



# Implementing research in personalized medicine

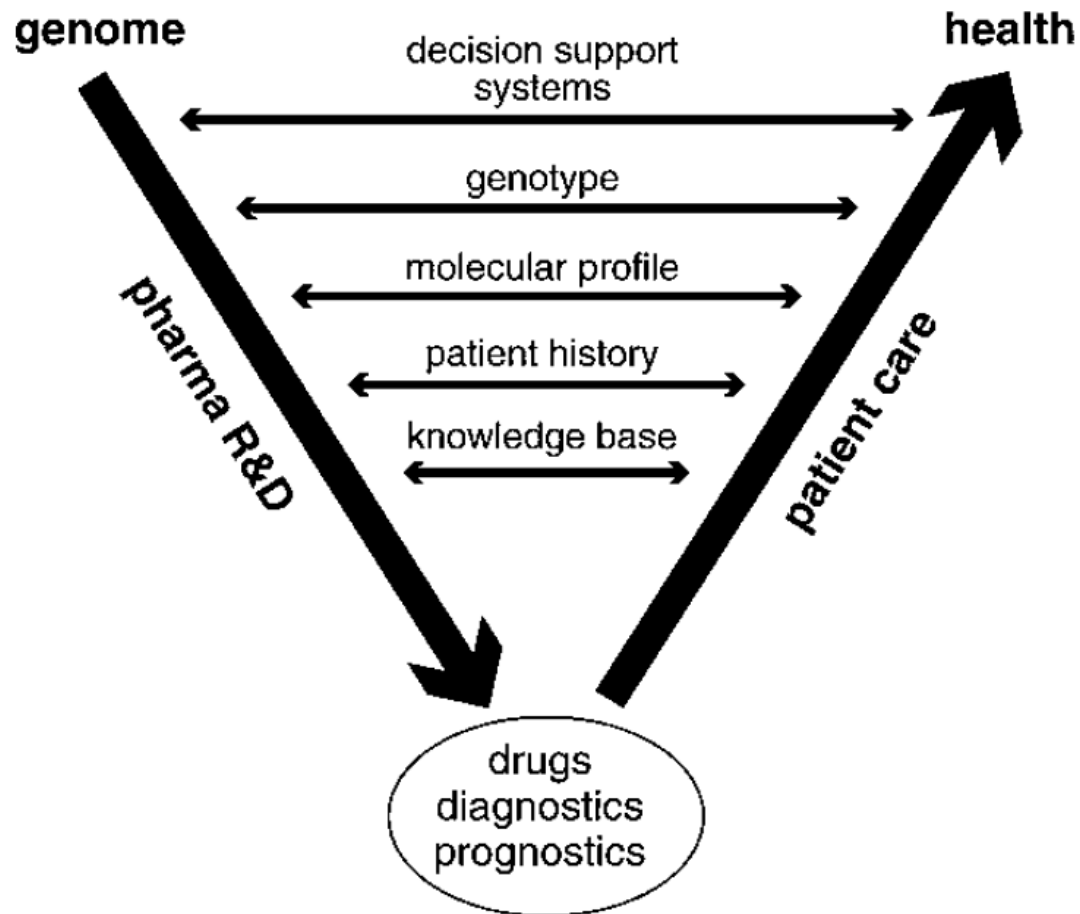
**Niko Beerenwinkel**

Department of Biosystems Science and Engineering, ETH Zurich

SIB Swiss Institute of Bioinformatics

Competence Center Personalized Medicine UZH/ETHZ

# The future of health care – 15 years ago

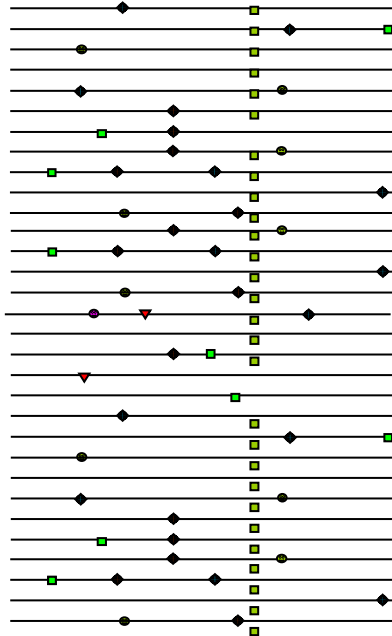
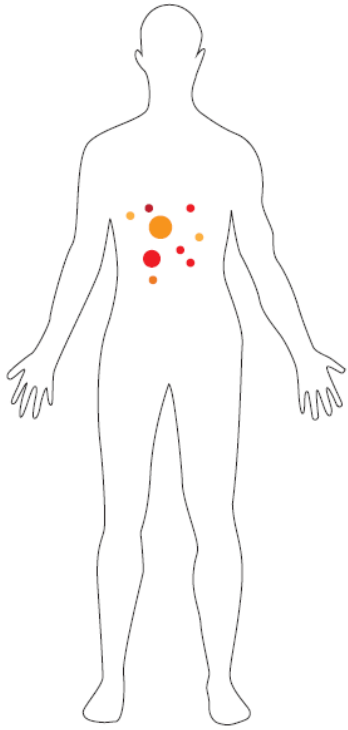


# Outline

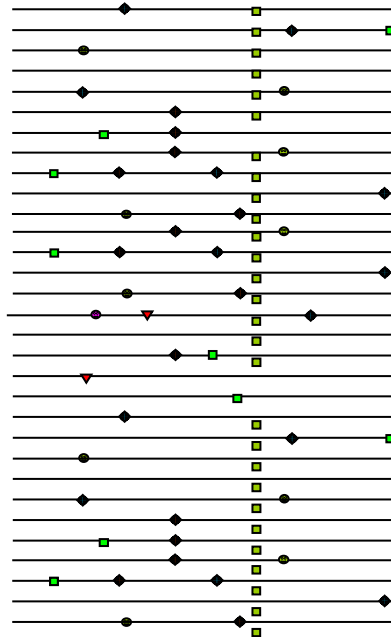
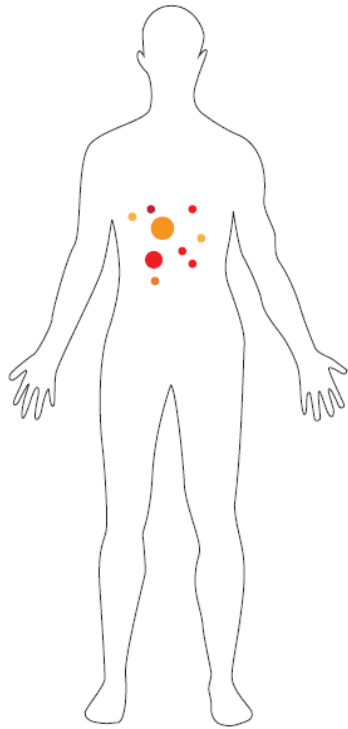
- Case study: Optimizing antiretroviral therapy
- Some (computational) challenges in Personalized Medicine
- Competence Center Personalized Medicine UZH/ETHZ

