

# **Methylation profiling identifies tissue of origin in brain metastases**

Christoph Geisenberger

---

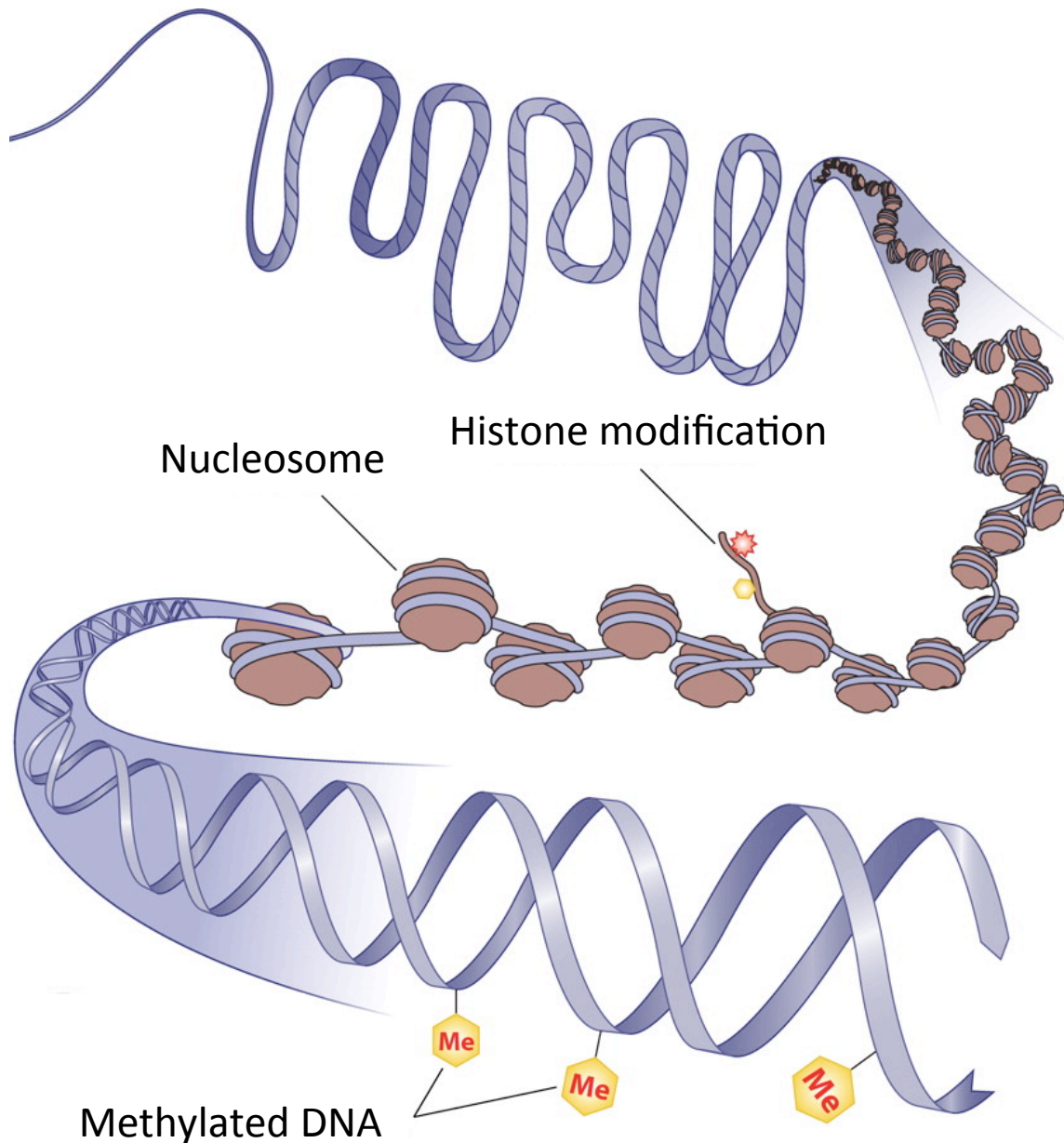
# Introduction

---

- (brain) metastases are fatal sequelae of cancer
- in some cases, primary is unknown (up to 11%)

