Universitatea de Medicina și Farmacie "Victor Babeș" Timișoara

PROGRAMUL DE COOPERARE ELVEȚIANO-ROMÂN
SWISS-ROMANIAN COOPERATION PROGRAMME

University of Zurich

ROMANIAN-SWISS RESEARCH PROGRAMME 2013-2015

FROM CHRONIC INFLAMMATORY DERMATOSES TO CUTANEOUS LYMPHOMA

FROM CHRONIC INFLAMMATORY DERMATOSES TO CUTANEOUS LYMPHOMA

Prof. Dr. Michael Baudis

Professor of Bioinformatics

Department of Molecular Life Sciences

Swiss Institute of Bioinformatics

Universität Zürich

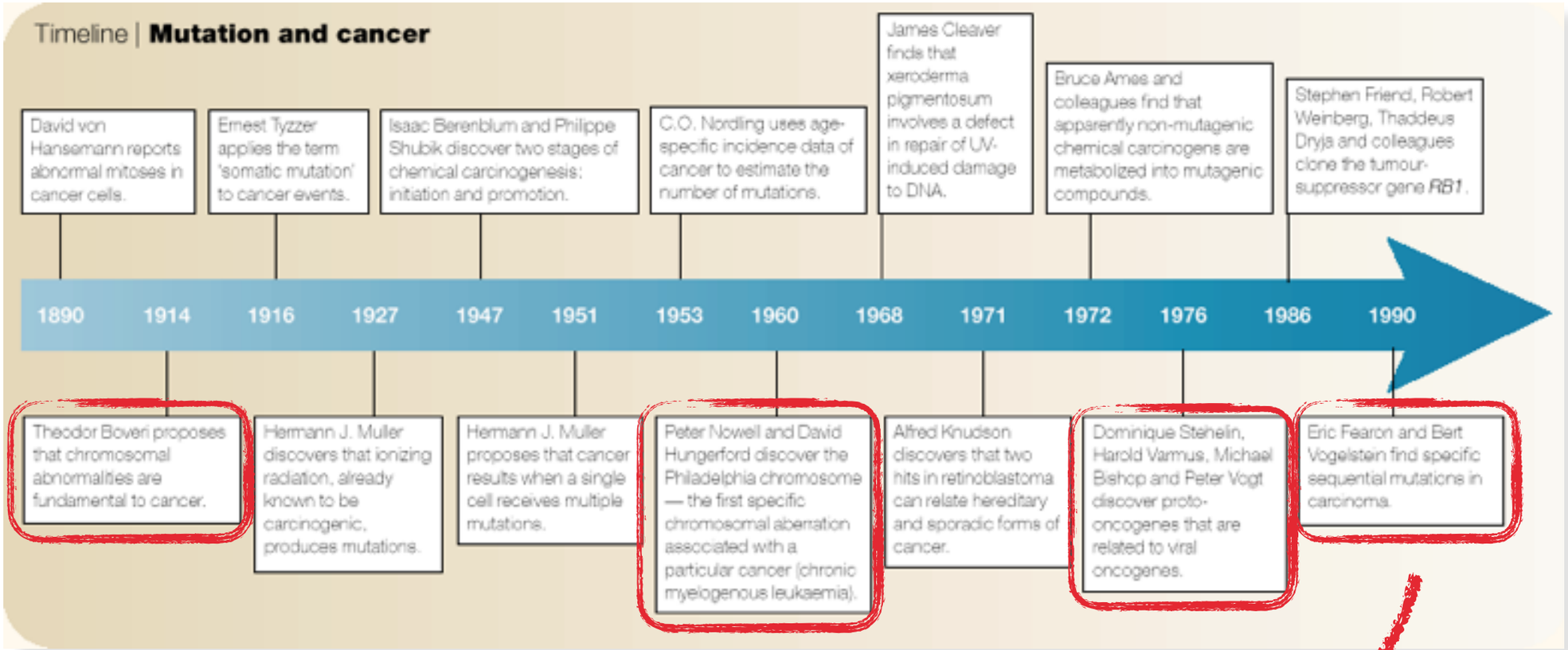
Switzerland

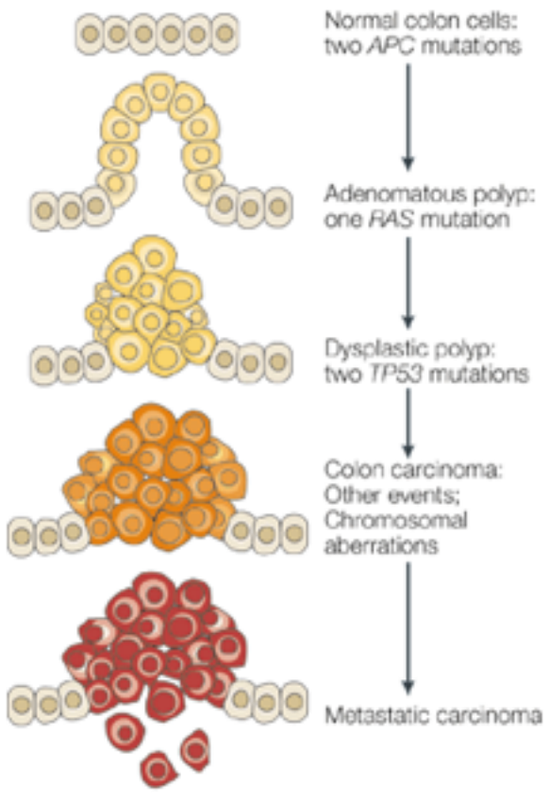
Prof. Dr. Caius S. Solovan

Professor of Dermatology

President, Romanian Society of Dermatopathology
University of Medicine and Farmacy "Victor Babes" Timisoara

