

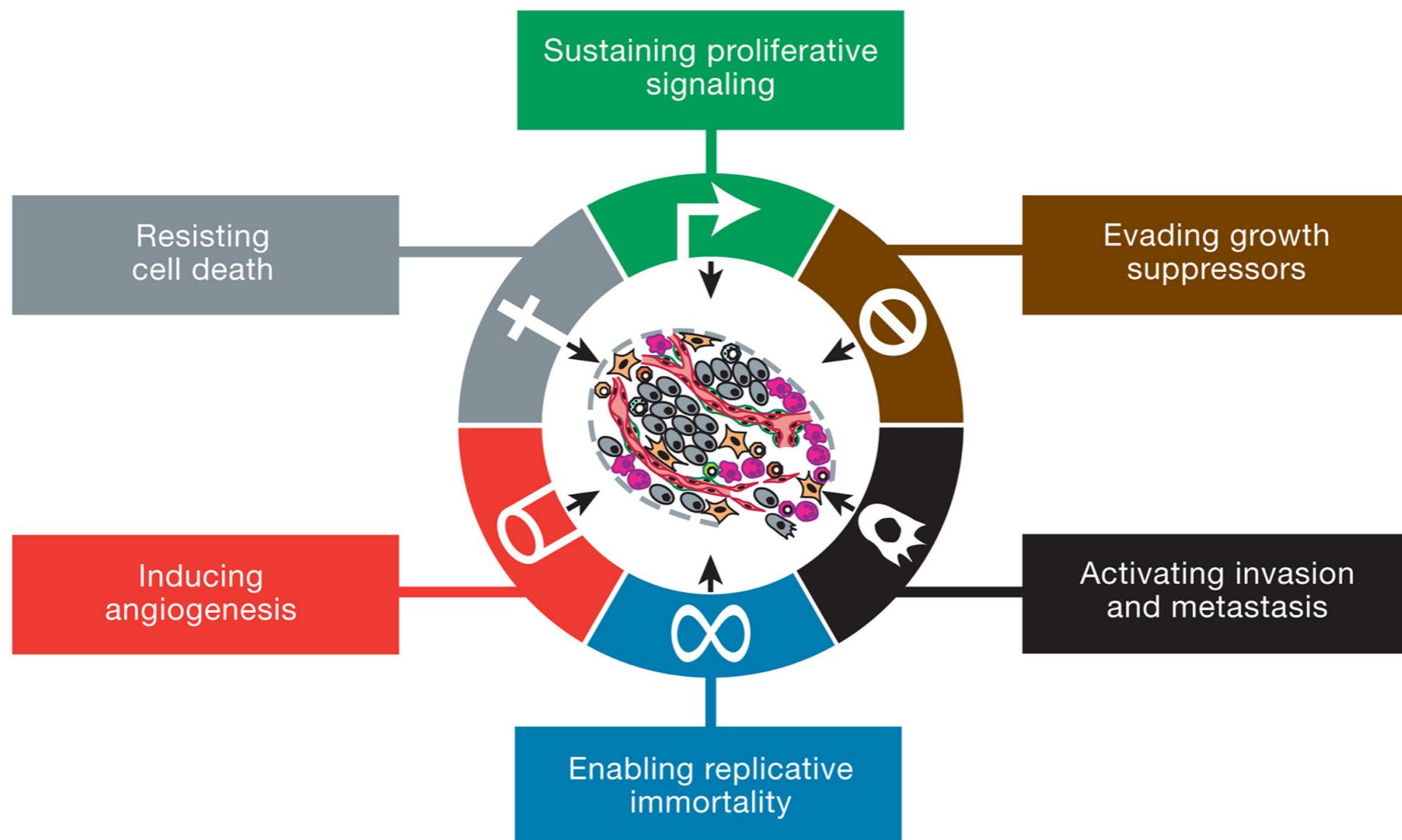
Genome instability analysis and provenance estimation in cancer cell lines, based on the large-scale evaluation of genome profiles and associated metadata