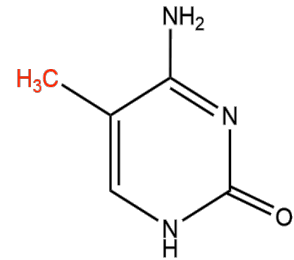
## **objectives**

1. differential methylation between primary tumors and metastases
2. epigenetic data as a potential marker to identify tissue of origin



### mechanism

- DNMT enzymes
- cytosine residues



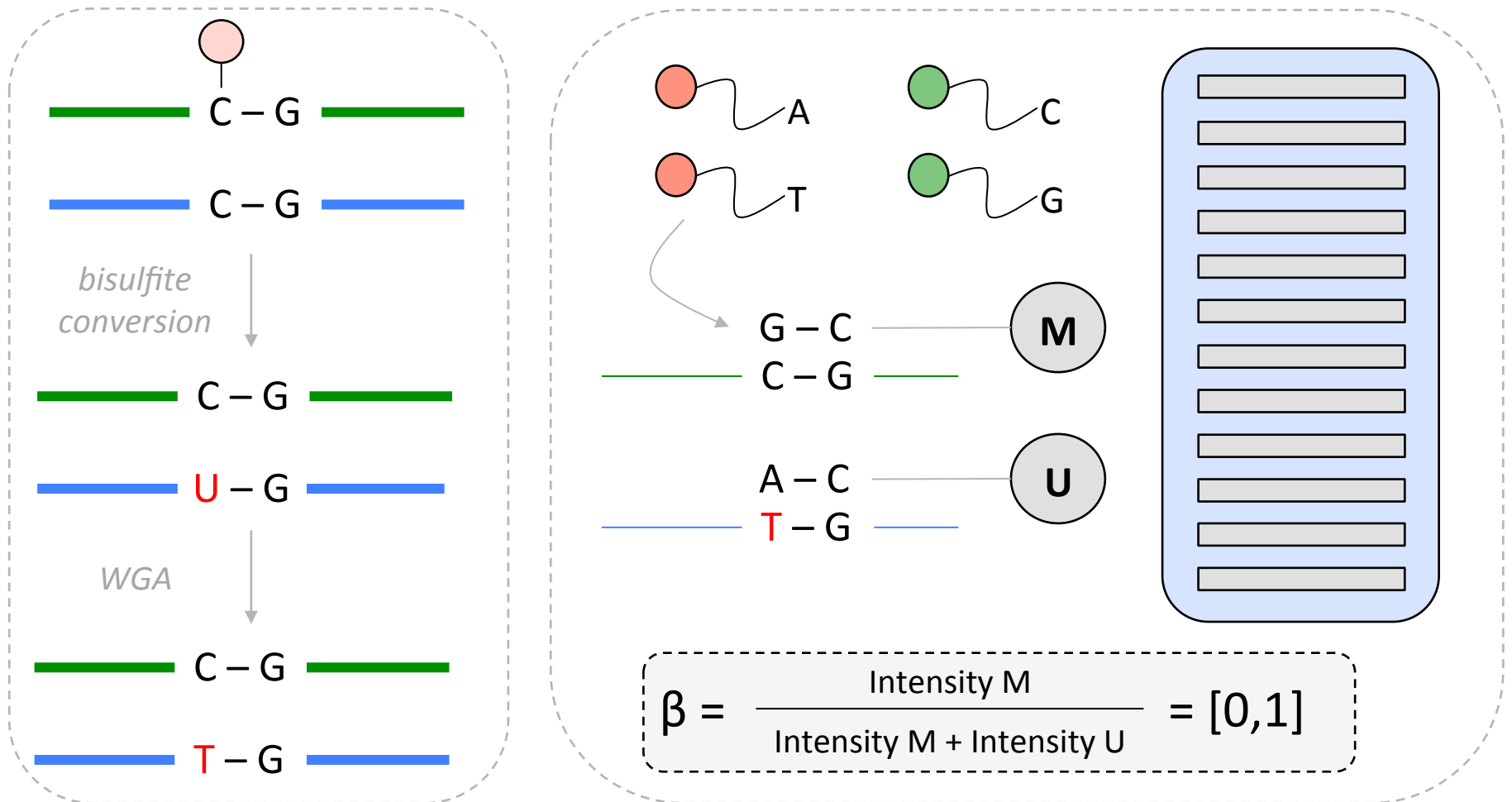
### consequence

- transcriptional repression
- differentiation/development
- tumor suppressor silencing