Romania

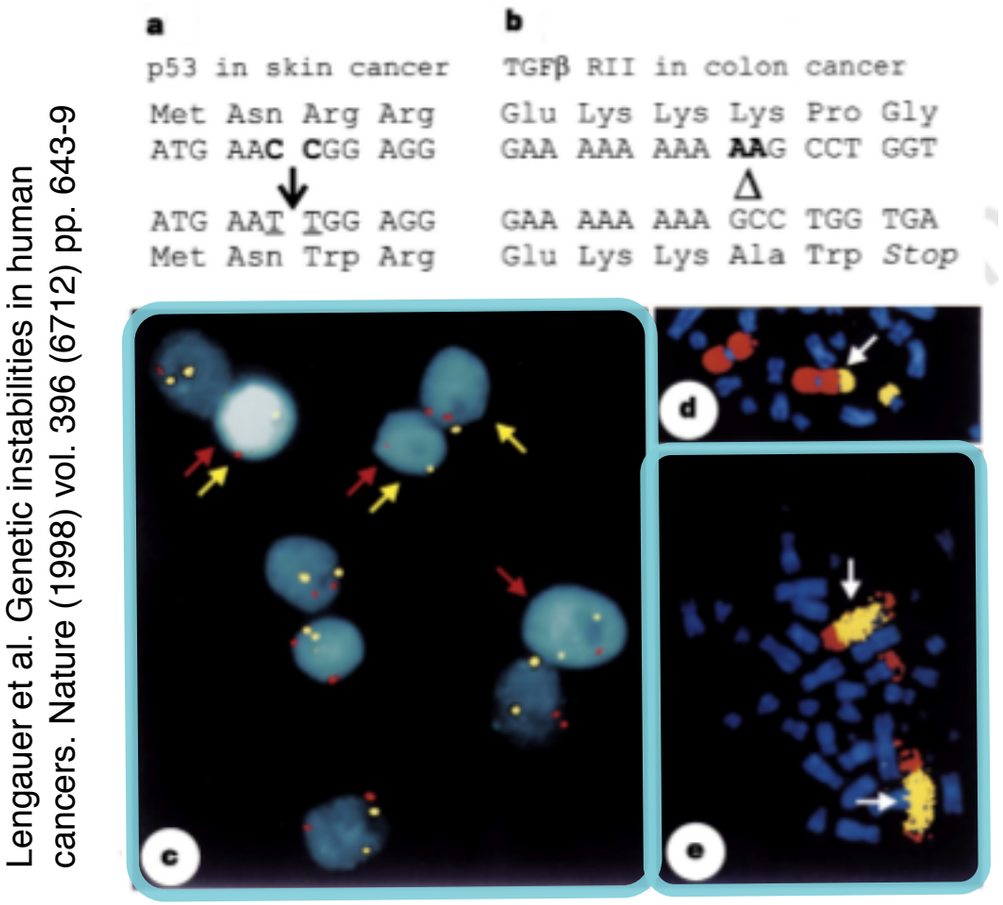




Cancers are based on acquired and inherited genomic mutations

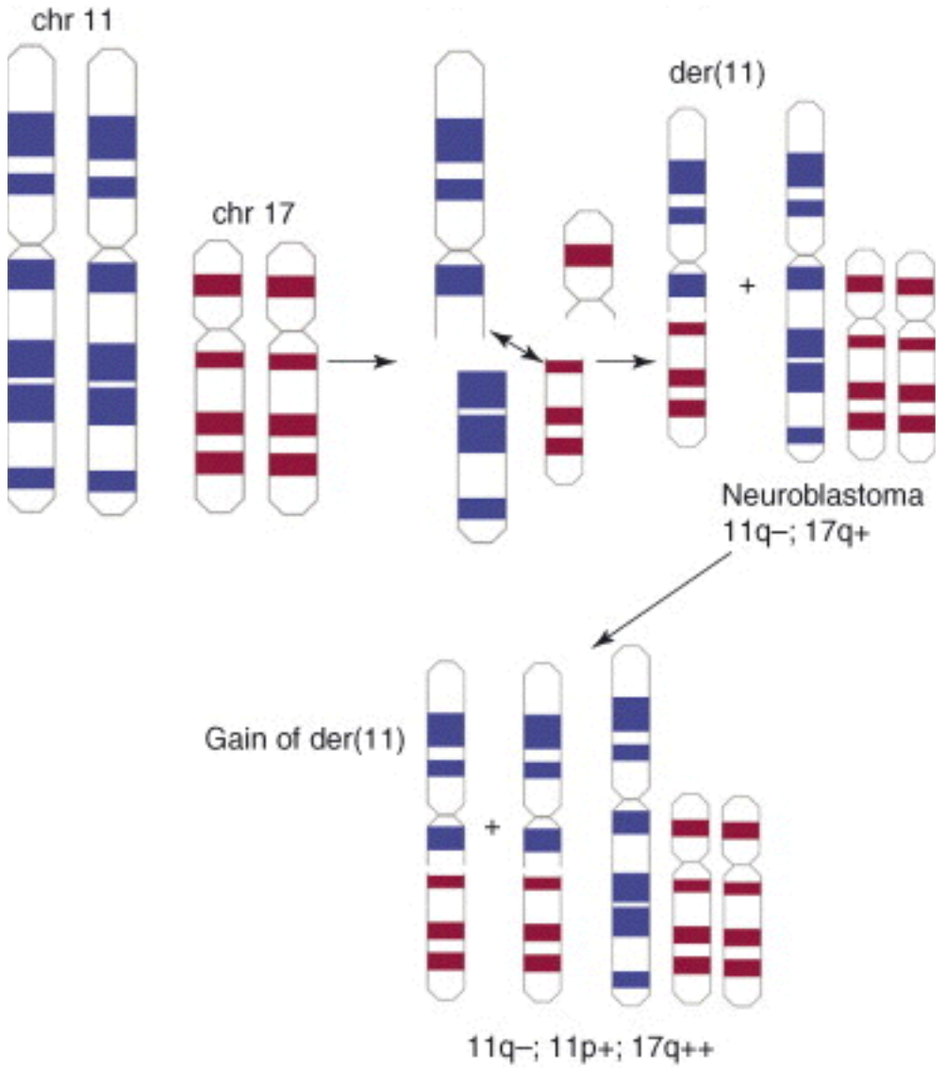
Knudson, A. G. (2001). Two genetic hits (more or less) to cancer. Nature Reviews Cancer, 1(2), 157–162.

Mutations & genomic rearrangements in cancer



Lengauer et al. Genetic instabilities in human cancers. Nature (1998) vol. 396 (6712) pp. 643-9

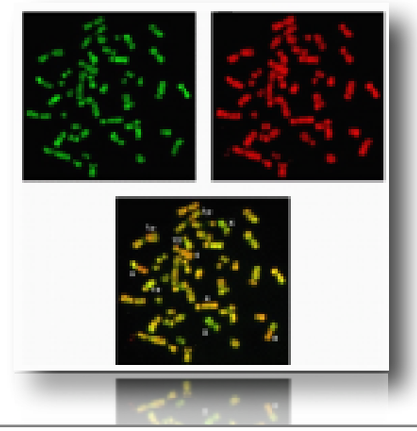
a. small mutation (di-pyrimidine exchange at p53 in Xeroderma pigmentosum patient)
b. two-base deletion in *TGFB* in a colorectal cancer patient with mismatch repair deficiency
c. chromosomal losses (FISH; red=3, yellow=12) in CRC
d. t(1;17) in neuroblastoma, whole-chromosomal painting
e. *MYCN* gene amplification (multiple copies inserted into chromosome 1 derived marker)

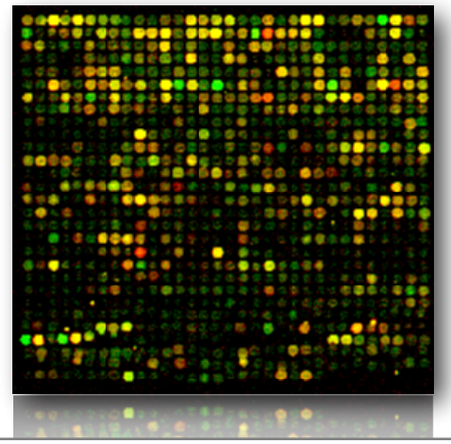


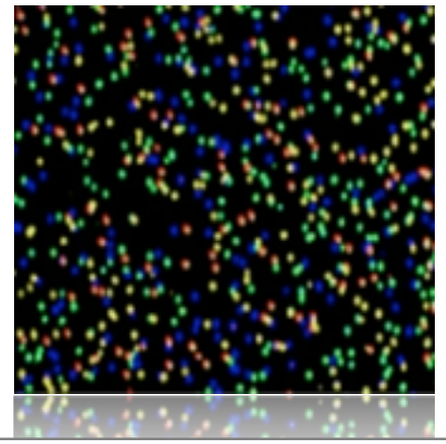
Generation of copy number imbalances in cancer through imbalanced cytogenetic rearrangements - partial deletion of 11q, gain of 11pterq21 and 2 addl. copies of 17q

RL Stallings: Are chromosomal imbalances important in cancer? Volume 23, Issue 6, p278–283, 2007