Rahel Paloots



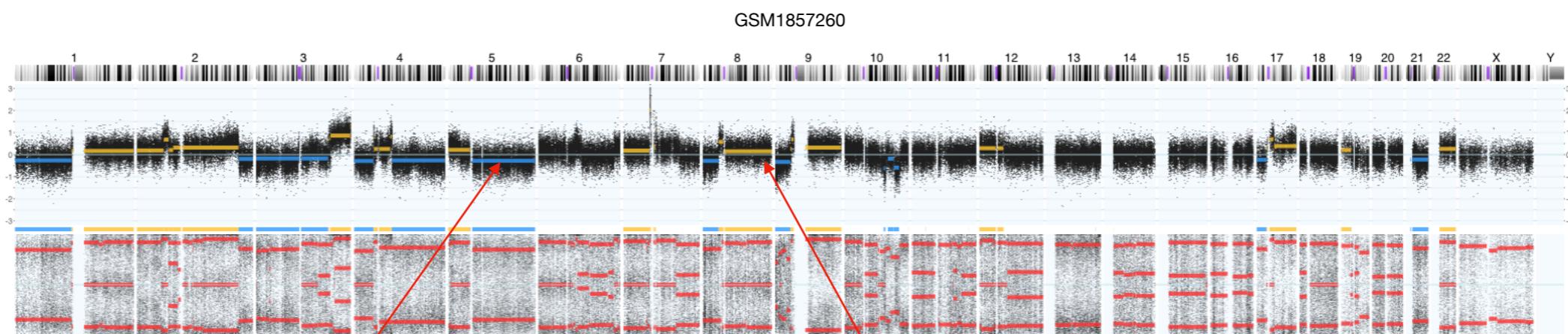
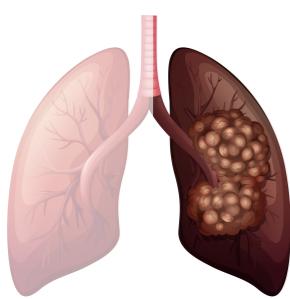
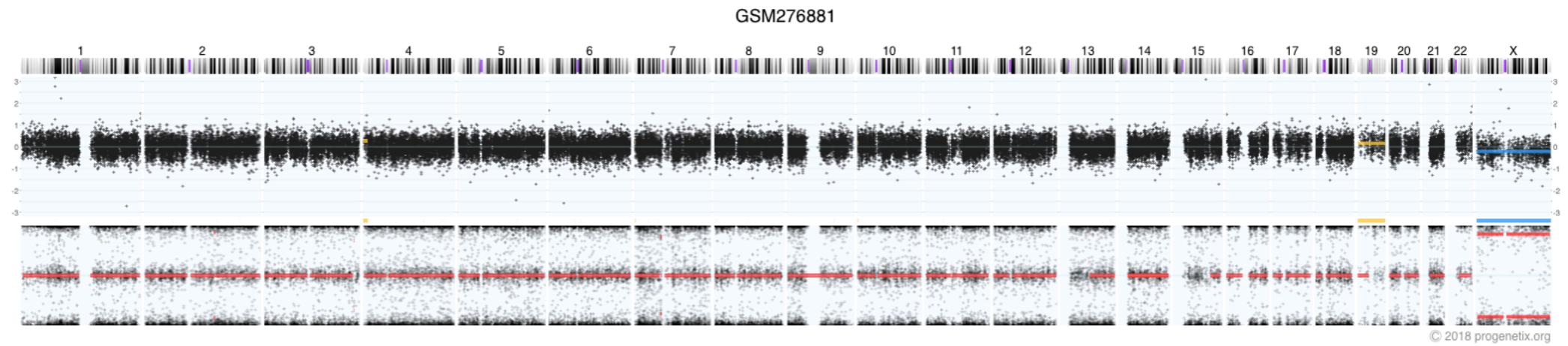
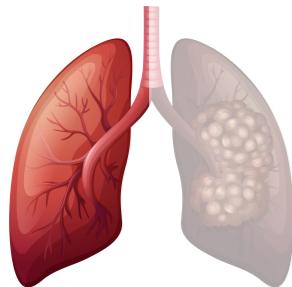
University of
Zurich UZH

The hallmarks of cancer



Hanahan D, Weinberg RA. Cell 2011; 144, 646-74.]