### assessment

- bisulfite sequencing
- array-based methods

# Infinium<sup>®</sup> HumanMethylation450 BeadChip



# Brain metastases data set

**bladder cancer** (n = 1)

- paired (n = 1)

**breast cancer** (n = 16)

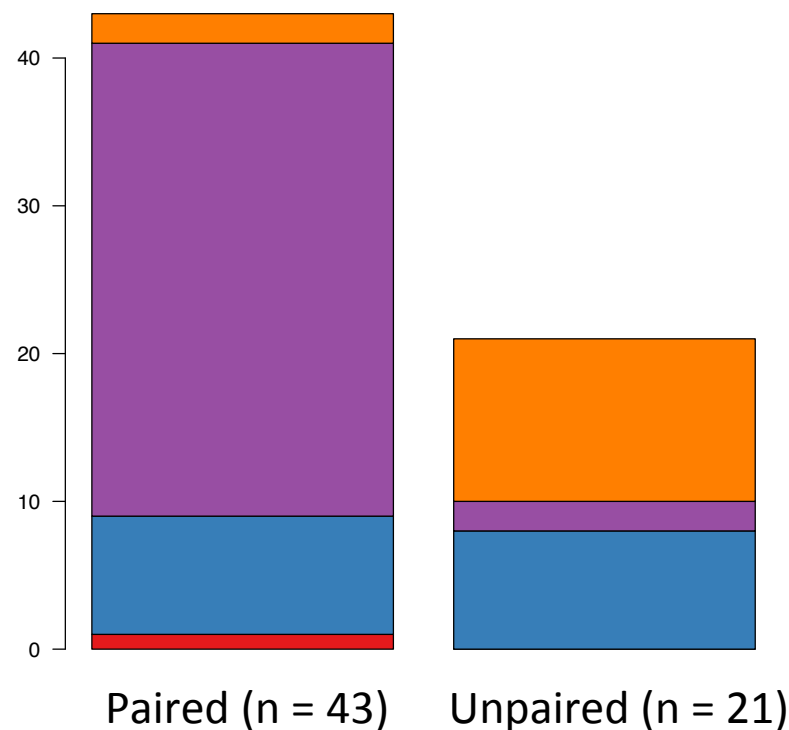
- paired (n = 8)
- unpaired (n = 8)

**lung cancer** (n = 34)

- paired (n = 32)
- unpaired (n = 2)

**melanoma** (n = 13)

- paired (n = 2)
- unpaired (n = 11)