# Optimizing antiretroviral therapy

# HIV drug resistance



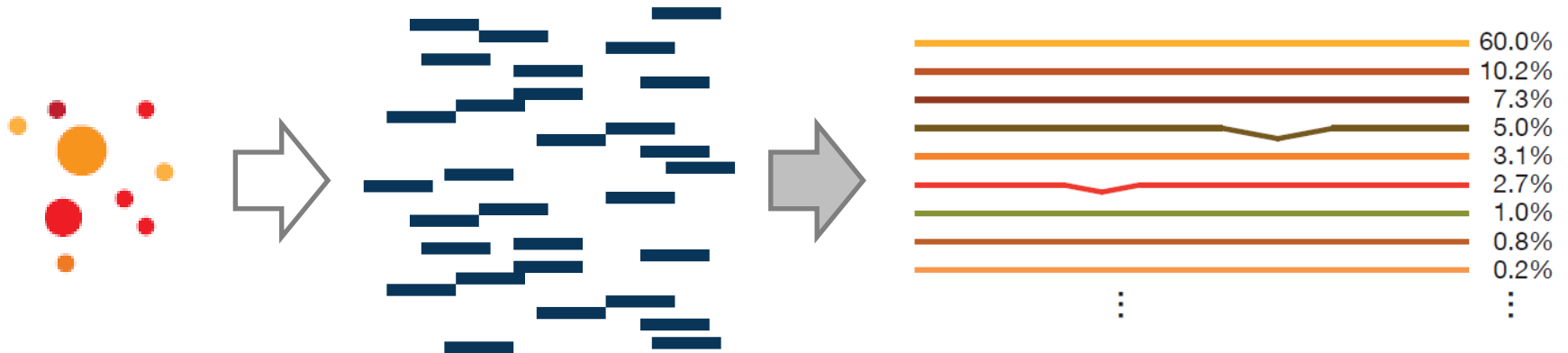
# HIV drug resistance – challenges



**Large genetic diversity**

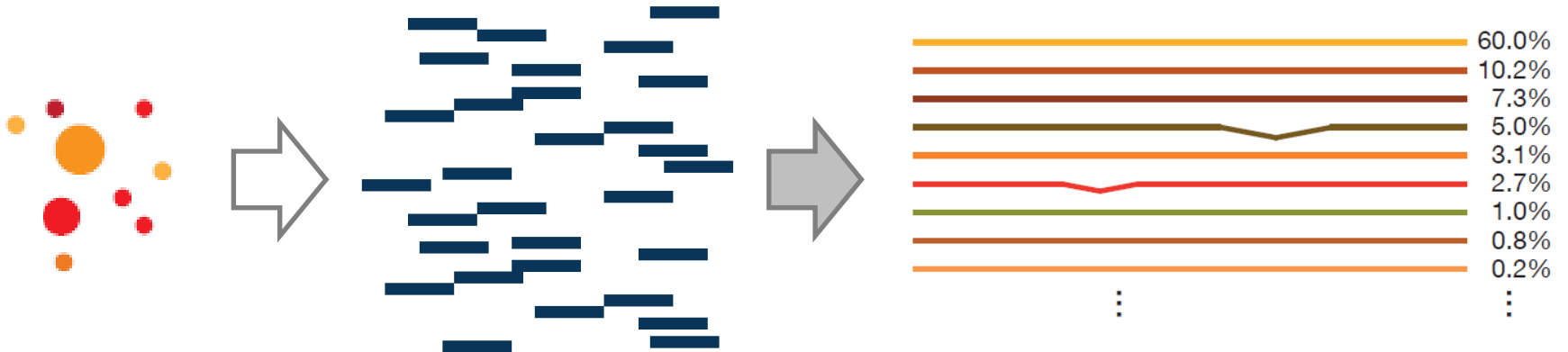
**Many drug combinations**

# Viral quasispecies assembly from NGS data



- Given many short error-prone reads from a mixed virus population, reconstruct all haplotype sequences and estimate their frequencies