Copy number aberrations (CNAs) in cancers

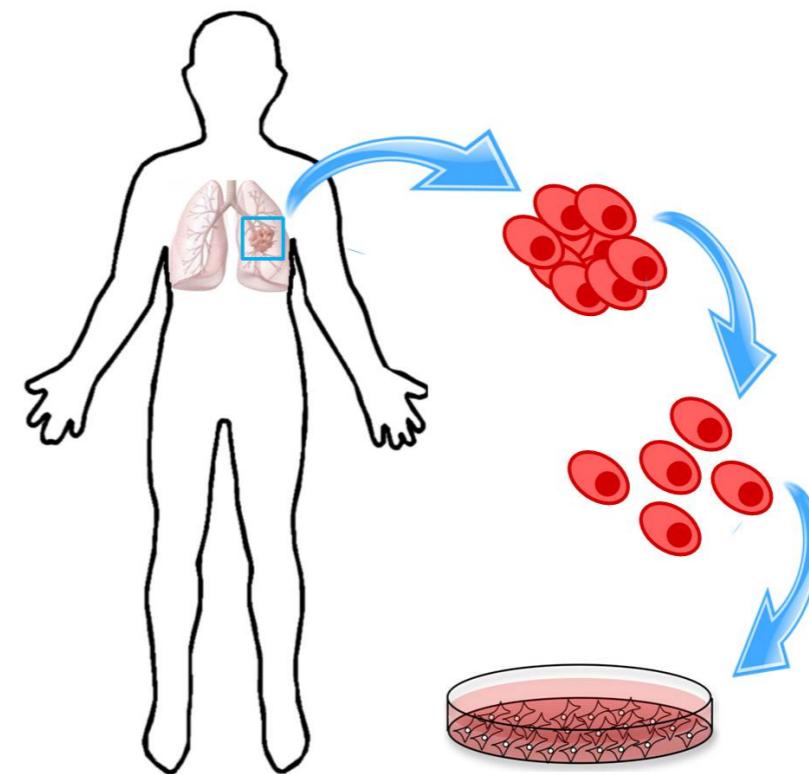


APC deletion

C-MYC amplification

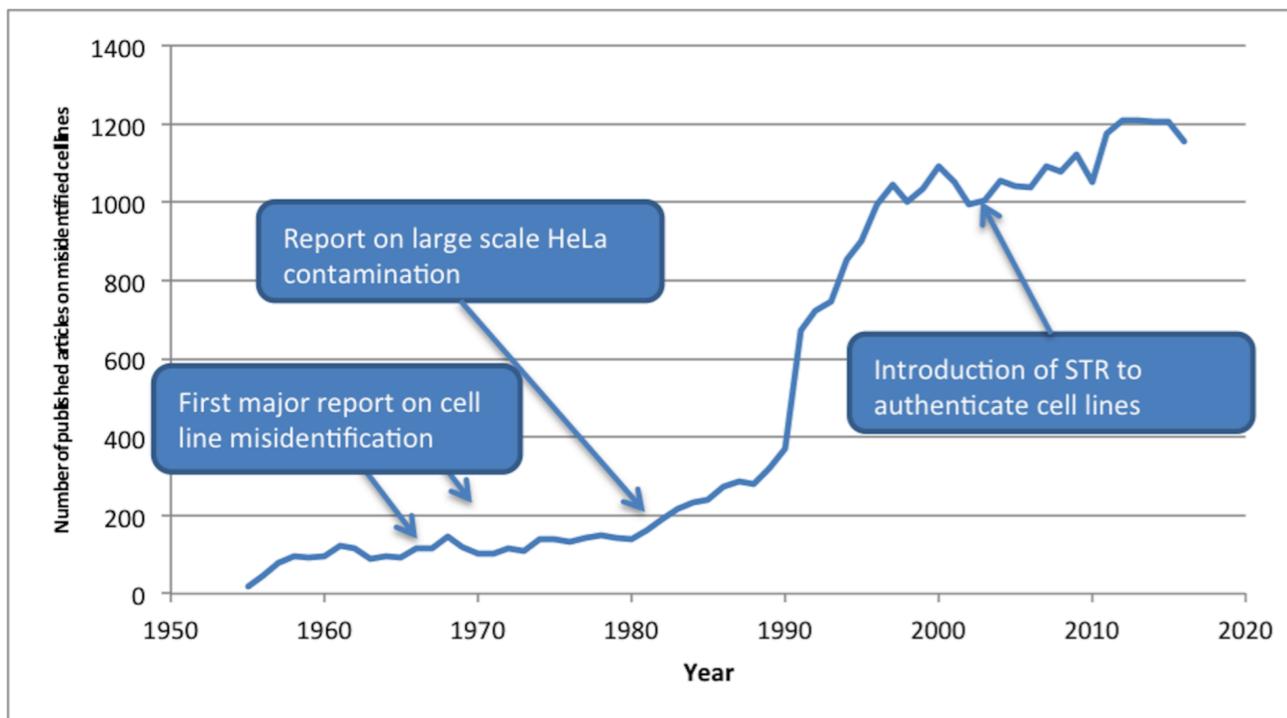
Cancer cell lines

- HeLa 1953
- NCI-60 panel
- Models for drug screening and disease study
- Concerns: are they actually good representatives of the disease?



www.signalsblog.ca/

Cell line misidentification



Everything was going along fine until they discovered their HeLa cell line expressed Y chromosome markers.

- Faulty handling of cell cultures and poor labeling
- Contaminated cell line - contains foreign material
- Misidentified cell line - no longer corresponds to the donor or species from which it was originally established.
- ICLAC - The International Cell Line Authentication Committee
- 488 cell lines.

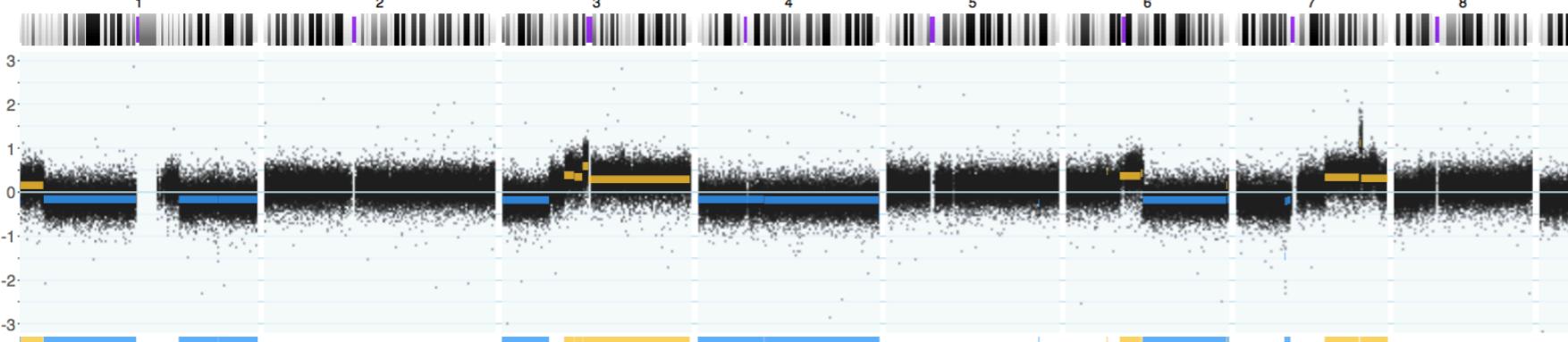
Horbach SPJM, Halfmann W. PLOS ONE 2017; 12(10), [6–10]

promega.de

Objectives

- To have a clearer insight into which cancer cell lines are “correct”
- Have an overview of the heterogeneity of the CNA profiles of cancer cell lines and detect outliers i.e. GEO samples that do not belong to this cell line
- Examine the similarity of the misidentified cell lines to their origins
- Observe the similarity between a given cell line to other cell lines from the same tumor type and to its primary tumor

Methods



www.arrayMap.org

<https://www.ncbi.nlm.nih.gov/geo/>

<https://web.expasy.org/cellosaurus/>

GSM750823	GSM750823
ARRAYID	GSM750823
BIOSAMPLEID	AM_BS_GSM750823
ARRAYNAME	ME:MDA-N (69213)
SERIESID	GSE30291
DIAGNOSTISTEXT	Amelanotic melanoma [cell line MDA-N]



Cellosaurus

Cellosaurus MDA-N (CVCL_1910)

