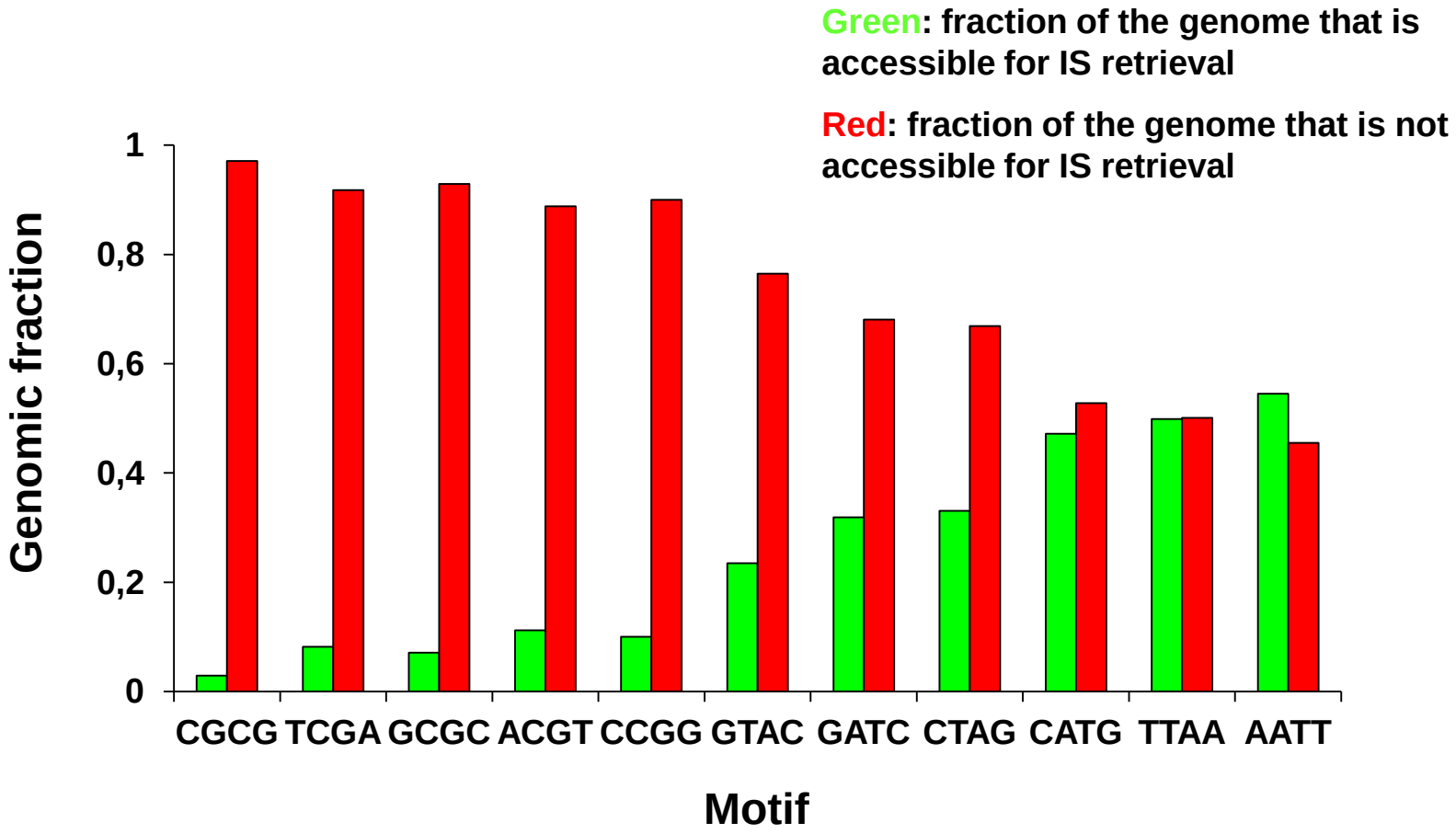
NCT

Modeling the Genome Accessibility



Gabriel *et al*., Nat Medicine 2009

3. Maximize the Access to Insertions *a priori*

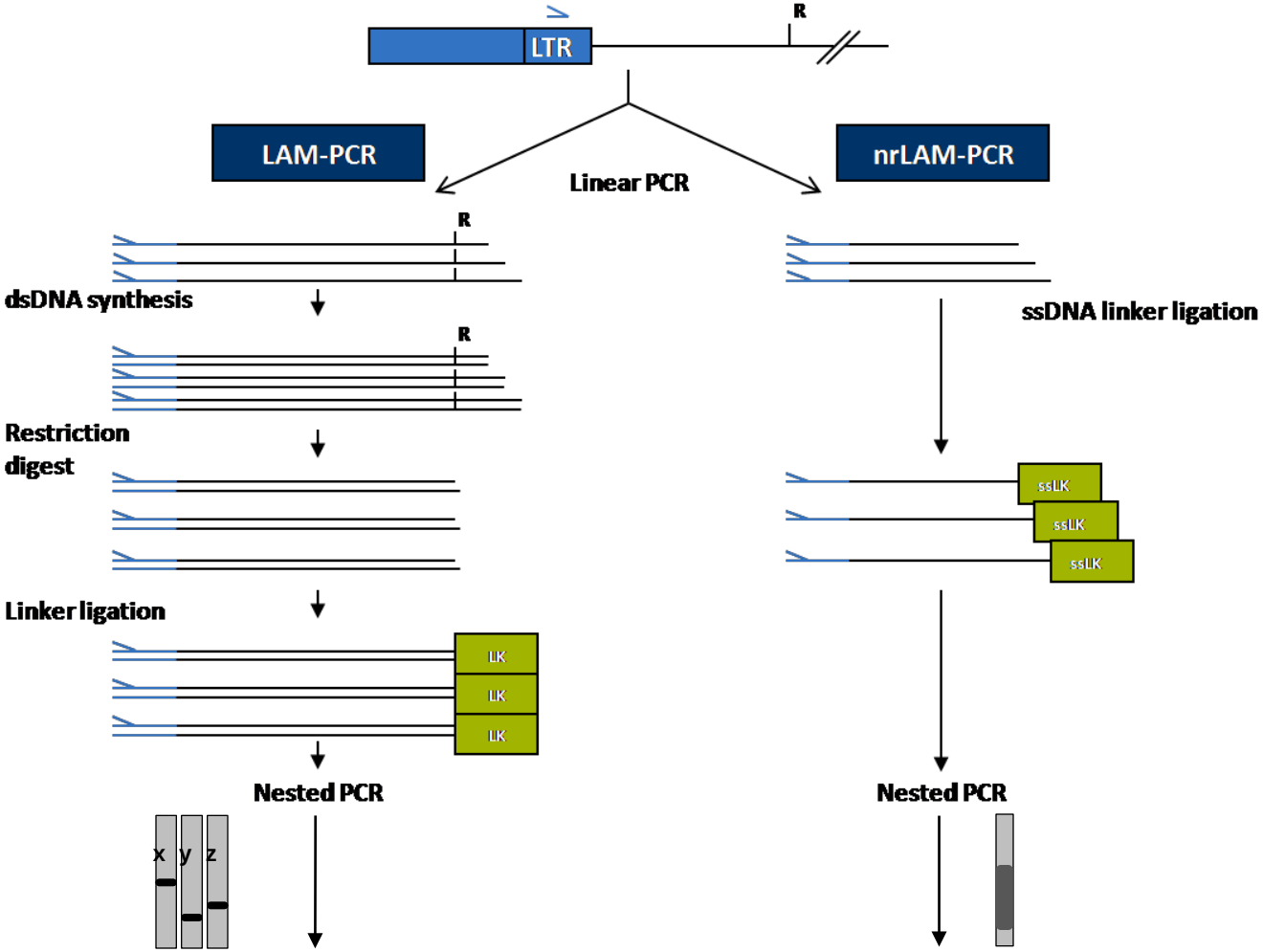


→ **CG motifs:** **2.9% - 11.2%**

→ **non-CG motifs:** **23.5% - 54.5%**

Gabriel *et al*., Nat Med 2009

4. Develop a Non-Restrictive Integration Site Approach (non-restrictive LAM-PCR)

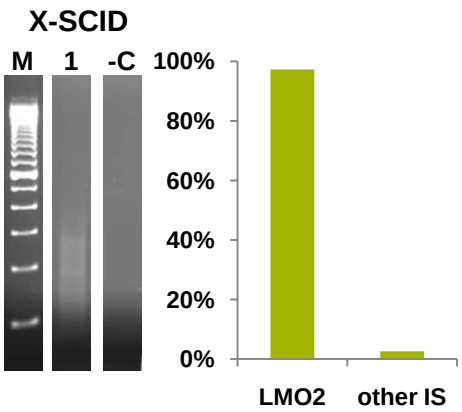


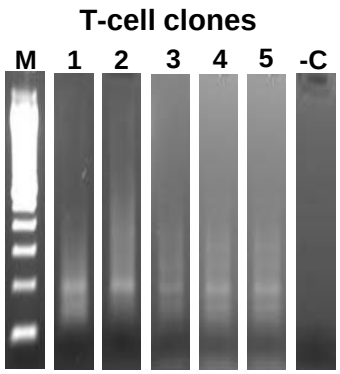
454 pyrosequencing