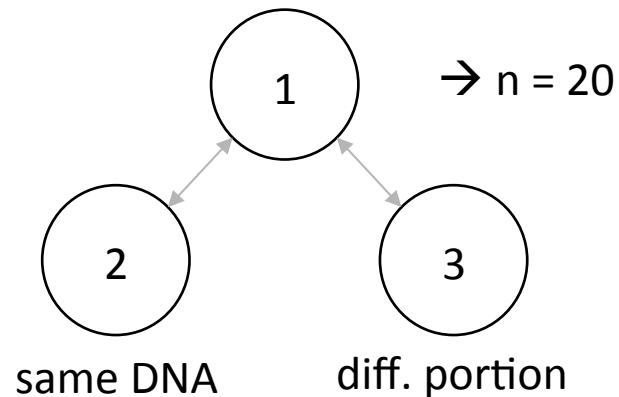


+ 40 primaries & 1 adrenal metastasis

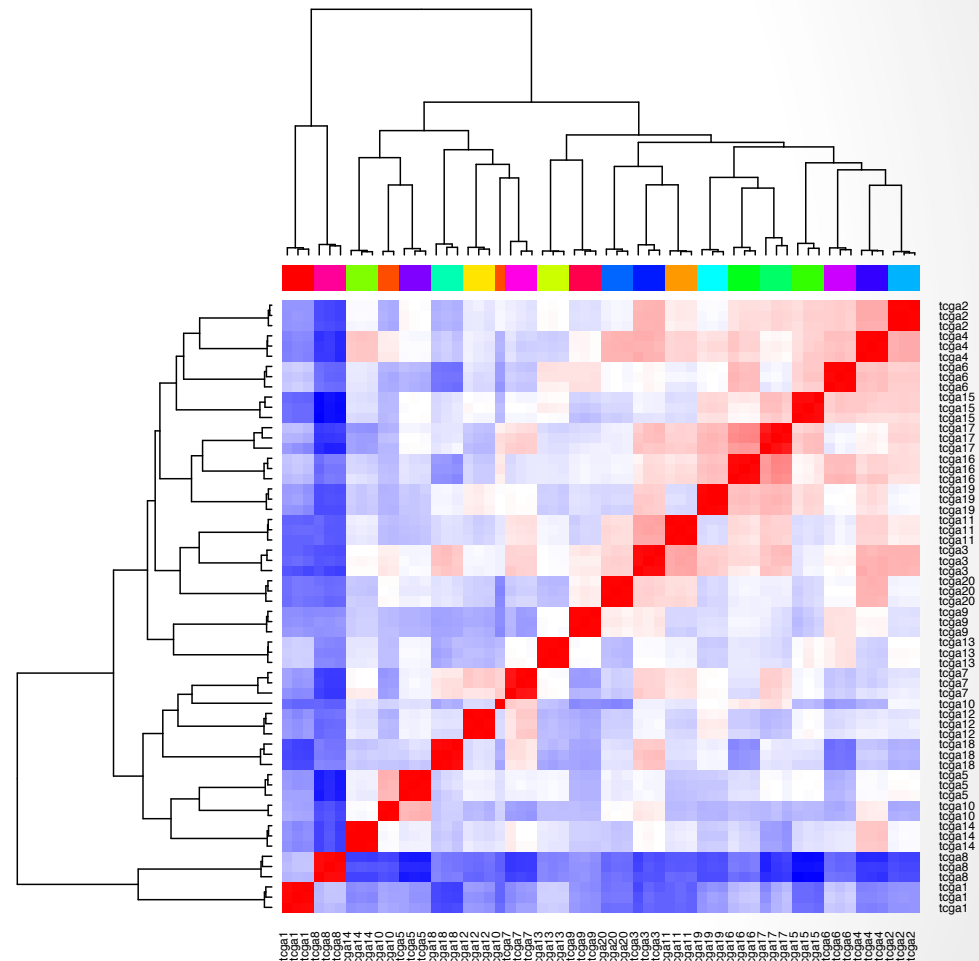
# Statistical difficulties

- t-test (and other statistics) detect *common* differences
- no statistical test for  $n = 1$