Cell line name	MDA-N
Synonyms	MDAN
Accession	CVCL_1910
Resource Identification Initiative	To cite this cell line use: MDA-N (RRID:CVCL_1910)
Comments	<p>Problematic cell line: Contaminated. Parent cell line (MDA-MB-435) has been part of NCI-60 cancer cell line panel.</p> <p>Doubling time: 22.5 hours (NCI-DTP).</p> <p>Microsatellite instability: Stable (MSS) (PubMed=12661003).</p> <p>Omics: Array-based CGH.</p> <p>Omics: CNV analysis.</p> <p>Omics: Transcriptome analysis.</p>
Sequence variations	Heterozygous for BRAF p.Val600Glu (c.1799T>A) (from parent cell line). Heterozygous for TP53 p.Gly266Glu (from parent cell line).
Disease	Amelanotic melanoma (NCIt: C3802) Derived from metastatic site: Subcutaneous; right buttock.
Species of origin	Homo sapiens (Human) (NCBI Taxonomy: 9606)
Hierarchy	Parent: CVCL_0417 (MDA-MB-435)
Sex of cell	Male
Age at sampling	33Y
Category	Cancer cell line

 NCBI
 [HOME](#) [SEARCH](#) [SITE MAP](#)

 Gene Expression Omnibus

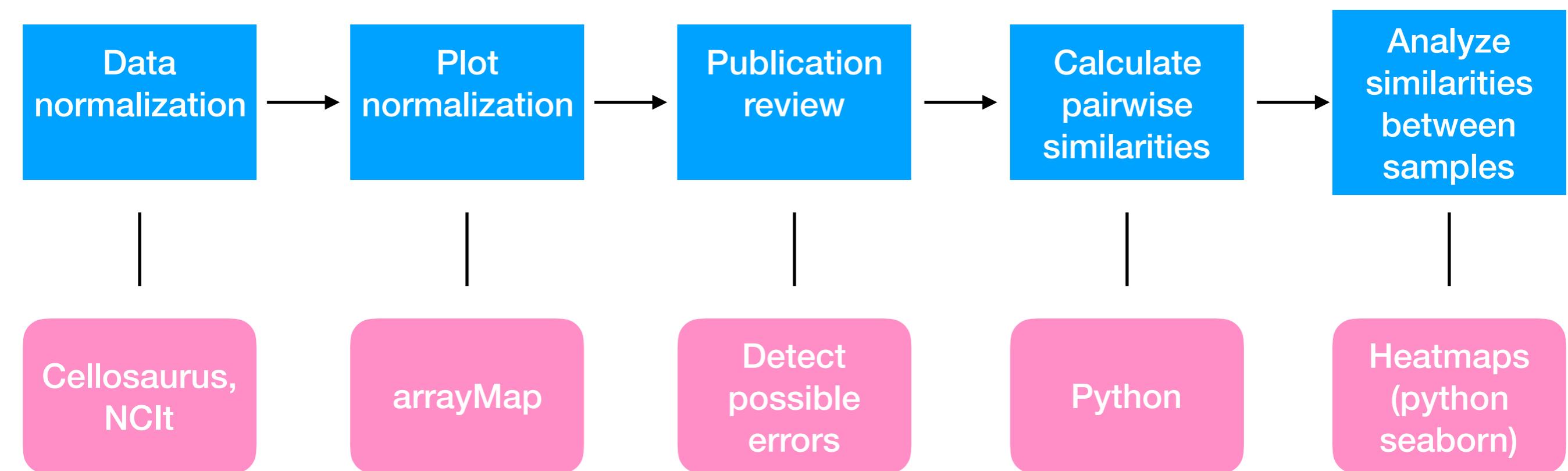
NCBI > GEO > Accession Display

GEO help: Mouse over screen elements for information.

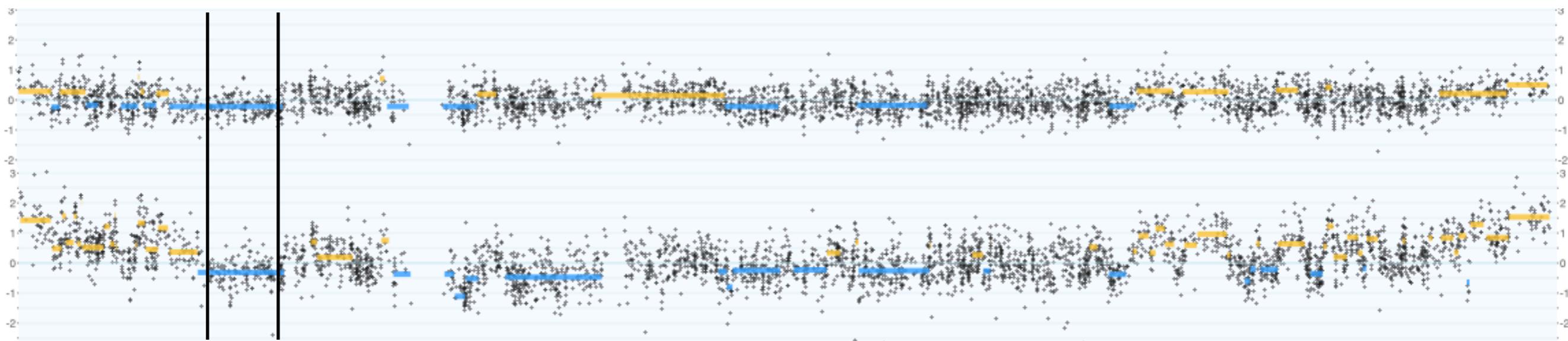
Scope: Self Format: HTML Amount: Quick GEO accession:

Sample GSM750823		Query DataSets for GSM750823
Status	Public on Oct 24, 2011	
Title	ME:MDA-N (69213)	
Sample type	genomic	
Channel 1		
Source name	ME:MDA-N	
Organism	Homo sapiens	
Characteristics	cell line: ME:MDA-N tissue of origin: Melanoma age: 31 Sex: F epithelial: no source: Pleural effusion p53 mutation: MT doubling time: 22.5 sample description: Ductal carcinoma-mammary gland; breast; duct; metastatic site: pleural effusion;	
Treatment protocol	Untreated	
Growth protocol	As described previously (Pubmed ID: 17272646) using conditions based on those adopted by DTP for their drug screen and harvested at ~80% confluence.	