# Challenges



1. Alignment uncertainty
2. Several confounding sources of variation
3. Short read length



↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓

A: CCTGAAATCACTCTATGGCAACGACCCATCGTCACAATAAAGATAGGG 60%

B: CCTCAAATCACTCTTTGGCAACGACGCATCGTCACAATATAGATAGGA 30%

C: CCTCAAATCTCTCTTTGGCACCGACCCATCGTCCCAATAAAGATAGGG 10%



1 CCTGAAATCACTCTATGGCA

2 GAAAACACTCTATGGCAACG

3 ATCACTCTTTGGCAAGGCCG

4 TCACTCTATGGCAACGACCC errors

5 CTCTTTGGCACCGACCCA

6 CTATGGTAACGACCCATCGT

7 TATGGCAACGACCATCGTC

8 ATGGCACGGACCCATCCCC

9 TGGCAACGACGCATCGTCAC

10 CAACGACCCATCGTCACAAT

11 CAACGACGCATCGTCACGAT

12 AACGACCC TTCGT CACAATA

13 CGACCCATCGTCTCAATAAA

14 GCATCGTCACAATATAGAGA

15 CATCGTCACAAAATAGATAG

16 TCGTCACAATAAAGATAGGG

17 TCACAATAAAGATGGGG

18 CCAATAAAGATAGGG

19 AATAAGGATGGGG

20 ATAGATAGGA

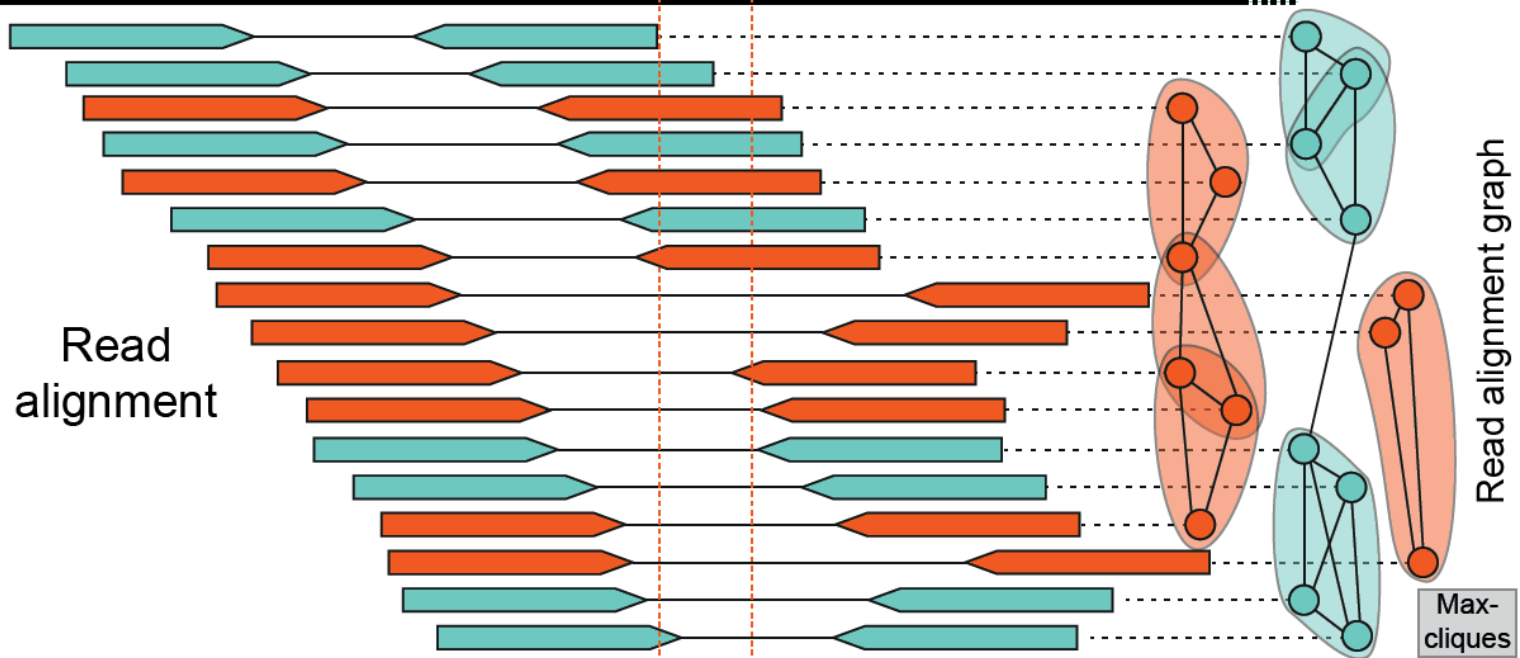
# Viral quasispecies assembly

Haplotype 1, strain A

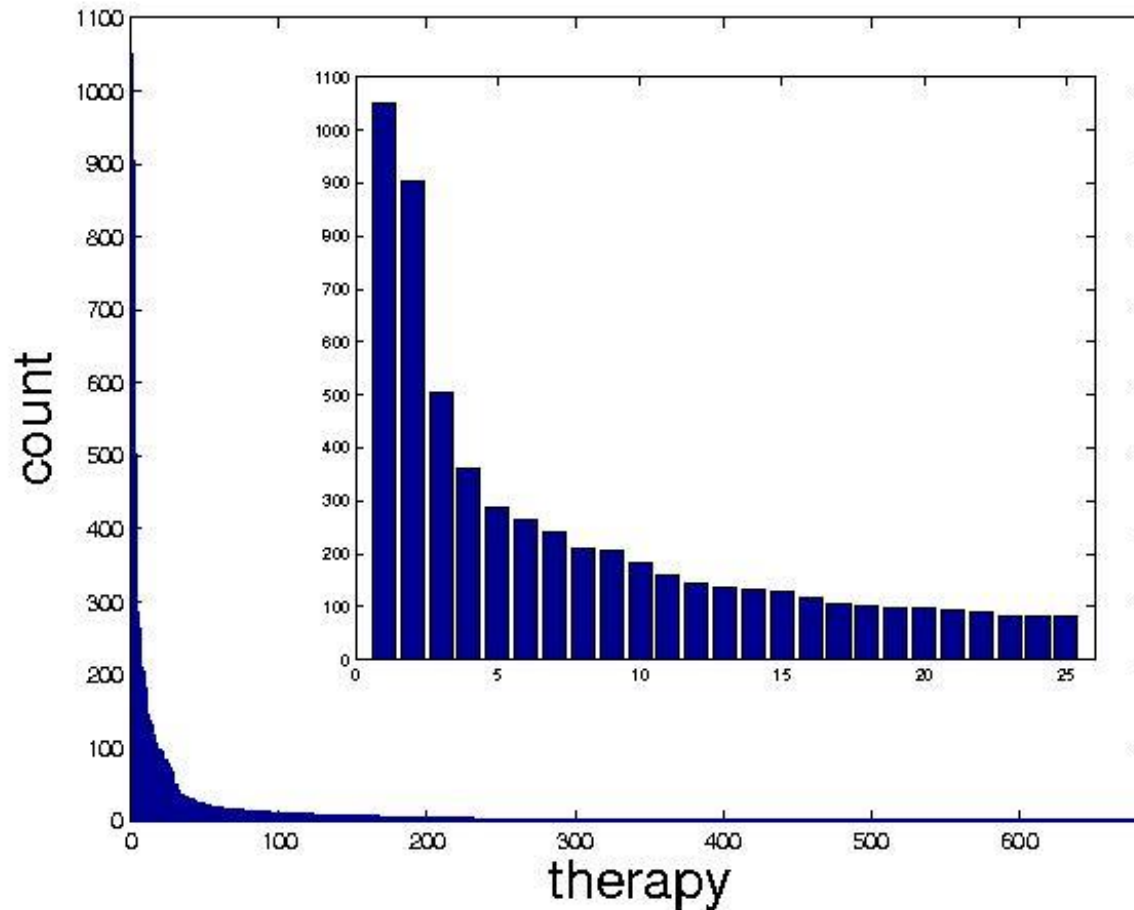
Haplotype 2, strain B

Haplotype 3, strain B with deletion

Reference genome

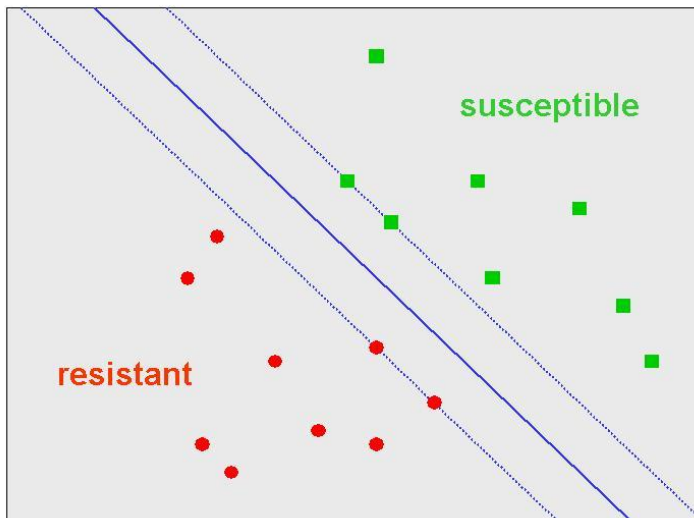


# Most drug combinations are rare



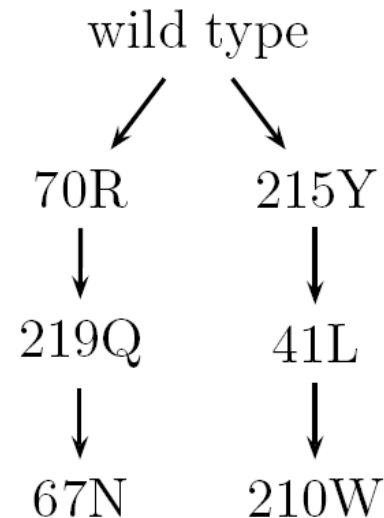
# Predicting HIV phenotypic drug resistance

Is the present virus resistant to drugs?



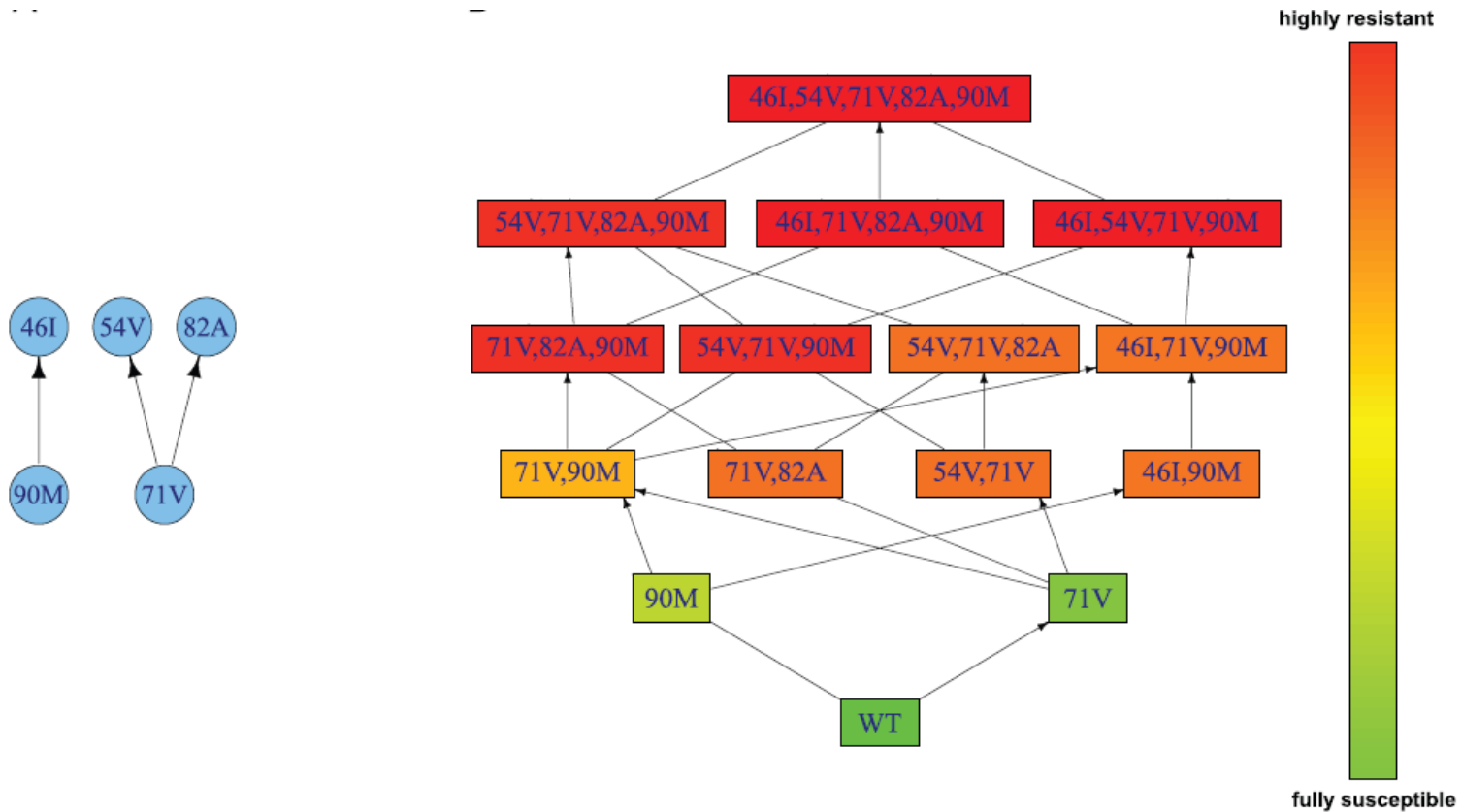
**Statistical learning**

How likely is it to evolve resistance?