## estimation of technical noise



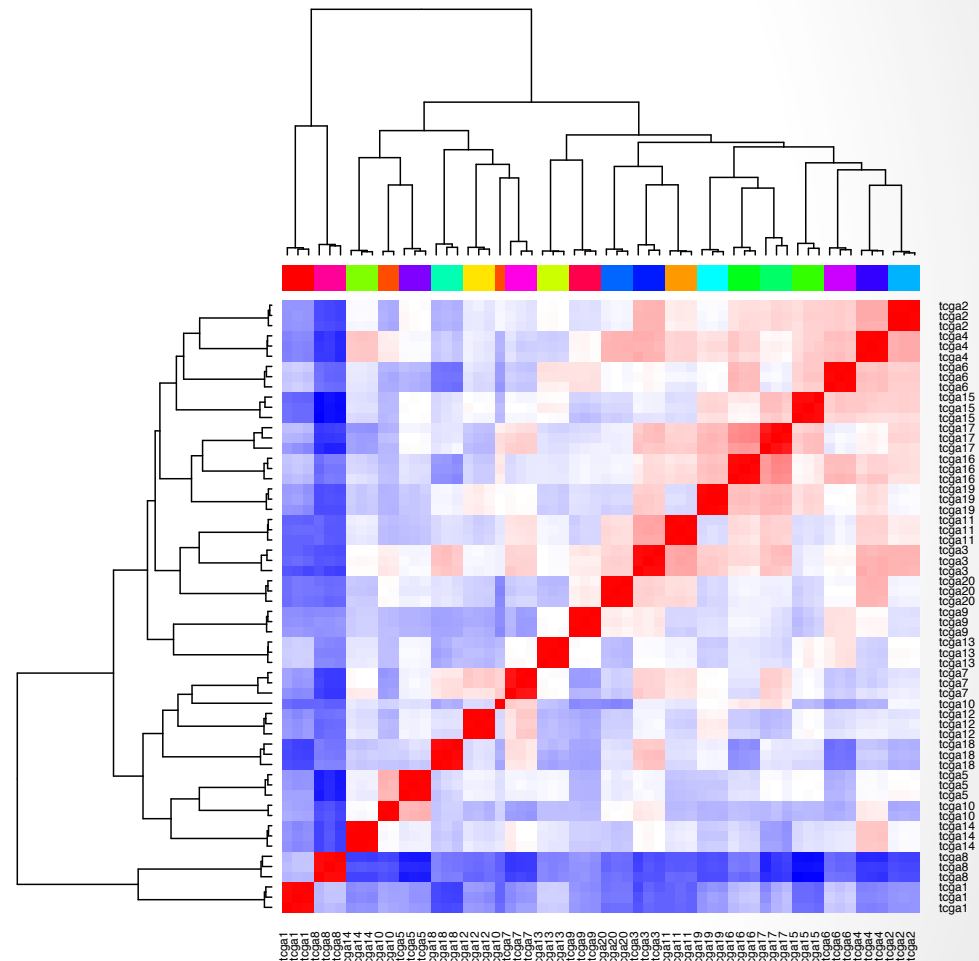
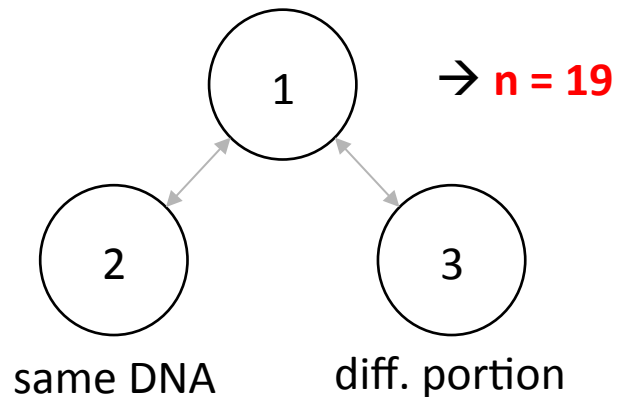
THE CANCER GENOME ATLAS 



# Statistical difficulties

- t-test (and other statistics) detect *common* differences
- no statistical test for  $n = 1$