Workflow



Overlap calculation



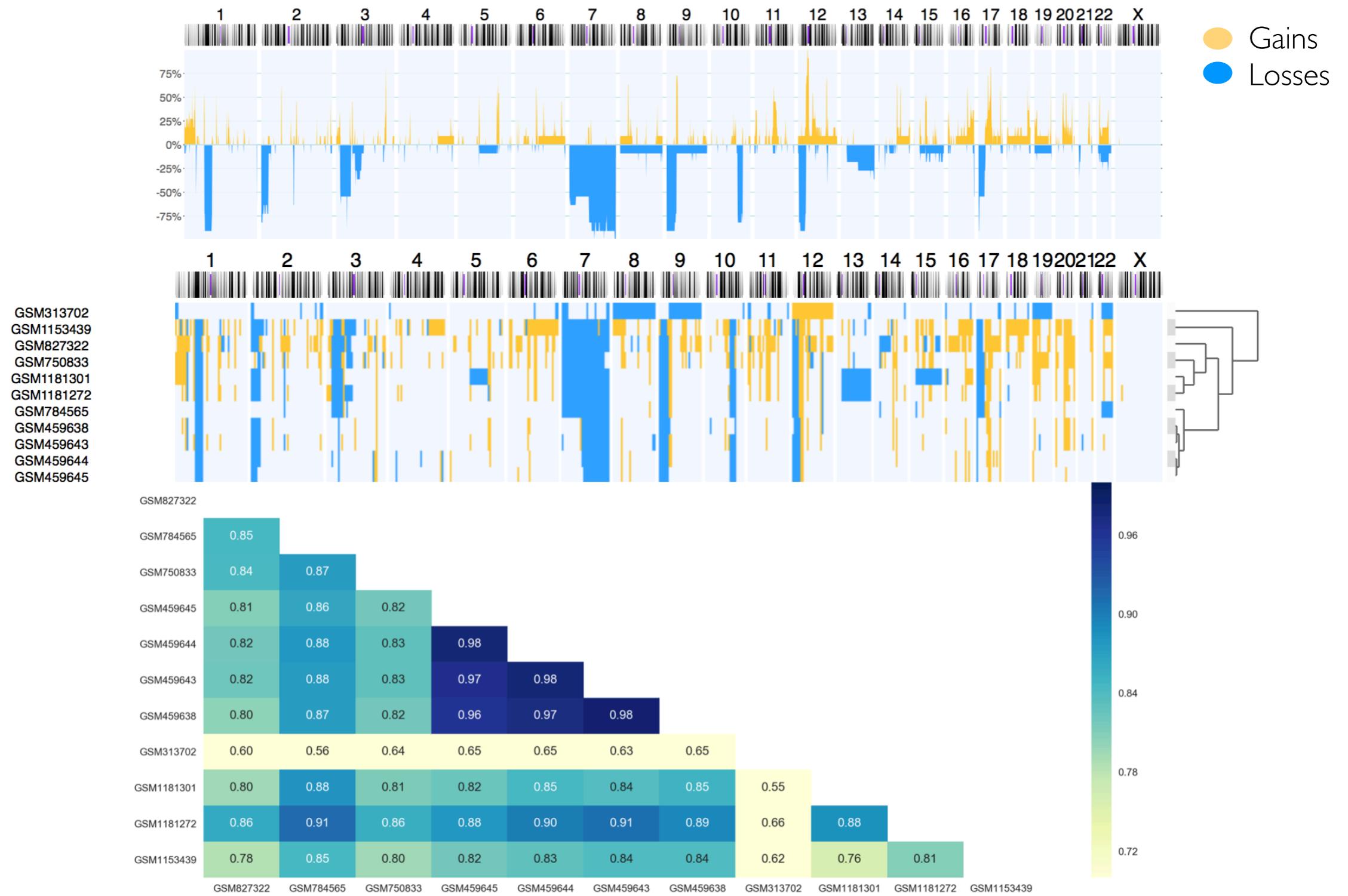
sample_id	chromosome	start	end	type	value	probes
GSM750835	1	73452	925778	DEL	-0.1928	25
GSM750835	1	50343345	50848705	DEL	-0.2942	75
GSM750835	1	50855802	51100877	DUP	0.1604	30
GSM750835	1	51107318	51207435	DEL	-0.1502	12
GSM750835	1	116767606	116851851	DEL	-0.243	14
GSM750835	1	116862294	150369966	DUP	0.2134	836
GSM750835	1	150953418	154316100	DUP	0.2728	408