Whole Genome Screening in Cancer

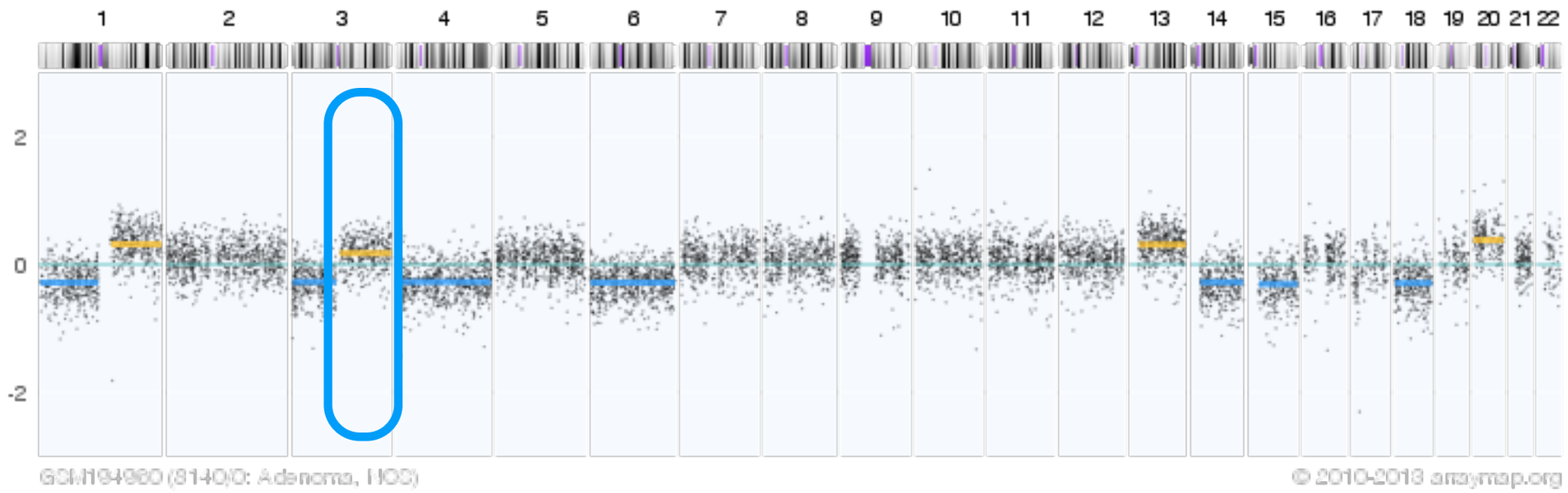




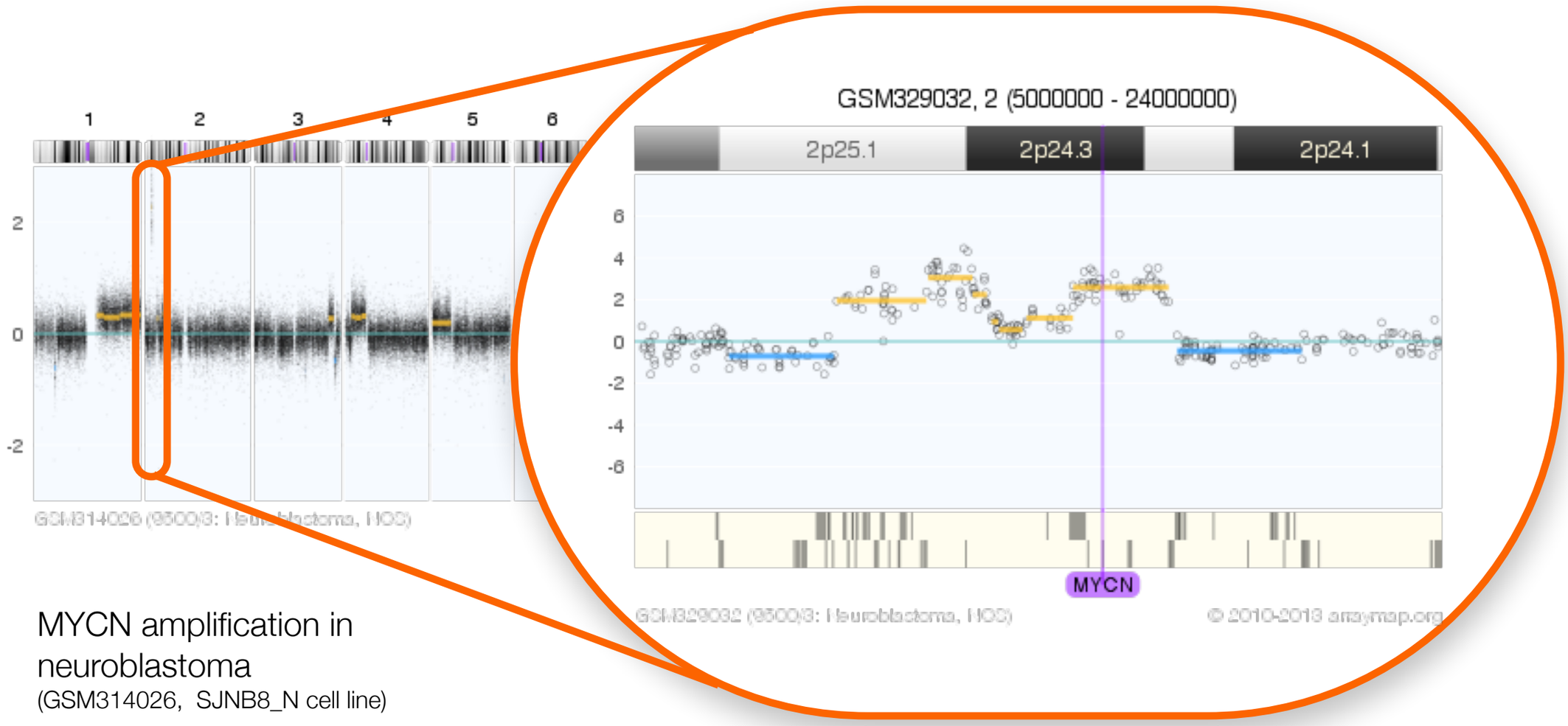


	chromosomal CGH	genomic arrays	"NGS" genome sequencing
1st application report	**1992**	**1997**	**2010**
source	DNA (paraffin, micro-dissected ...)	DNA (paraffin, micro-dissected ...)	DNA (paraffin, micro-dissected ...)
main source problems	mixed/degraded source tissue	mixed/degraded source tissue	mixed/degraded source tissue
resolution	chromosomal bands = few megabases	mostly in the 100kb range, but tiling possible	single bases
target identification	surrogate (position)	"semidirect" (segmentation spanning probes)	direct quantitative and qualitative
structural	no	depending on type	yes
available data	>24,000 cases (57%) through **Progenetix**	raw data repositories (GEO, EMBL, SMD), **arrayMap**	limited (few entities, study consortia…); variant call data in dbgap, clinvar …
predominant data format	ISCN = static	raw => depends on bioinformatics	mostly selected variant calls