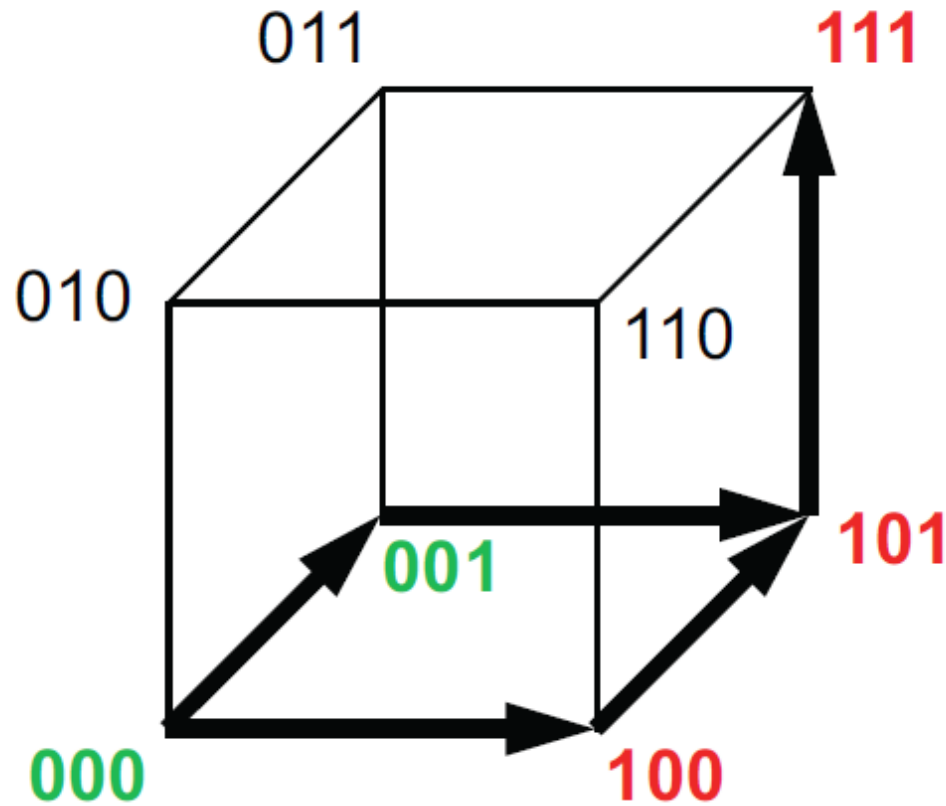


**Evolutionary modeling**

# A model of drug resistance development

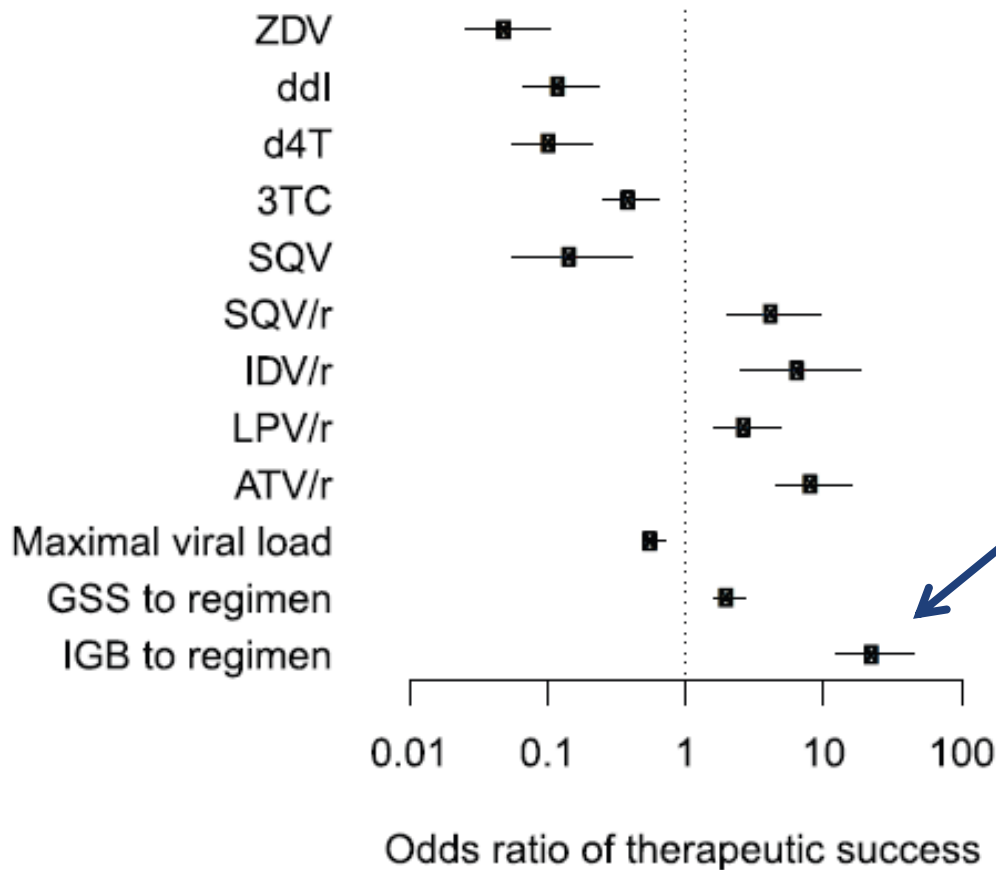
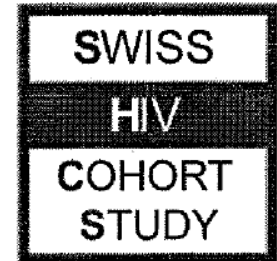


# The individualized genetic barrier



What is the probability of a **susceptible** virus to become **resistant**?

# Predicting treatment response

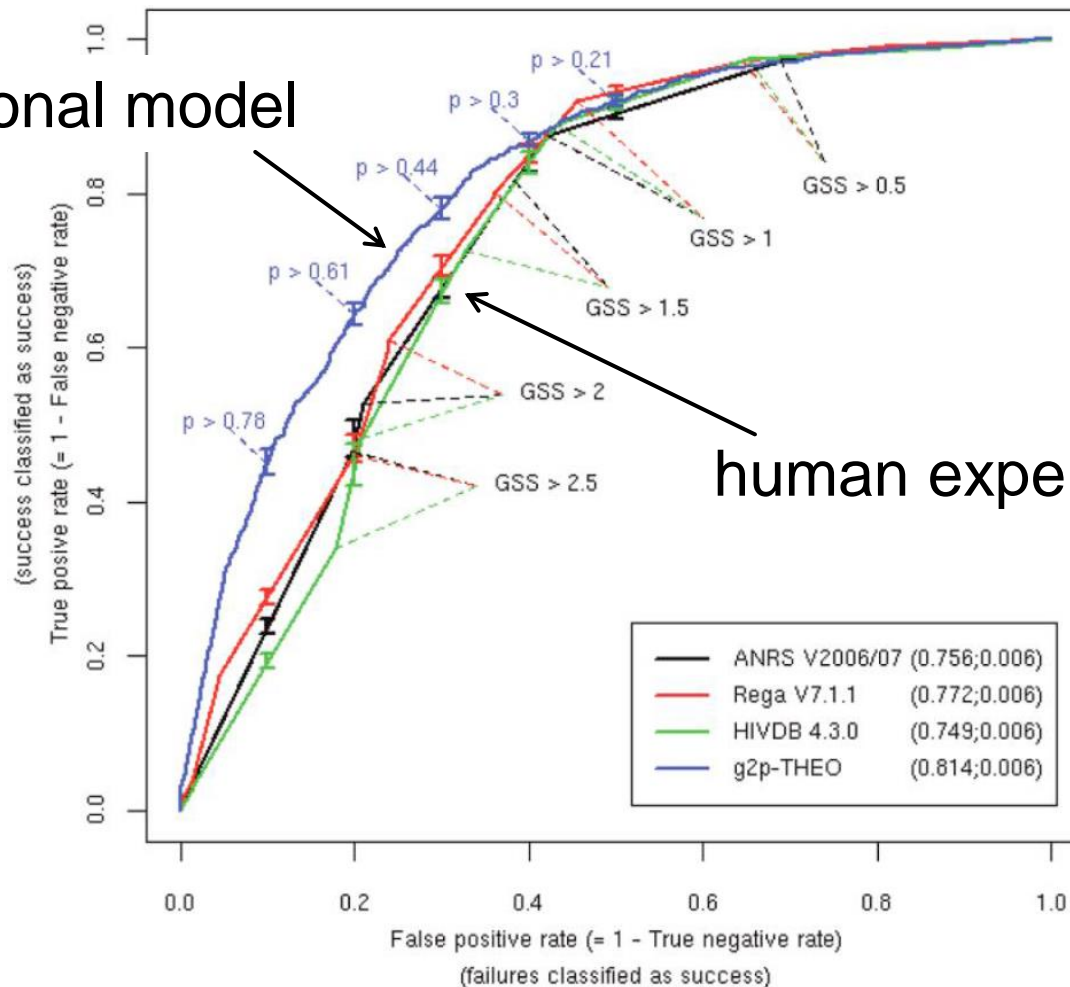


New predictor based on genomic data and evolutionary model



# Comparative performance

computational model



human expert panels



# Clinical decision support

ranked  
treatment  
options



No. of drugs <= - <- >

No. of pills per day <= - <- >

NRTIs: NNRTIs: PIs:

>= - <- > >= - <- > >= - <- >

ZDV= - NVP= - IDV= -  
 ddC= - DLV= - RTV= -  
 ddI= - EFV= - SQV= -  
 d4T= - NFV= exclude  
 3TC= - Reset APV= -  
 ABC= - Compute LPV= -  
 TDF= - ATV= -

Selected drug combinations:

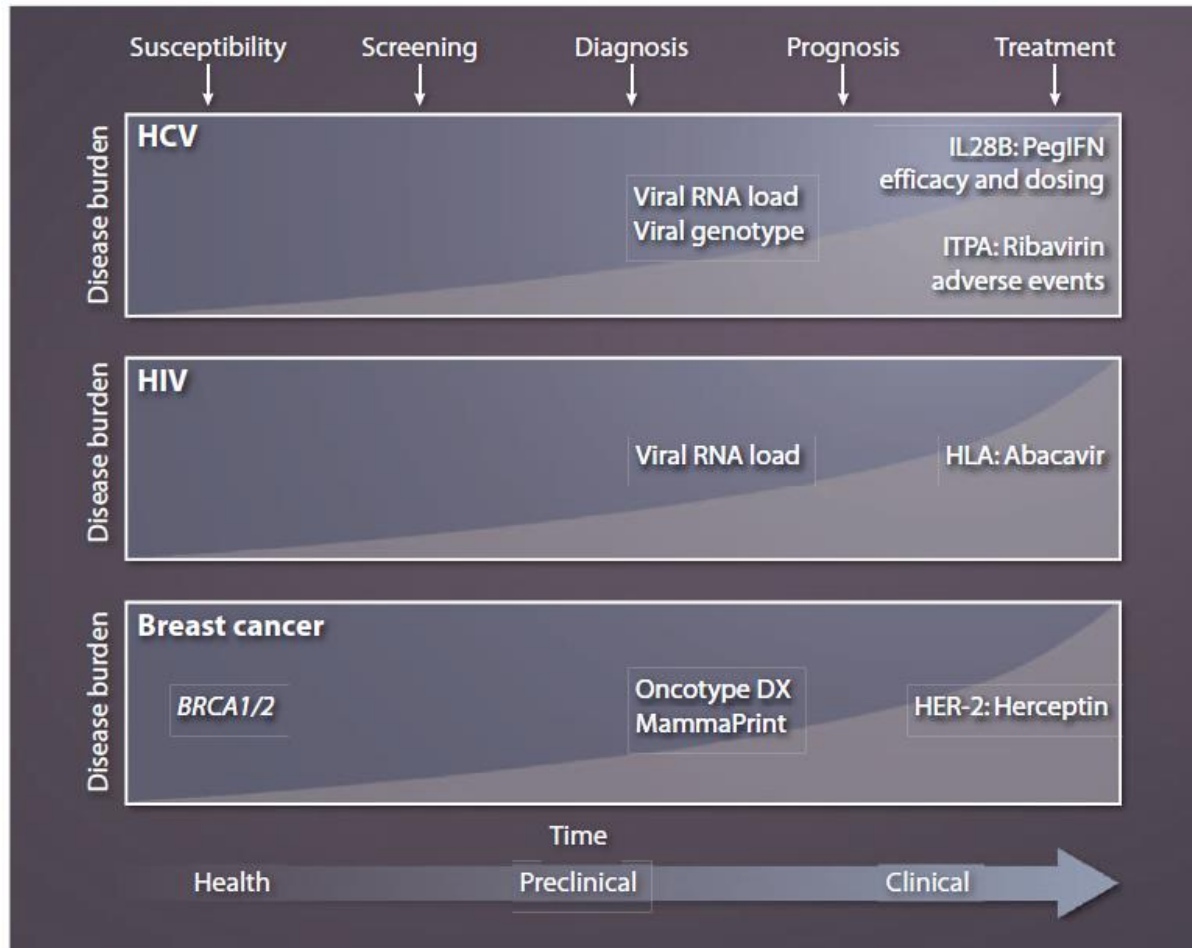
Success*	Regimen	Pills	Comment
0.72	SQV LPV	18	LPV/r(6) SQV/r(10)
0.71	APV LPV	22	LPV/r(6) APV(16)
0.66	d4T TDF LPV	9	d4T(2) TDF(1) LPV/r(6)
0.64	ZDV ddI LPV	9	ZDV(2) ddI(1) LPV/r(6)

\*) PREDICTED probability of virological success

Histogramm of all (all selected) therapies

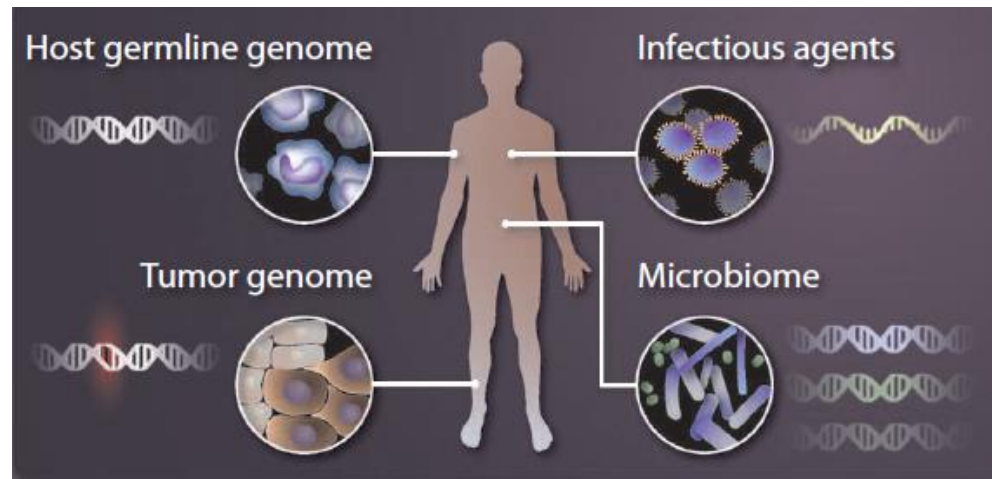
Probability of virological success over 24+ weeks

# Genomic medicine in action



# Goals of genome-based personalized medicine

- Improved, individualized
  - **diagnostics** and **prognostics**: earlier, more precise, broad molecular profiling