Results

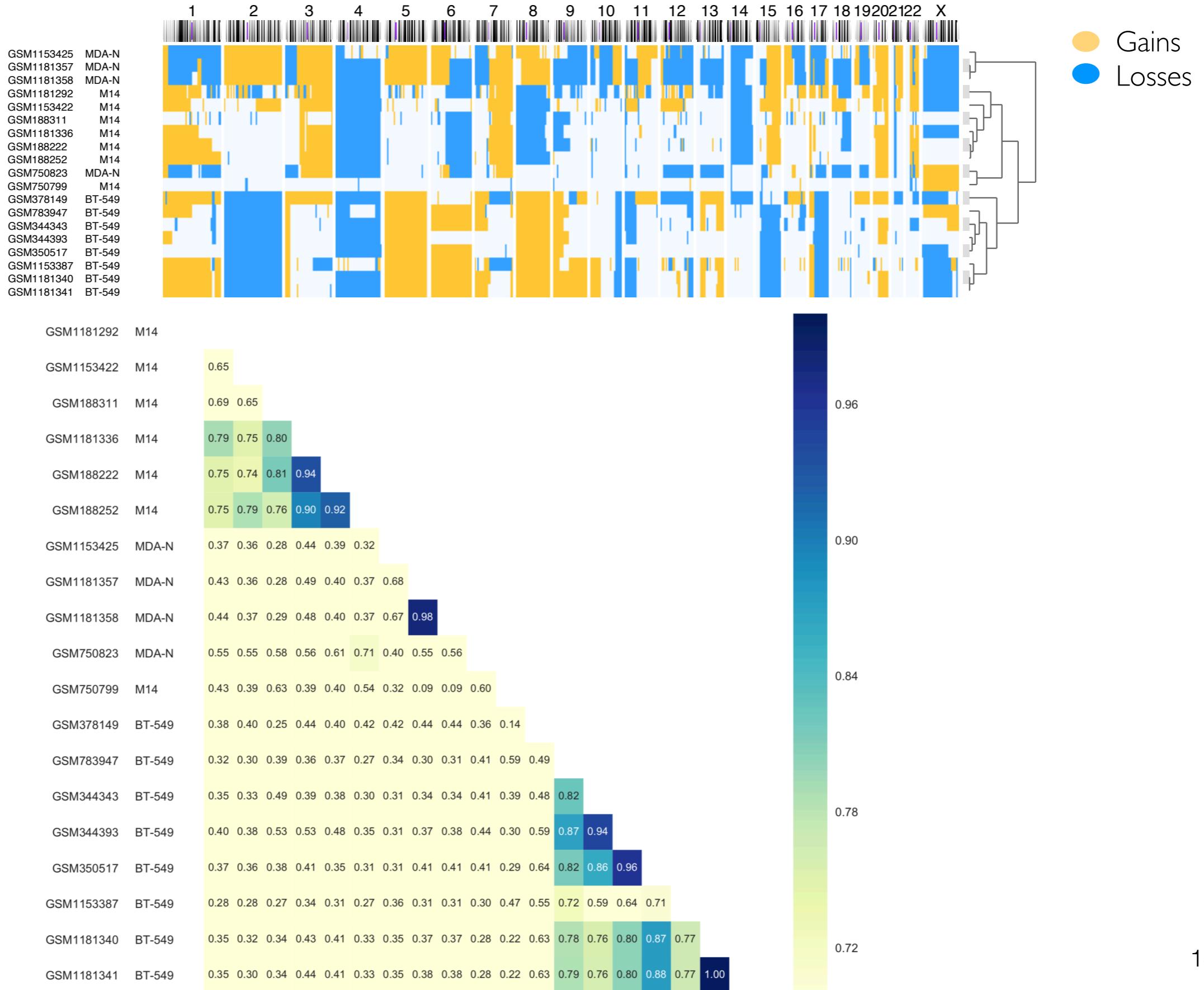
HeLa