Genomic arrays: Many probes + bioinformatics determine copy number aberrations



Gain of chromosome arm 3q in colorectal carcinoma



MYCN amplification in neuroblastoma
(GSM314026, SJNB8_N cell line)

low level/**high level** copy number alterations (CNAs)

arrayMap

Bioinformatics 2001

progenetix

arrayMap

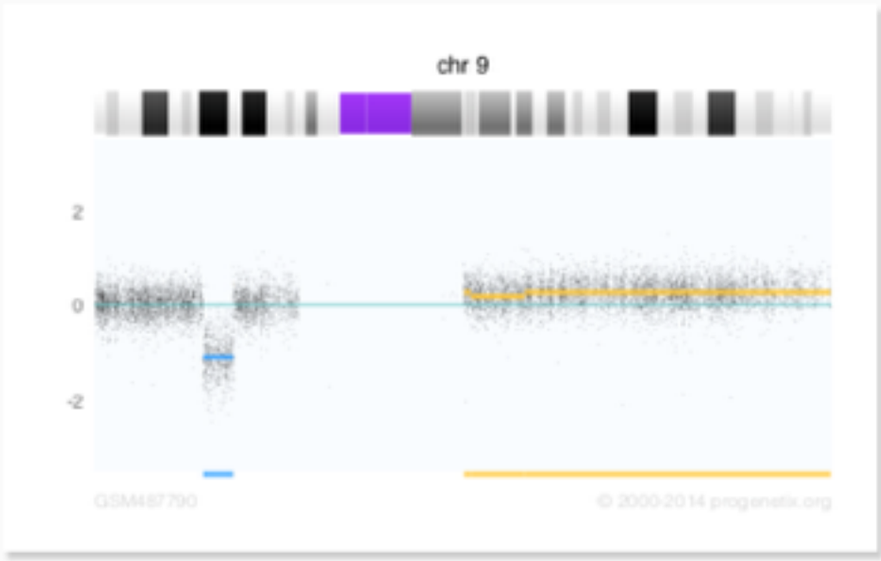
PLOS ONE 2012

	progenetix	arrayMap
techniques	cCGH, aCGH, WES, WGS	aCGH (+?)
scope	**sample** (e.g. combination of several experiments); **literature** tracking	**experiment**
content	**>31000** samples **>2700** publications	**>60000** arrays
raw data presentation	no (link to sources if available)	yes (raw, log2, segmentation if available)
per sample re-analysis	no; supervised result (mostly as provided through publication)	yes (re-segmentation, thresholding, size filters ...)
final data	annotated/interpreted CN status for GP and cytogenetic regions	unsupervised CN status for GP and cytogenetic regions
main purposes	• Distribution of CNA target regions in most tumor types (>350 ICD-O) • Cancer classification	• Gene specific hits • Genome feature correlation (fragile sites ...)

visualizing cancer genome array data @ arraymap.org

arrayMap is a curated reference database and bioinformatics resource targeting copy number profiling data in human cancer. The arrayMap database provides an entry point for meta-analysis and systems level data integration of high-resolution oncogenomic CNA data. The current data reflects:

- 65042 genomic copy number arrays
- 986 experimental series
- 333 array platforms
- ICD-O 253 ICD-O cancer entities
- 716 publications (Pubmed entries)



For the majority of the samples, probe level visualization as well as customized data representation facilitate gene level and genome wide data review. Results from multi-case selections can be connected to downstream data analysis and visualization tools, as we provide through our Progenetix project.

arrayMap is developed by the group "Theoretical Cytogenetics and Oncogenomics" at the Institute of Molecular Life Sciences of the University of Zurich.

BRAIN TUMOURS	5791 samples ↗	[?]
BREAST CANCER	8594 samples ↗	[?]
COLORECTAL CANCER	3470 samples ↗	[?]
PROSTATE CANCER	1366 samples ↗	[?]
STOMACH CANCER	1457 samples ↗	[?]

ARRAYMAP NEWS

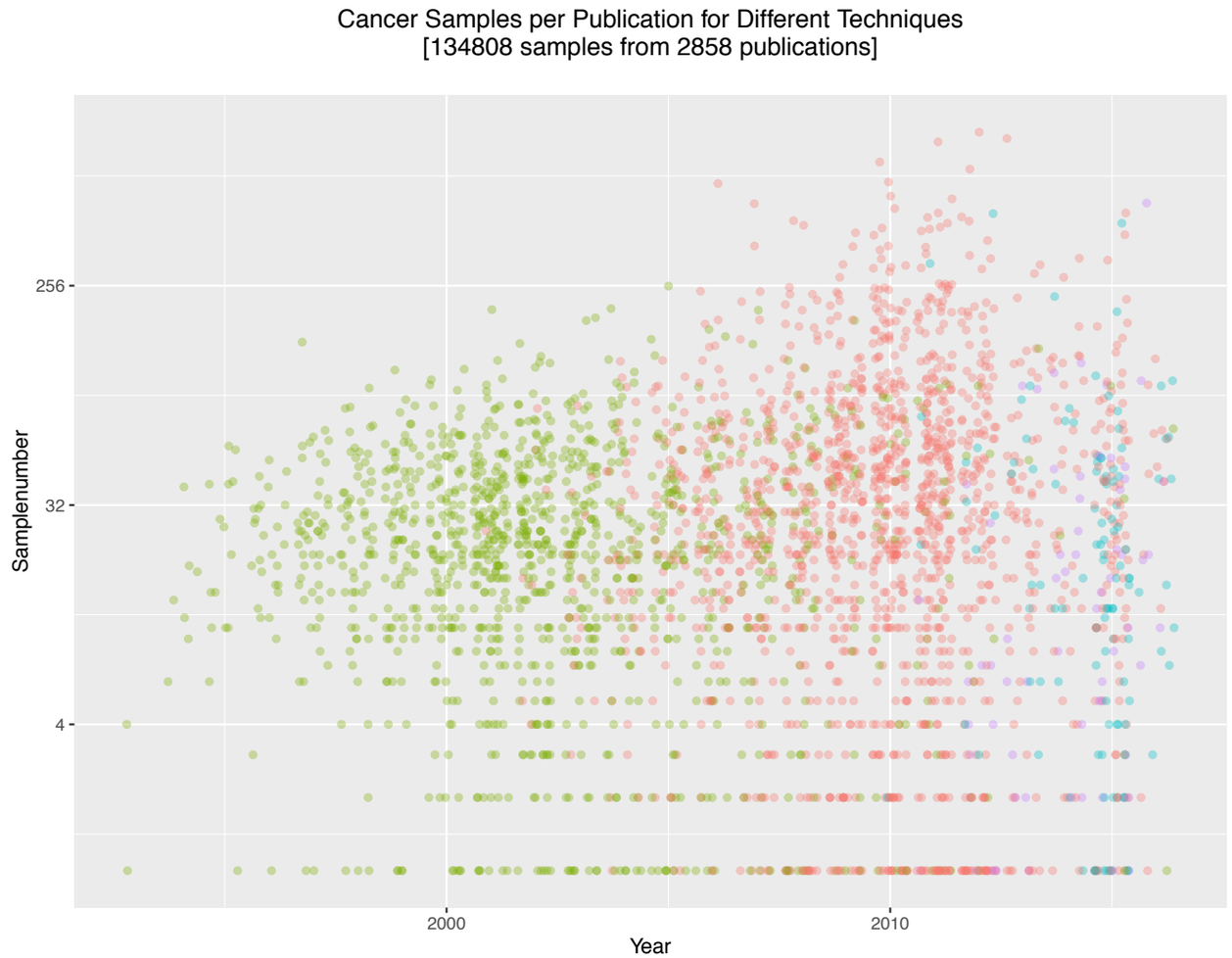
2016-04-11: Sorting cancer subset tables

2015-03-23: SIB Profile 2015

More news ...

Feel free to use the data and tools for academic research projects and other applications. If more support and/or custom analysis is needed, please contact Michael Baudis regarding a collaborative project or a special license.

MOLECULAR CYTOGENETICS & SEQUENCING STUDIES FOR WHOLE GENOME PROFILING



SURVEY OF >2800 CANCER GENOME ARTICLES . . .