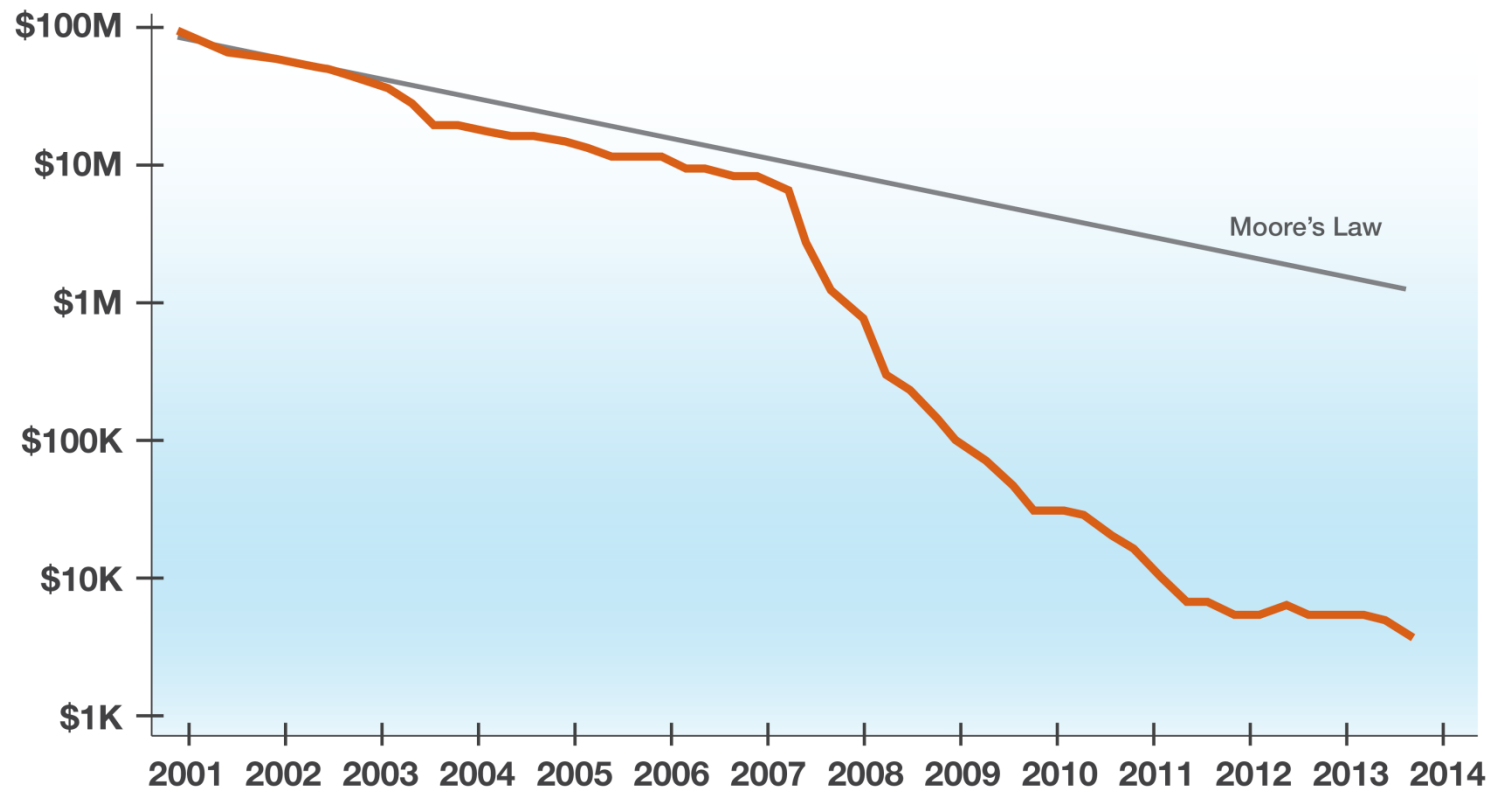


McCarthy et al. 2013

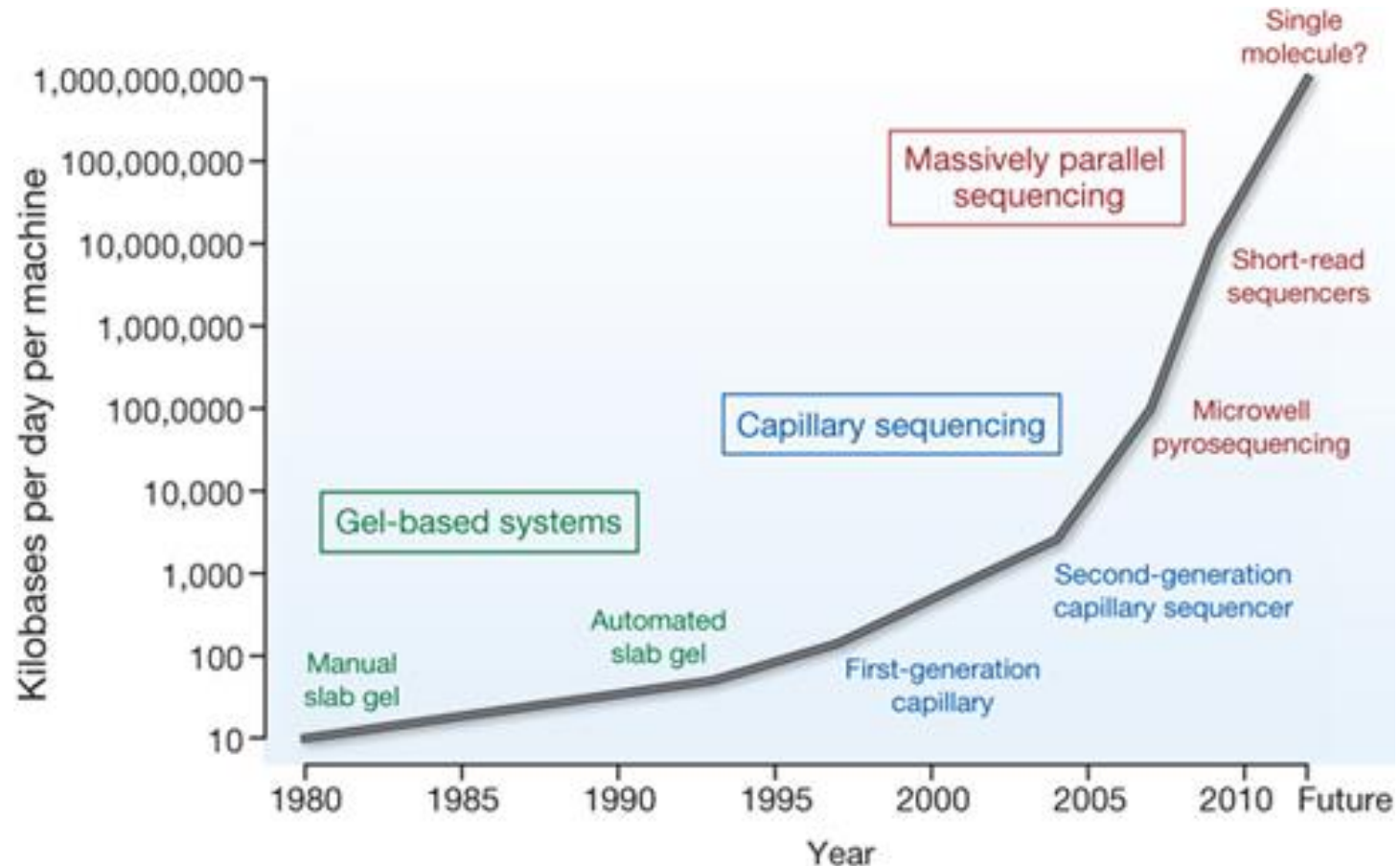
- **treatments**: targeting specific disease markers, higher likelihood of response, enhanced drug safety, optimized dosing