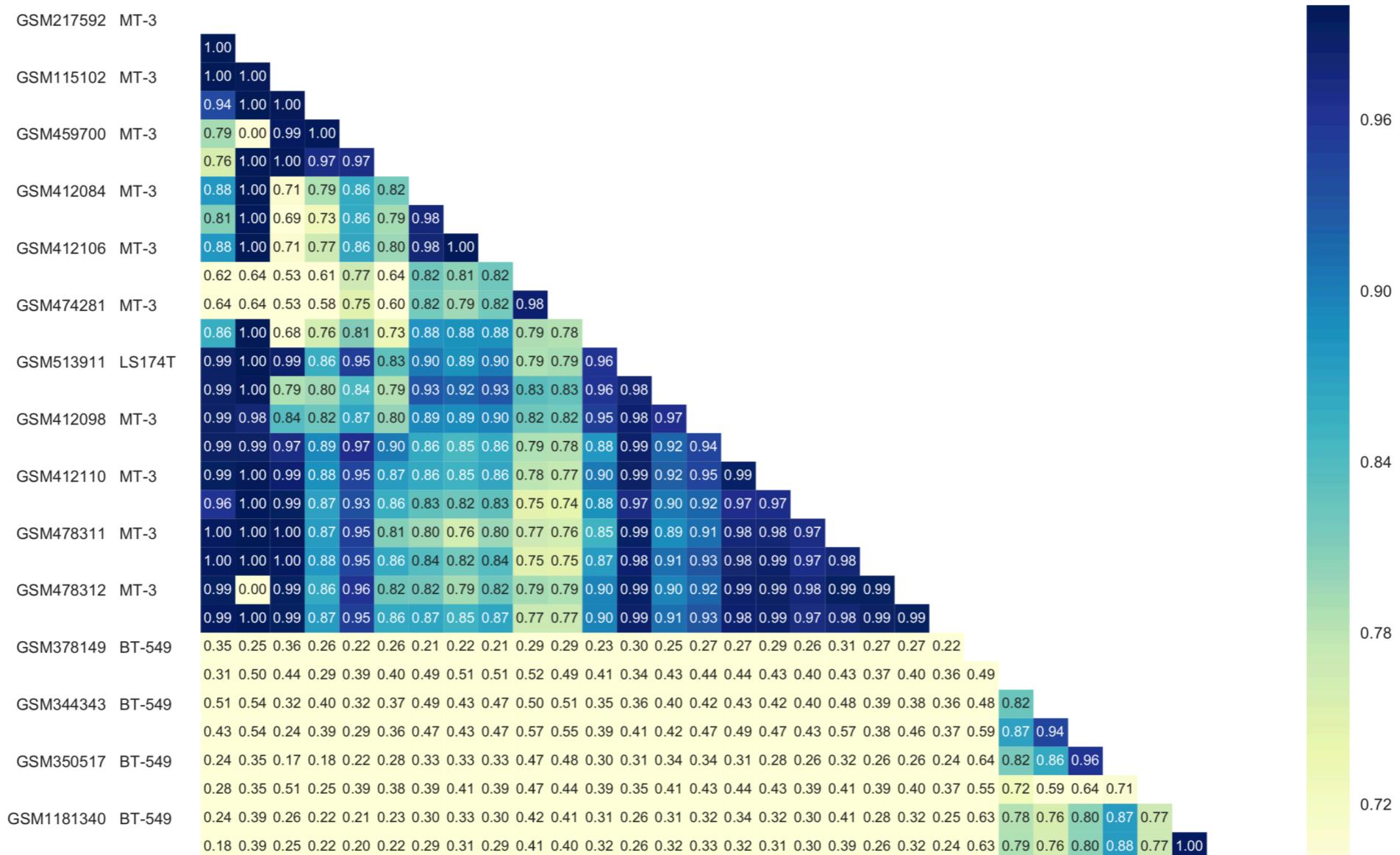
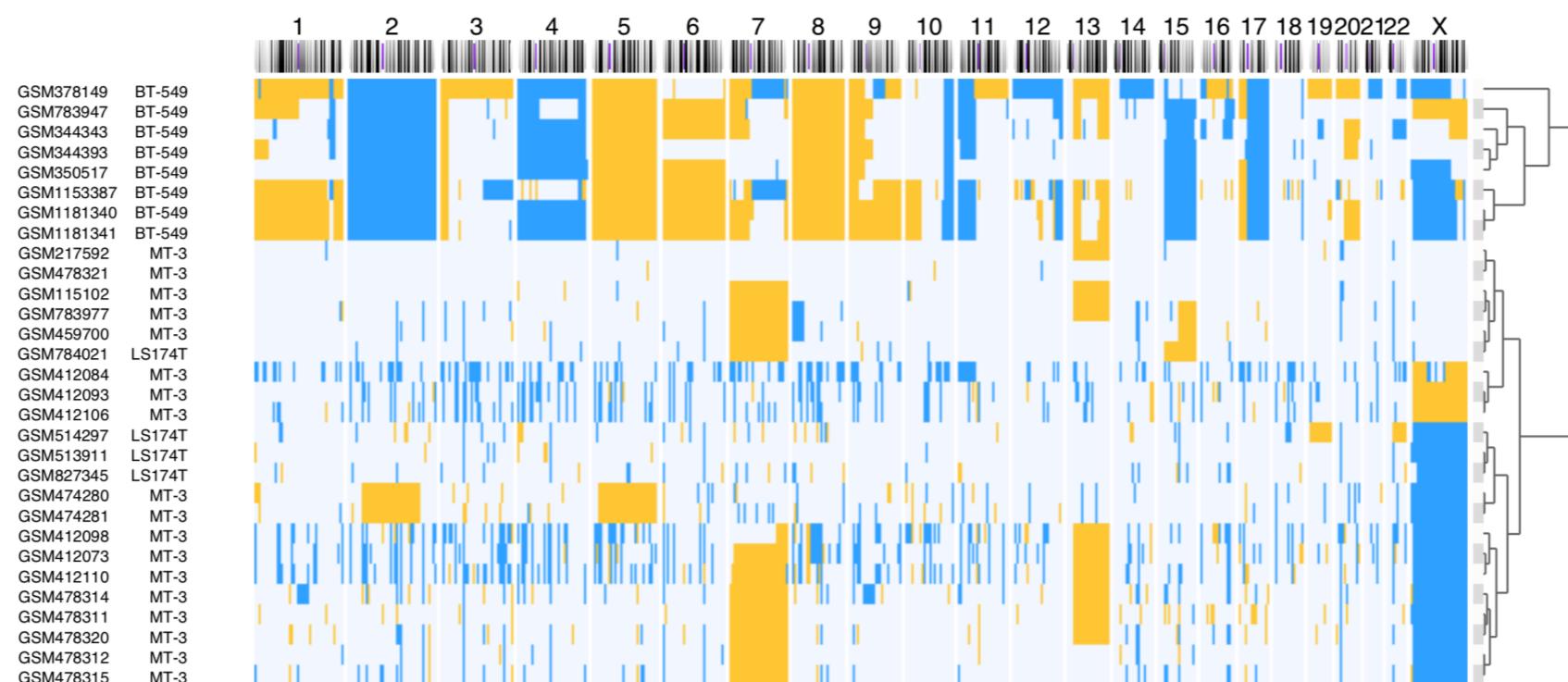
SK-OV-3