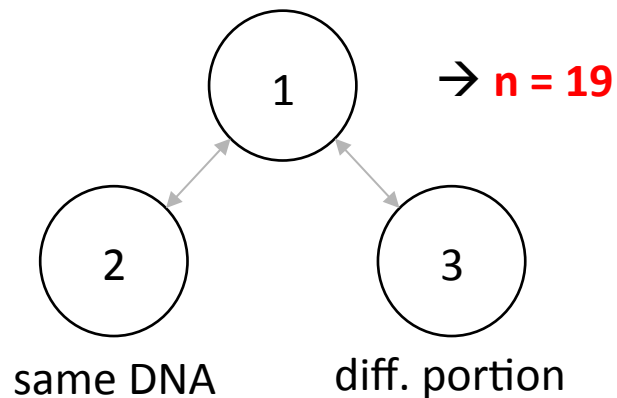
## estimation of technical noise



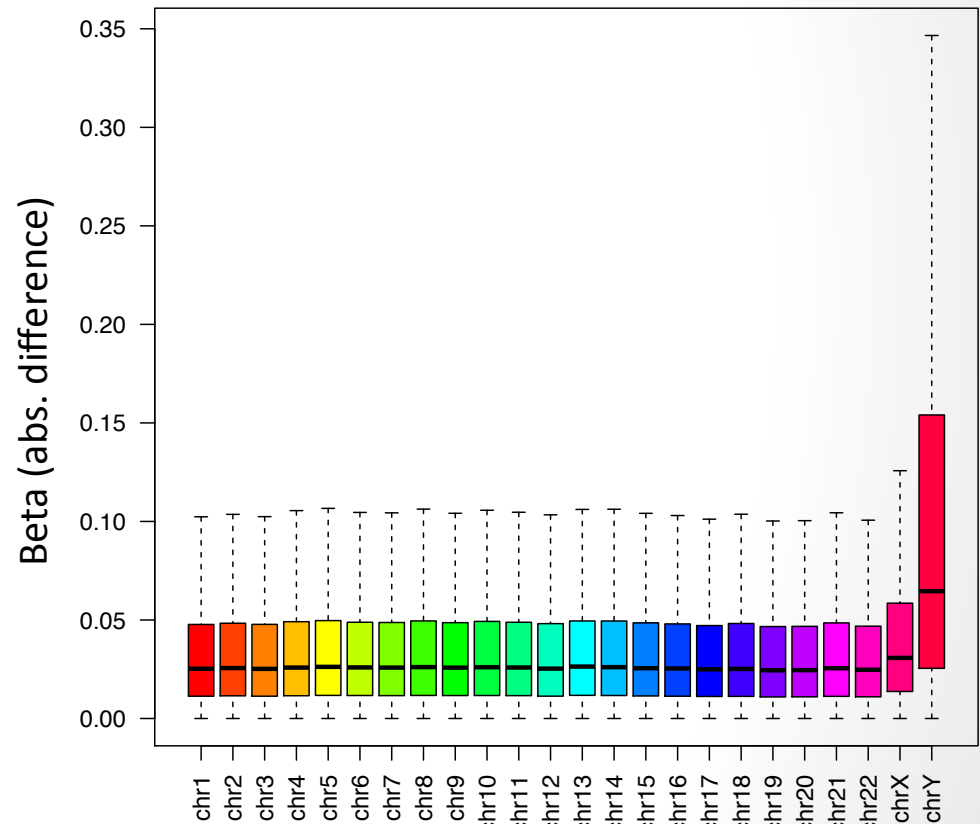
# Statistical difficulties

- t-test (and other statistics) detect *common* differences
- no statistical test for  $n = 1$