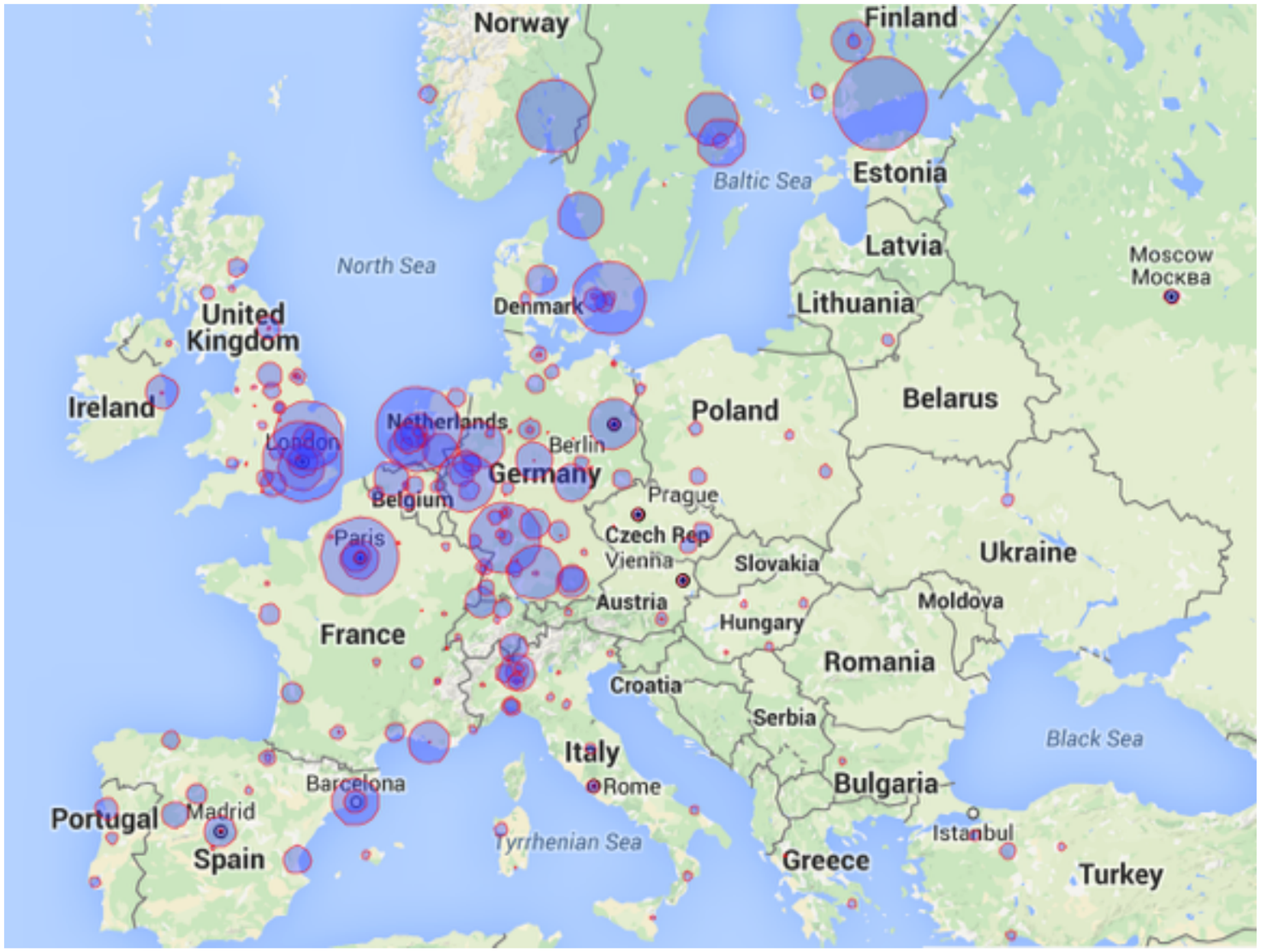
Universitatea de Medicina și Farmacie "Victor Babeș" Timișoara

PROGRAMUL DE COOPERARE ELVEȚIANO-ROMÂN
SWISS-ROMANIAN COOPERATION PROGRAMME

University of Zurich

. . . FINDS A GEOGRAPHICALLY BIASED PUBLICATION LANDSCAPE

In a content analysis of >2800 articles from 1994-2016 for molecular-cytogenetic, genomic array and whole exome/genome sequencing methods, no publication with primary Romanian stakeholder could be identified.



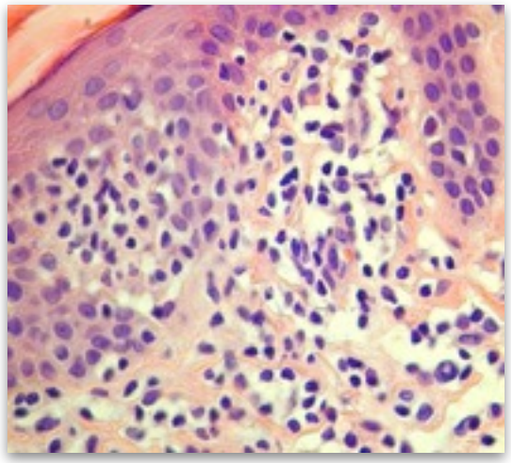
A COLLABORATION STUDY IN THE BIOLOGY OF RARE SKIN DISEASES

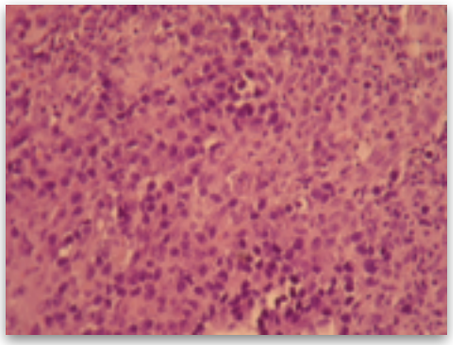
- Establishment of a Romanian Biobank for high-quality specimen from rare, lymphoid skin diseases
- Implementation of procedures and protocols for the preparation of high quality biosamples
- Molecular analysis (genotyping by genomic arrays, whole exome sequencing, immunotyping)
- Improvement of bioinformatic methods for genome data analysis
- Identification of candidate disease markers

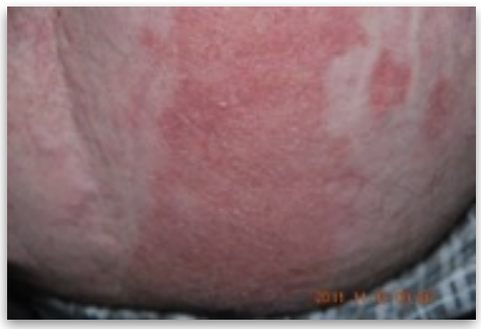
A BIOBANK SUPPORTING RARE DISEASE RESEARCH IN DERMATOPATHOLOGY

Biobanks

- wide array of biospecimens
  - tissue biopsies
  - blood samples
  - purified specimen
- associated with extensive clinical data
- basis for refined diagnostics and selected/targeted therapy