Misidentified cell line MDA-N



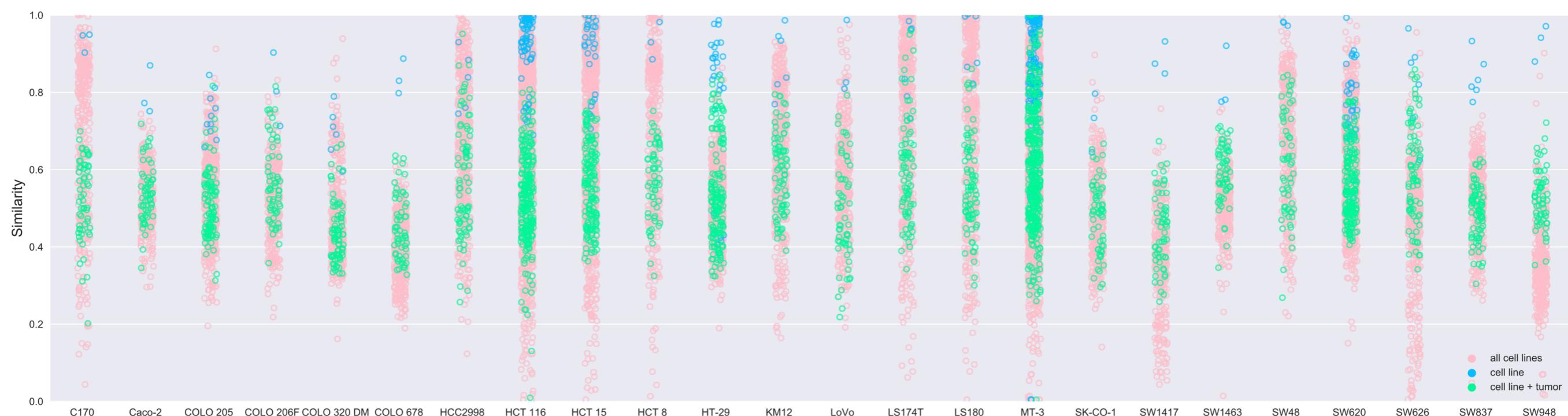
MT-3



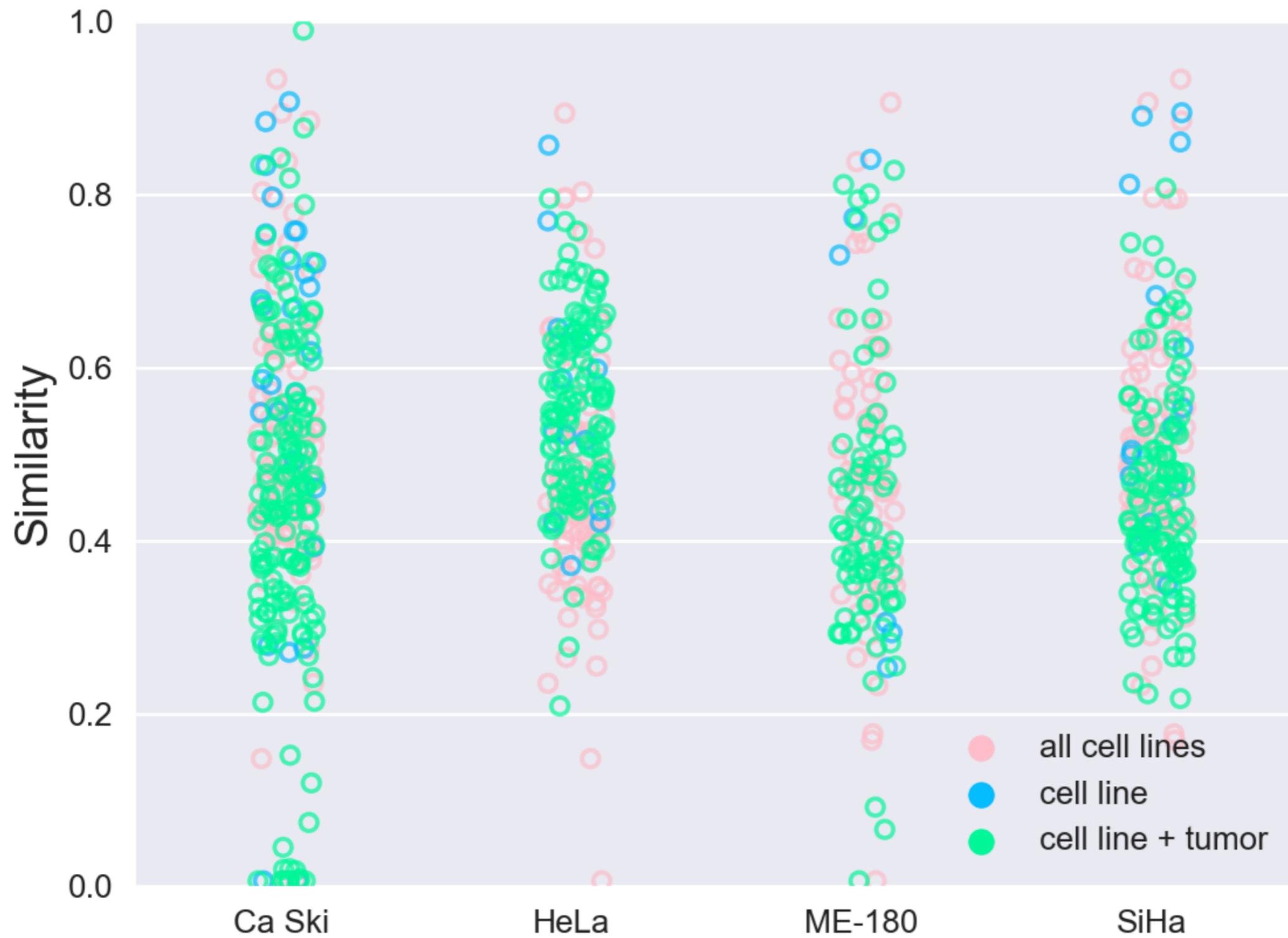
Ovarian carcinoma



Colorectal carcinoma



Cervical carcinoma



Summary

- Outliers for each cell line could be detected by using this method.
- The level of heterogeneity depends on the cell line - some cell lines are more homogenous.
- Misidentified cell lines can be different from their both “supposed” and “actual” origin.
- Microsatellite instable cancer cell lines exhibit lower number of CNAs similarly to microsatellite instable colorectal carcinomas.
- Some cell lines show higher similarity to the primary tumor than others. These cell lines should be used in research.

Outlook

- Improve overlap calculation method by adding length coefficients or weighted segments
- Compare cell lines that showed low similarity to the primary tumor to other tumor types to detect possible misidentification.
- Analyze more cell lines and establish CNA profiles for all cancer cell lines

Thank you!



**University of
Zurich UZH**