## estimation of technical noise



THE CANCER GENOME ATLAS 

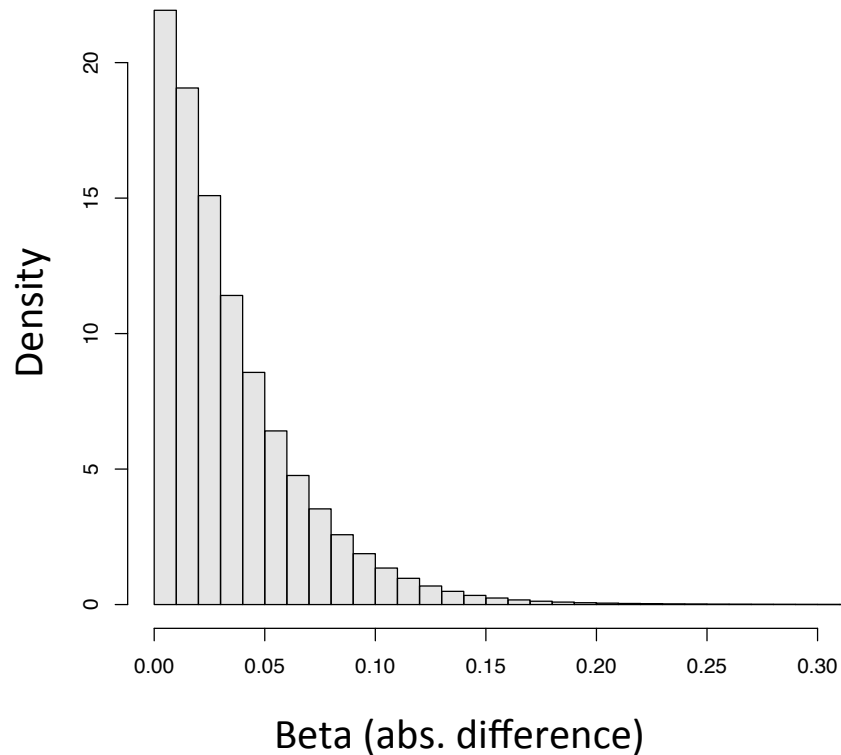




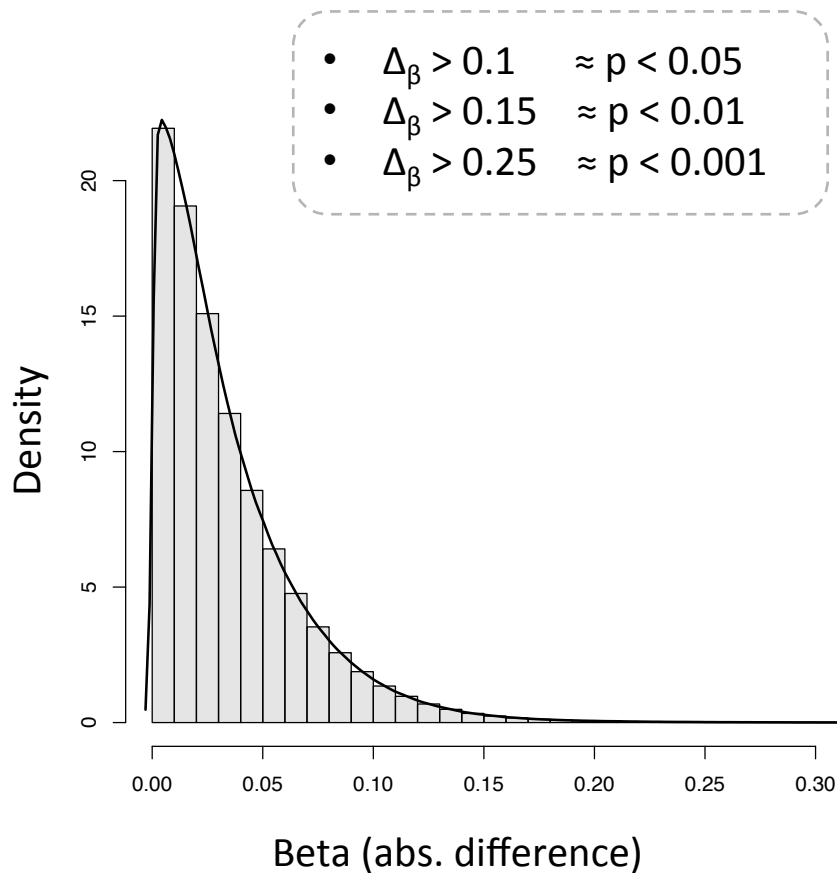