=> **Personalised medicine**

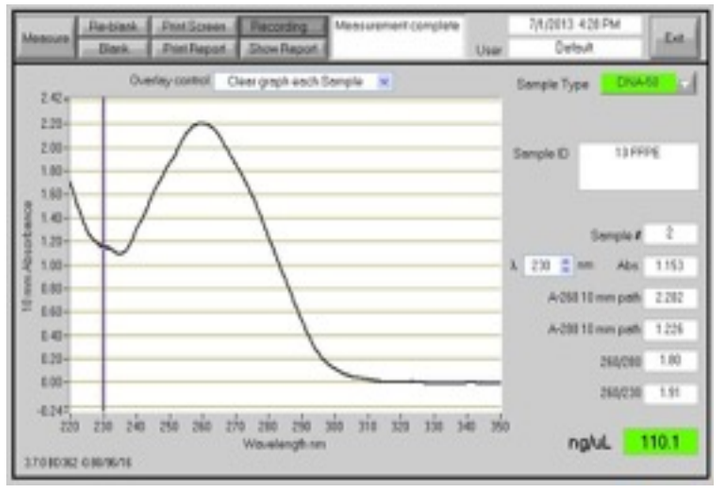
=> **Biomedical research**

TIMISOARA DERMATOPATHOLOGY BIOBANK

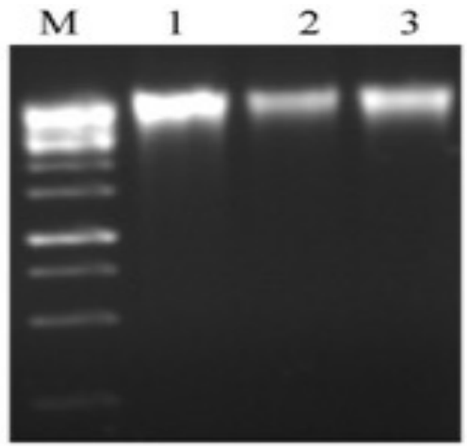
- Samples
  - skin biopsies, fresh frozen
  - formalin fixed, paraffin embedded tissues
  - peripheral blood samples
  - extracted DNA
- Patient registry
- Standardised collection protocols

TIMISOARA DERMATOPATHOLOGY BIOBANK: PILOT STUDY

- Material
  - genomic, high molecular weight DNA
- Quality assessment
  - DNA purity, concentration and fragmentation were assessed at the Timisoara biobank prior to genome analyses
- Genome screening analyses
  - high-density SNP arrays (AROS, Denmark)
  - Whole Exome Sequencing (FGCZ, Universität Zürich)



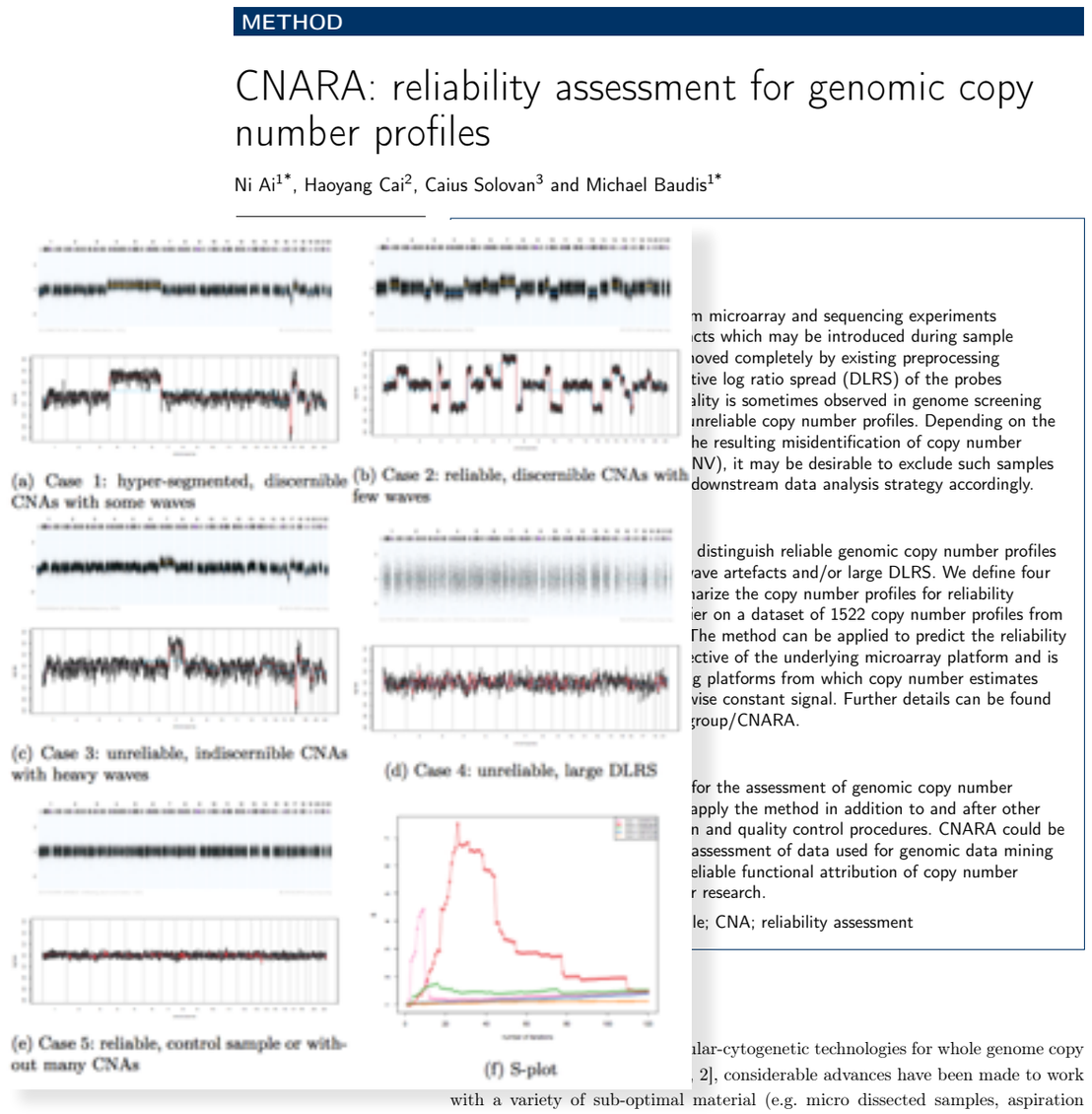
DNA quality assessment



High molecular weight DNA

CNARA: RELIABILITY ASSESSMENT FOR GENOMIC COPY NUMBER PROFILES

- DNA copy number profiles from microarray and sequencing experiments sometimes contain artefacts
- such artefacts cannot be removed completely by existing preprocessing methods
- we have developed a custom data evaluation method (CNARA) for assessing suitability of segmented genome data for downstream analyses



Ai *et al.*

METHOD

CNARA: reliability assessment for genomic copy number profiles

Ni Ai[1*], Haoyang Cai[2], Caius Solovan[3] and Michael Baudis[1*]

m microarray and sequencing experiments
cts which may be introduced during sample
oved completely by existing preprocessing
tive log ratio spread (DLRS) of the probes
ality is sometimes observed in genome screening
nreliable copy number profiles. Depending on the
he resulting misidentification of copy number
NV), it may be desirable to exclude such samples
downstream data analysis strategy accordingly.

distinguish reliable genomic copy number profiles
ave artefacts and/or large DLRS. We define four
arize the copy number profiles for reliability
ier on a dataset of 1522 copy number profiles from
The method can be applied to predict the reliability
ective of the underlying microarray platform and is
g platforms from which copy number estimates
vise constant signal. Further details can be found
group/CNARA.

for the assessment of genomic copy number
apply the method in addition to and after other
n and quality control procedures. CNARA could be
assessment of data used for genomic data mining
eliable functional attribution of copy number
r research.

le; CNA; reliability assessment

(a) Case 1: hyper-segmented, discernible CNAs with some waves

(b) Case 2: reliable, discernible CNAs with few waves

(c) Case 3: unreliable, indiscernible CNAs with heavy waves

(d) Case 4: unreliable, large DLRS

(e) Case 5: reliable, control sample or without many CNAs

(f) S-plot

ular-cytogenetic technologies for whole genome copy
2], considerable advances have been made to work with a variety of sub-optimal material (e.g. micro dissected samples, aspiration biopsies, paraffin embedded tissue), both in the areas of DNA preparation, labeling and platform technologies as well as in bioinformatic processing of the experimental