# **Some challenges in Personalized Medicine**

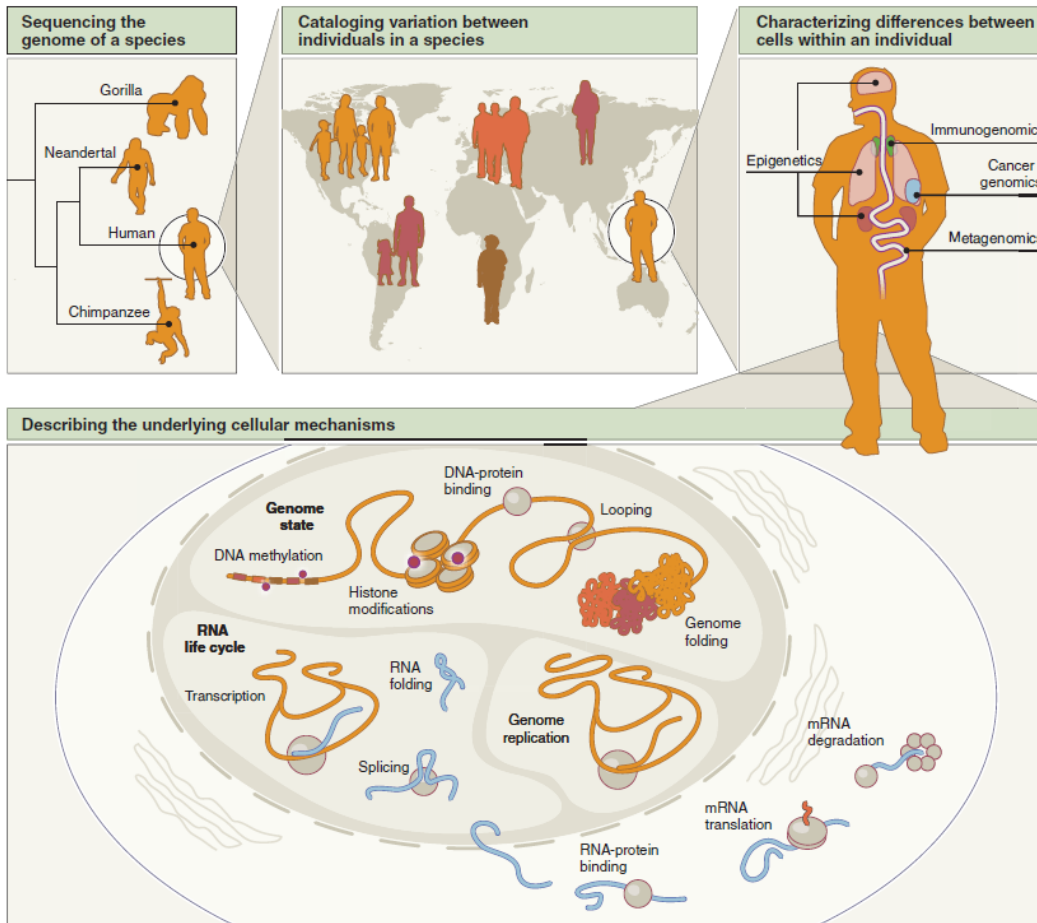
# Cost per genome



# High throughput due to parallelization

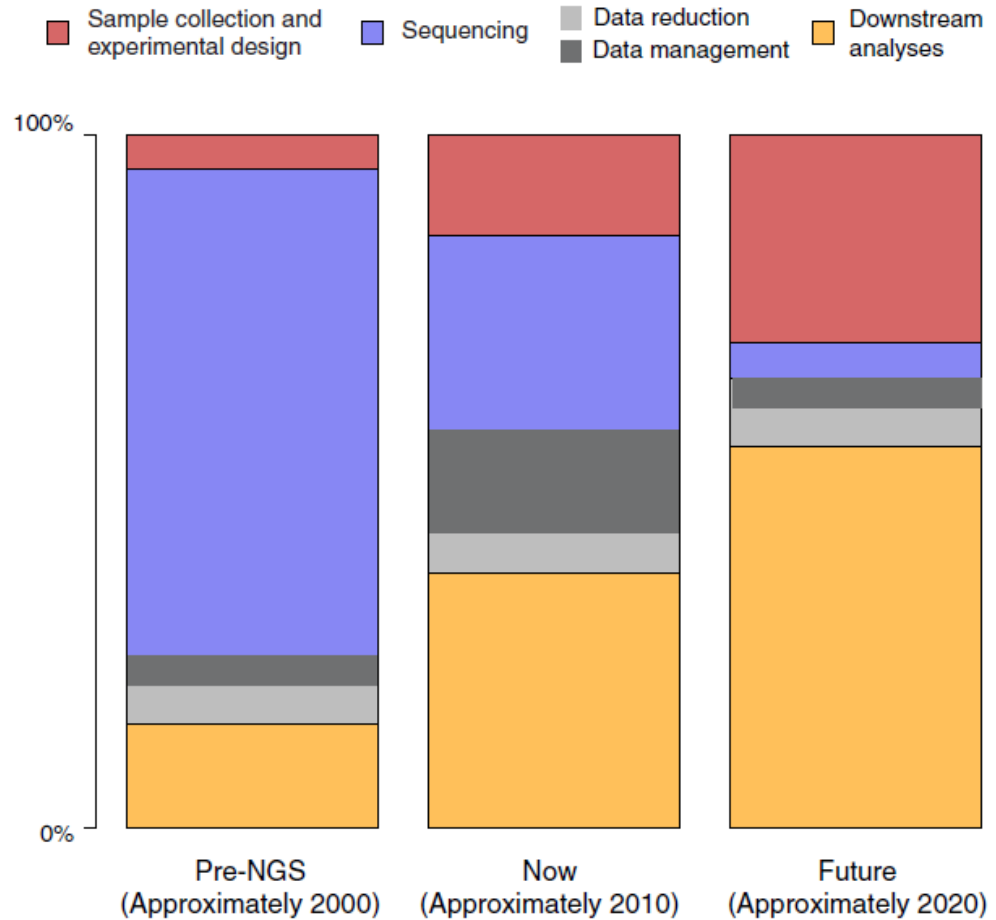
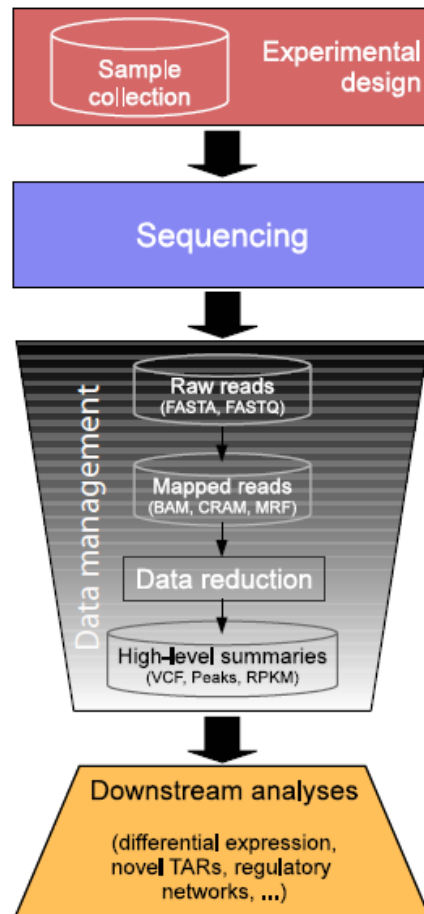


# Applications of next-generation sequencing



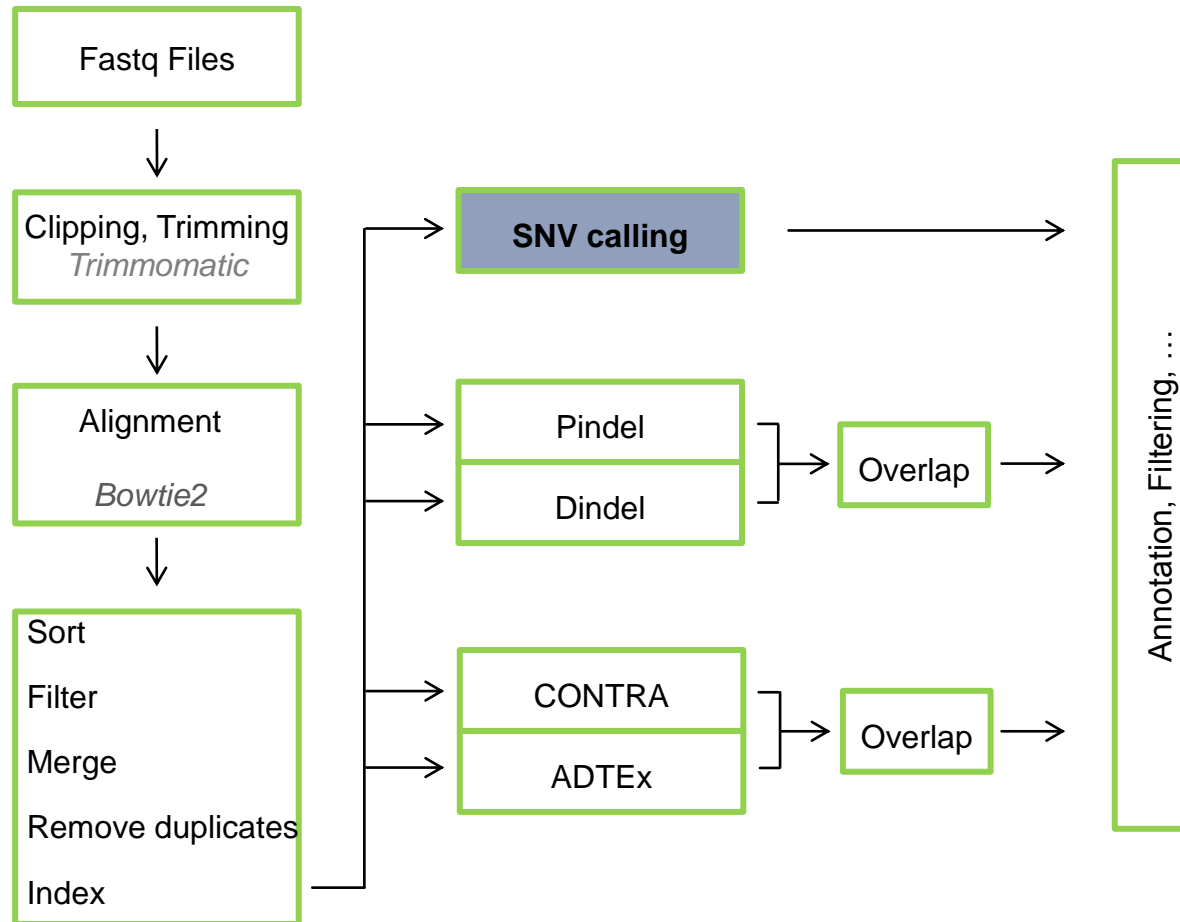
Method	Sequencing to determine:
DNA-Seq	A genome sequence
Targeted DNA-Seq	A subset of a genome (for example, an exome)
Methyl-Seq	Sites of DNA methylation, genome-wide
Targeted methyl-Seq	DNA methylation in a subset of the genome
DNase-Seq, Sono-Seq and FAIRE-Seq	Active regulatory chromatin (that is, nucleosome-depleted)
MAINE-Seq	Histone-bound DNA (nucleosome positioning)
ChIP-Seq	Protein-DNA interactions (using chromatin immunoprecipitation)
RIP-Seq, CLIP-Seq, HITS-CLIP	Protein-RNA interactions
RNA-Seq	RNA (that is, the transcriptome)
FRT-Seq	Amplification-free, strand-specific transcriptome sequencing
NET-Seq	Nascent transcription
Hi-C	Three-dimensional genome structure
Chia-PET	Long-range interactions mediated by a protein
Ribo-Seq	Ribosome-protected mRNA fragments (that is, active translation)
TRAP	Genetically targeted purification of polysomal mRNAs
PARS	Parallel analysis of RNA structure
Synthetic saturation mutagenesis	Functional consequences of genetic variation
Immuno-Seq	The B-cell and T-cell repertoires
Deep protein mutagenesis	Protein binding activity of synthetic peptide libraries or variants
PhIT-Seq	Relative fitness of cells containing disruptive insertions in diverse genes