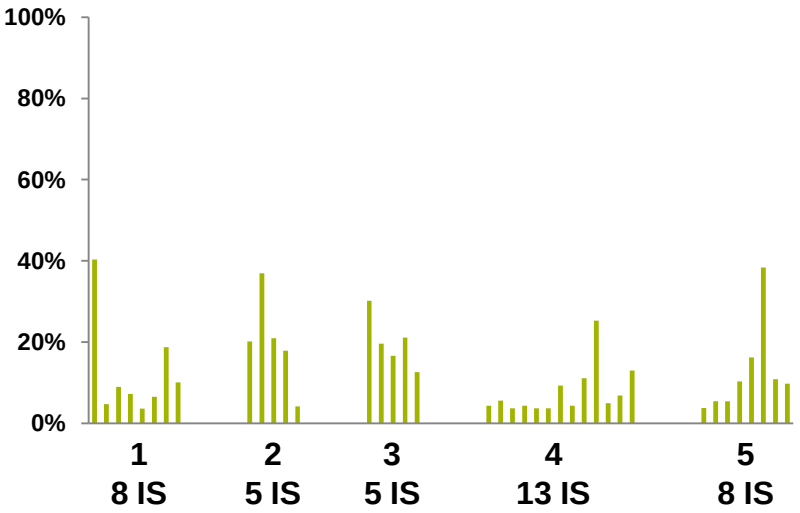
454 pyrosequencing

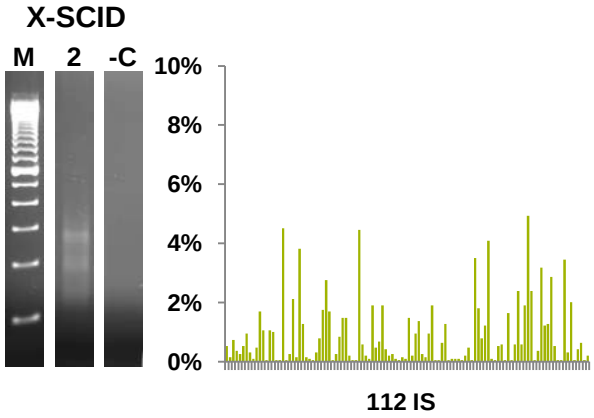
Gabriel *et al*., Nat Med 2009

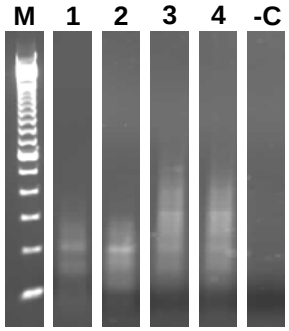
Analysis of Different Preclinical and Clinical Samples with nrLAM-PCR

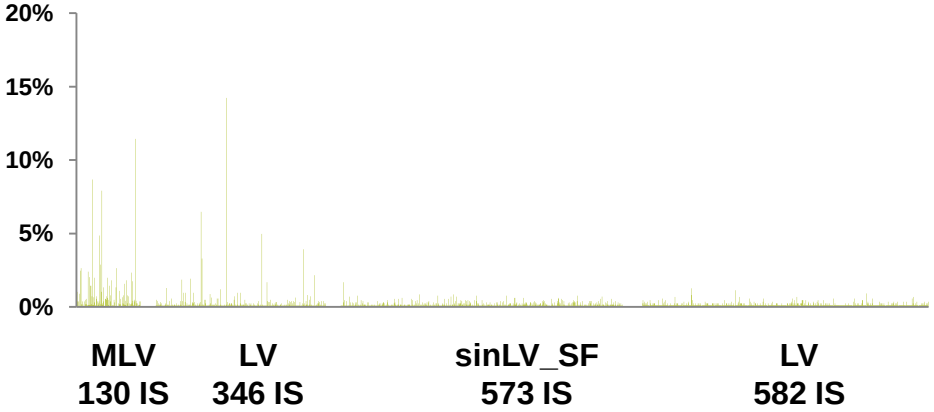












Gabriel *et al*., Nat Medicine 2009

NCT

Acknowledgements

Christof von Kalle
Hanno Glimm
Cynthia Bartholomae
Ali Nowrouzi
Annette Deichmann
Anna Paruzynski
Wei Wang
Richard Gabriel
Simone Scholz
Christine Käppel
Eliana Ruggiero
Ina Kutschera
Nadine Krenzer
Christina Lulay
Claudia Ball
Ulrich Abel
Bernhard Korn
Anne Arens
Ralph Eckenberg
William Saurin

EU, BMBF, DFG, Helmholtz

Hans-Jörg Hauser
Dagmar Wirth
Odile C.-Haguenauer
Christopher Baum
Alessandro Aiuti
Philippe Moullier
Nicolas Ferry
David Williams
Donald Kohn
Luigi Naldini
Fulvio Mavilio
Alain Fischer
Marina Cvazzana-Calvo
Gerard Wagemaker
Adrian Thrasher
Bobby Gaspar
Manuel Grez
Patrick Aubourg
Nathalie Cartier
Reinhard Seger

NCT