(Under review at BMC Genomics, April 2016)

PILOT STUDY IN LYMPHOPROLIFERATIVE SKIN DISEASES AND CUTANEOUS LYMPHOMAS

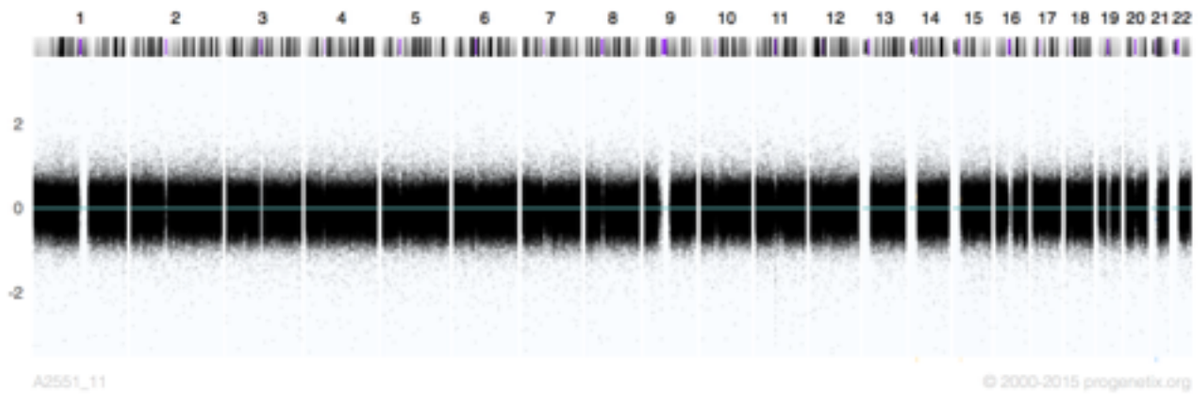
Diagnostic Group	Number of Successful Array Analyses
Sezary	2
Mycosis fungoides	16
T-NHL (other)	0
LPP/SPP	24
cBNHL	3
Others	29
Total	**74**

Diagnostic Group	Number of Successful WES Analyses
Sezary	0
Mycosis fungoides	6
T-NHL (other)	4
LPP/SPP	4
cBNHL	3
Others	6
Total	**23**

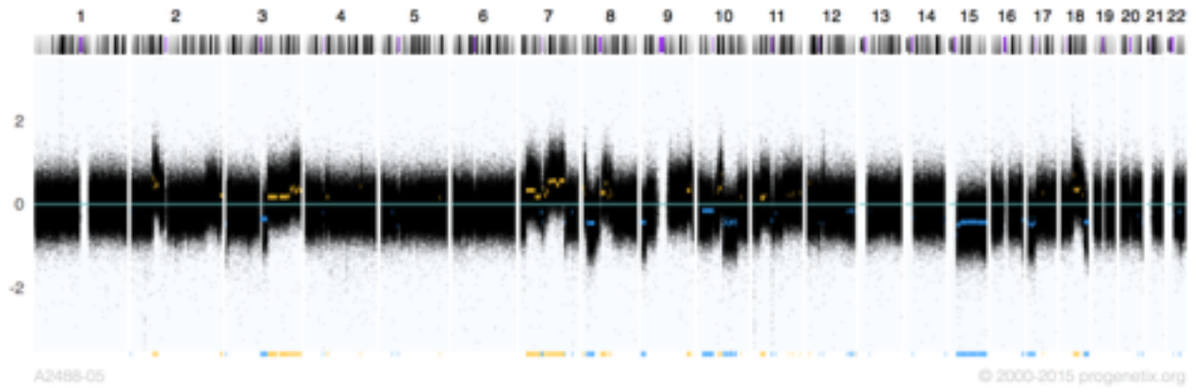
In total, 74 of 90 array datasets and all 23 Whole Exome Sequencing data sets could be evaluated for genomic aberrations

WHOLE-GENOME SNP ARRAYS IN C-LPD

- A total of 90 DNA samples were submitted for SNP-array hybridisation (AROS, Denmark)
- Data quality evaluation and bioinformatic analysis were performed at the University of Zurich
- 74/90 arrays (82%) passed quality assessment using a custom data evaluation method (CNARA)



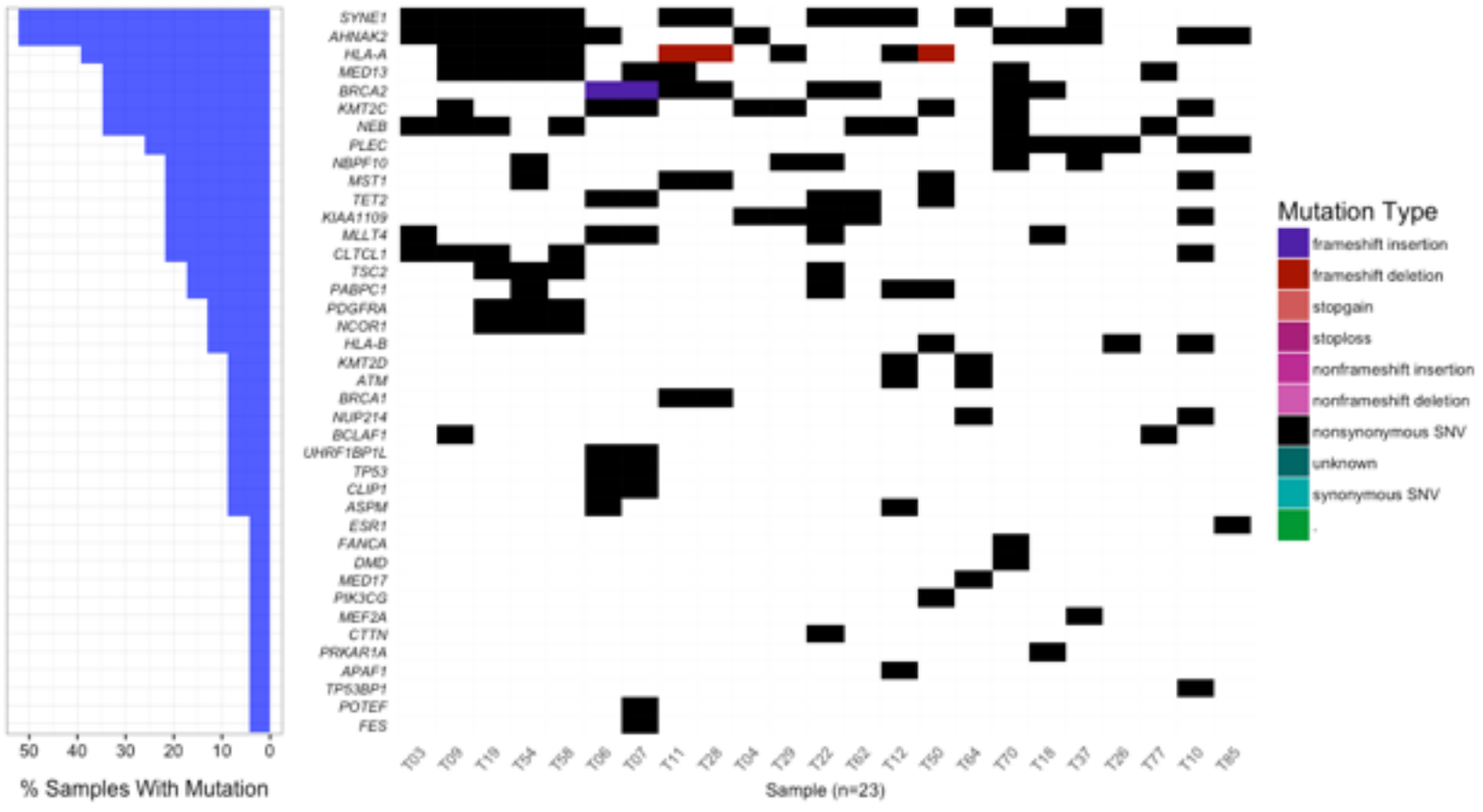
Lack of gross genomic changes in a case of Small plaque parapsoriasis (SPP)