# The real cost of sequencing

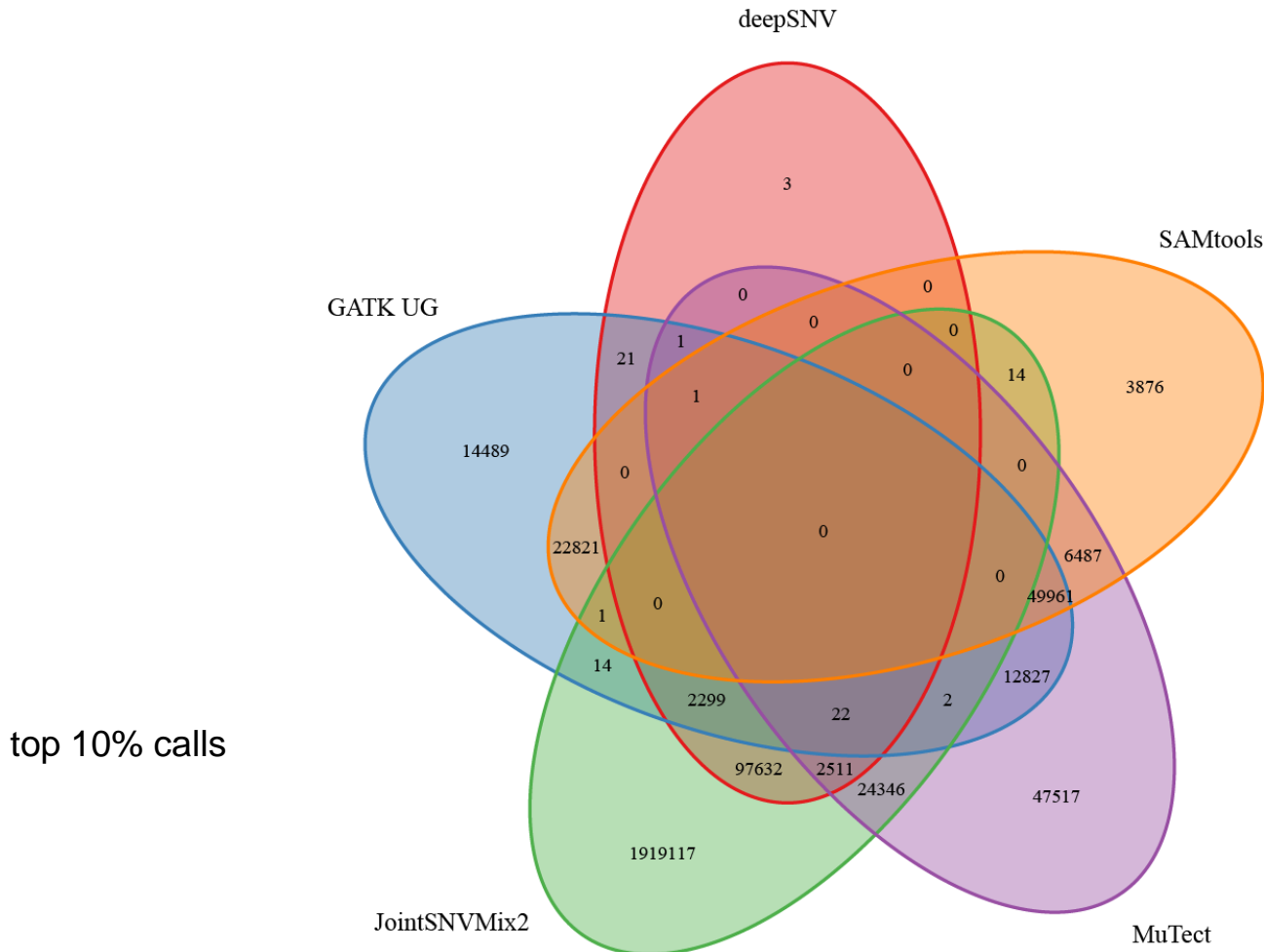




# Exome sequencing data analysis pipeline



# Five common SNV callers on a cancer exome

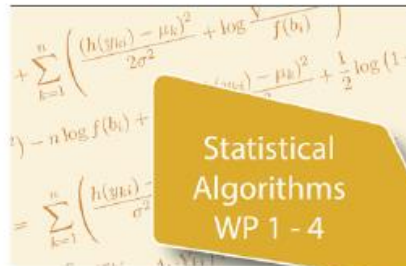




# Developing open-source software for PM

## SOUND

Statistical Multi-Omics Understanding


$$+ \sum_{k=1}^n \left( \frac{(h(y_{ki}) - \mu_k)^2}{2\sigma^2} + \log \frac{1}{f(b_i)} \right)$$
$$- n \log f(b_i) +$$
$$= \sum_{k=1}^n \left( \frac{h(y_{ki}) - \mu_k}{\sigma^2} \right)$$

Statistical  
Algorithms  
WP 1 - 4



Methodological  
and Computational  
Integration  
WP 8 - 12

Bioinformatic Tools  
for R&D in  
Personalized  
Medicine



Clinician  
Scientists



Statistical Omics  
Community



# **Competence Center Personalized Medicine UZH/ETHZ (CC-PM)**

## Hochschulmedizin Zurich

## Competence Center Personalized Medicine (CC-PM)



- 25 PIs from USZ/UZH and ETHZ
- PhD program Molecular and Translational Biomedicine (~50 students)

## CC-PM flagship projects

- *Towards individualized prevention and therapy in hereditary cancer diseases*
- *Brown fat tissue – personalized strategies to achieve weight loss*
- *Metastatic ccRCC 1 – Personalising prognosis and therapy of metastatic renal carcinoma*
- *Digital biobanking*
- *The intra-tumor heterogeneity census project – decoding mechanisms of intra-tumor heterogeneity for precision medicine*



# Technology for personalized medicine research

Molecular and clinical data management, analysis, and integration

**PM-ICT**

Lab automation and robotics tools and advanced cell systems

**Theragnostics  
Discovery**

Tissue and serum biobanking

**Biobank**

Hochschulmedizin Zurich

Competence Center Personalized Medicine (CC-PM)

*joint PM research projects*

PM-ICT  
Unit

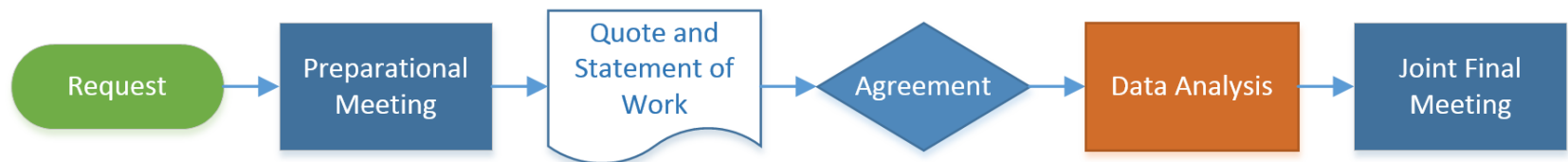
TD  
Unit

Biobanking  
(UZH/USZ)

**nexus (ETHZ)**

...

- Bioinformatics and statistics services
  - genome analysis (WGS, WES, targeted sequencing)
  - transcriptome analysis (RNA-seq)
  - proteome analysis
  - association studies
  - image analysis for digital pathology
  - ...
- Consulting



- Education: courses, internships, etc.

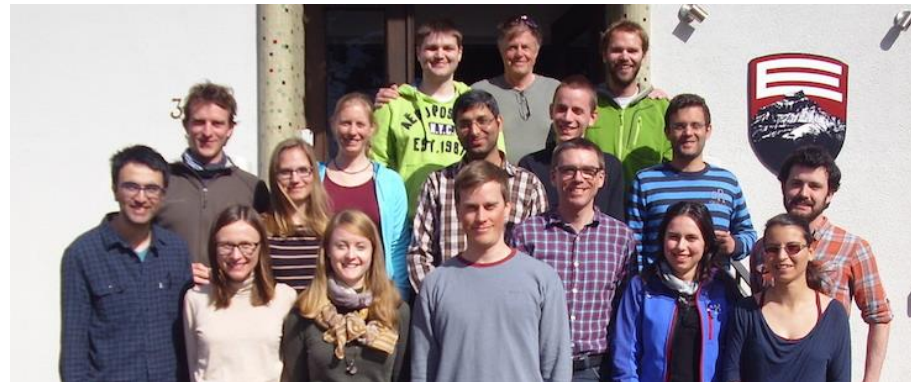
# Acknowledgements

## CGB, D-BSSE, ETH Zurich

- Simona Constantinescu
- Madeline Diekmann
- Christos Dimitrakopoulos
- Monica Golumbeanu
- Ariane Hofmann
- Katharina Jahn
- Vinay Jethava
- Jack Kuipers
- Hesam Montazeri
- Caroline Perraudin
- Susana Posada Cespedes
- Hans-Joachim Ruscheweyh
- Fabian Schmich
- David Seifert
- Jochen Singer
- Ewa Szczurek

## Collaborators

- Huldrych Günthard, USZ
- Willy Krek, ETH Zurich, NEXUS
- Holger Moch, USZ, CC-PM
- Daniel Stekhoven, NEXUS



**Funding:** ETH Zurich, SNSF, Swiss HIV Cohort Study, SystemsX.ch, ERASysAPP, Swiss Cancer League, EU H2020, ERC Synergy

[www.cbg.ethz.ch](http://www.cbg.ethz.ch)