Multiple genomic copy number changes in a case of cutaneous B-cell lymphoma

WHOLE-EXOME SEQUENCING STUDY IN C-LPD

- First WES pilot study comparing genomic mutations across various lymphoproliferative and malignant skin diseases

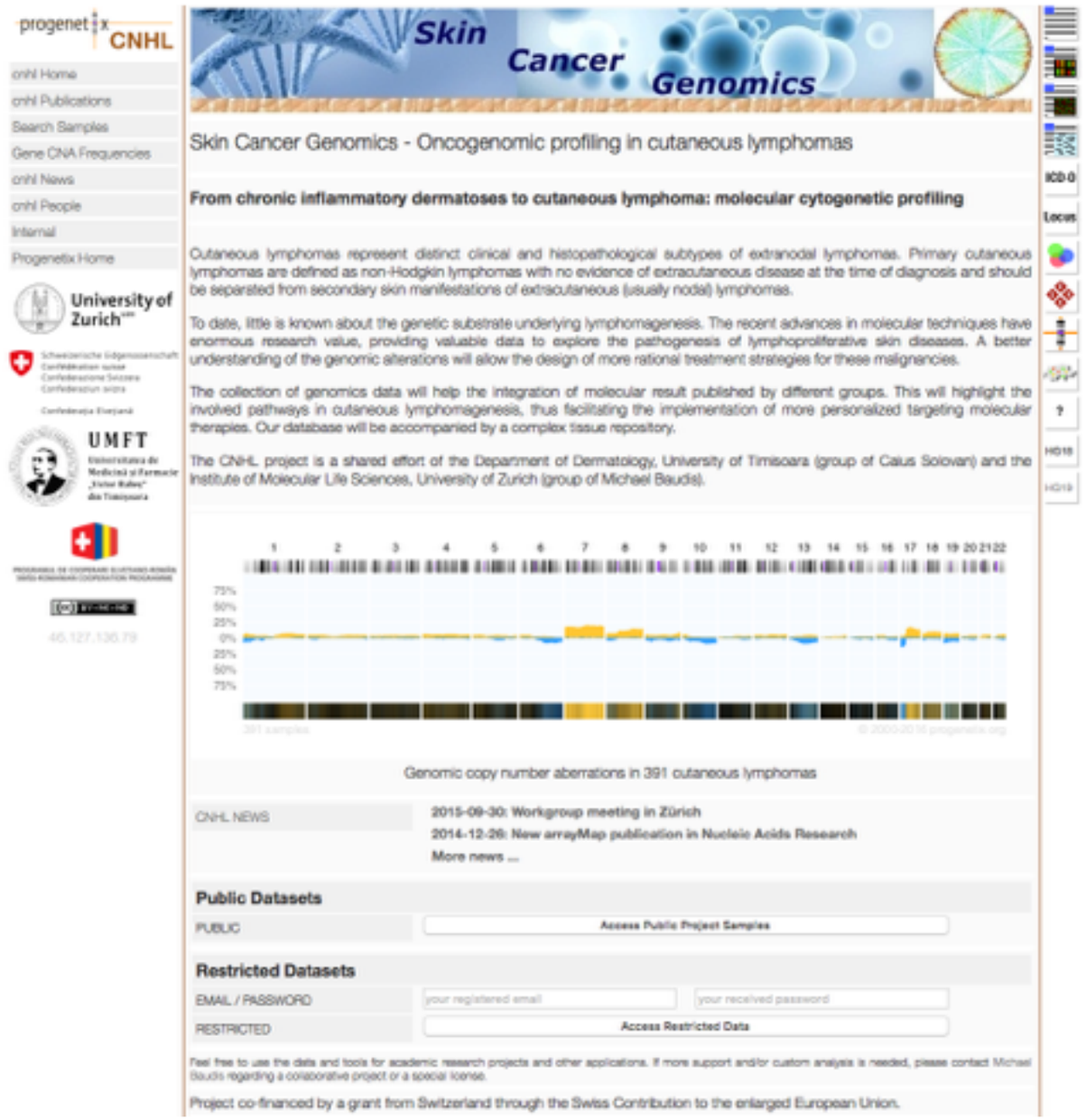


Sample	Clinical diagnosis	Average Coverage	Identified SNV
T03 *	cTNHL	105.08	148534
T04	MF	101.85	131925
T06	LPP	104.51	149307
T07	cBNHL	105.36	150033
T09 *	cALCL	105.19	118457
T10	LPP/MF	101.97	121289
T11	MF	109.7	122354
T12	MF	124.34	127385
T18	MF/Sarcoidosis	111.5	122857
T19 *	cALCL	114.54	157448
T22	MF/SS	96.55	138184
T26	MF	106.01	118446
T28	MF	112.09	121867
T29	cTNHL	110.56	122467
T37	SPP	96.3	114809
T50	REM	116.09	121829
T54	cBNHL	113.79	121041
T58	cBNHL	169.1	179496
T62	MF	52.5	95533
T64	SPP	128.96	121373
T70	SPP	105.99	119617
T77	MF-like	71.96	102493
T85	SDR-like	51.28	85112

* denotes samples retrieved from the same patient

WEB-BASED INFORMATION RESOURCE

- Representation of project details
- Resource for collected genomic profiling data from T-cell related cutaneous malignancies
- Portal/workbench for genomic screening data produced inside of the RSRP project



cnhl.progenetix.org

PROJECT: BIOBANK

- consisting of nucleic acids, corresponding blood and tissue samples
- full donor authorisation and ethics committee approval for use in research activities
- enables the use of the material and anonymised supporting information in research studies

PROJECT: TECHNICAL SUMMARY

- Molecular screening analysis of heterogeneous, lymphoproliferative skin diseases and cutaneous lymphomas, using state-of-the-art techniques
- Bioinformatics method development
- Detection of genomic aberrations with possible implications for disease classification and prognosis

PROJECT: ADMINISTRATION & GENERAL NOTES

- excellent control & support through the grant management organisations UEFISCDI and SNSF
- overall acceptable level of “transactional overhead”
- main problem: Lack of followup or extension options
  - biobank as seed resource project with huge potential, but without long-term support strategy
  - still ongoing data analysis w/o further grant support
  - no specific framework for CH-Ro followup proposals

WHOLE GENOME PROFILING IN CANCER: PUTTING ROMANIA ON THE MAP

Universität Zürich[UZH]

NI AI
MICHAEL BAUDIS
(HAOYANG CAI)
LINDA GROB
SAUMYA GUPTA

UMFT
Universitatea de Medicină și Farmacie „Victor Babeș" din Timișoara

CAIUS SOLOVAN
EDWARD SECLAMAN
MANFRED BELEUT
MARIA IORDACHE
ELENA CHITICARIU
FLAVIA BADERCA

Danke!

Thank You!

Mulţumim!

PROGRAMUL DE COOPERARE ELVEȚIANO-ROMÂN
SWISS-ROMANIAN COOPERATION PROGRAMME

FNSNF

KARIN BRÖNNIMANN
TIMOTHY RYAN

UEFISCDI

MONICA CRUCERU