---

# Methylation changes depend on cancer type

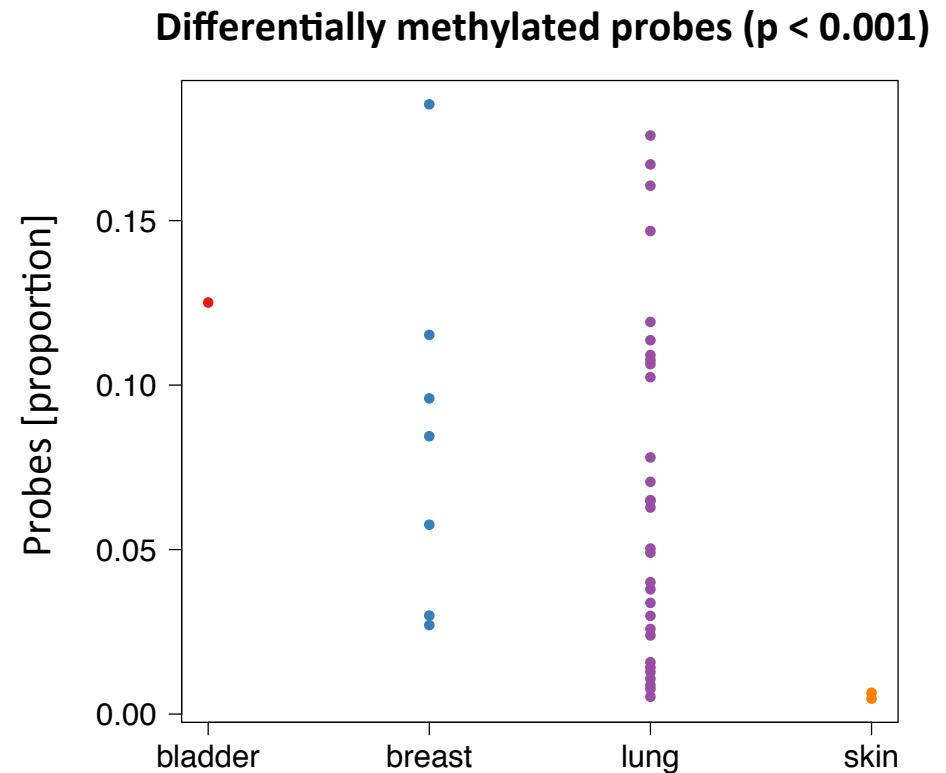
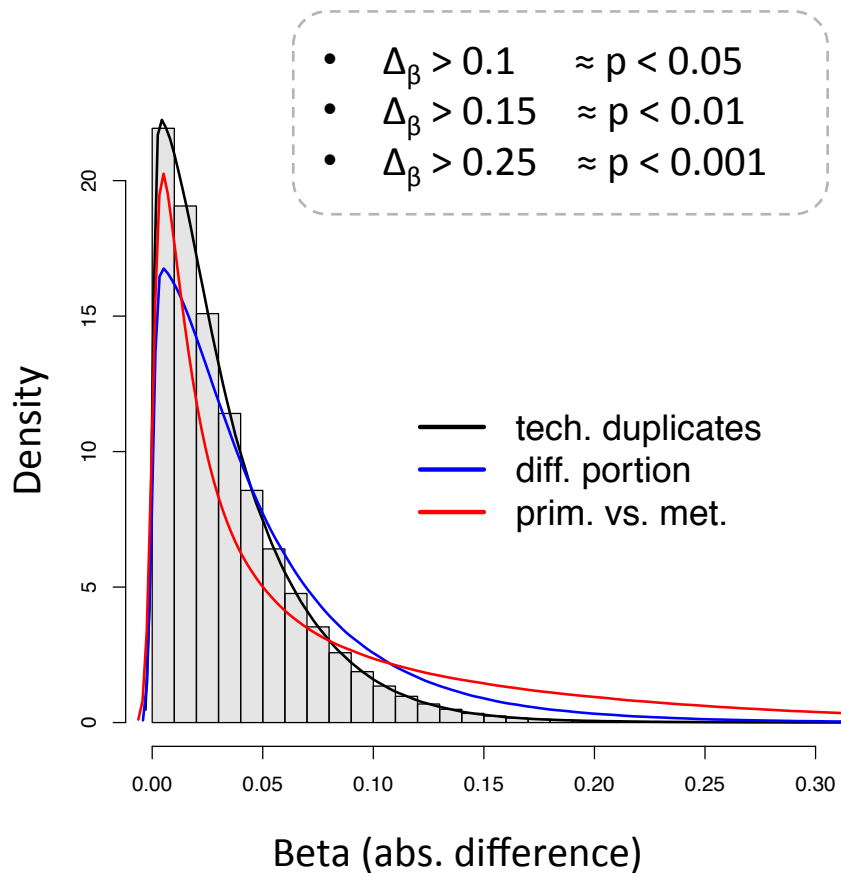
---



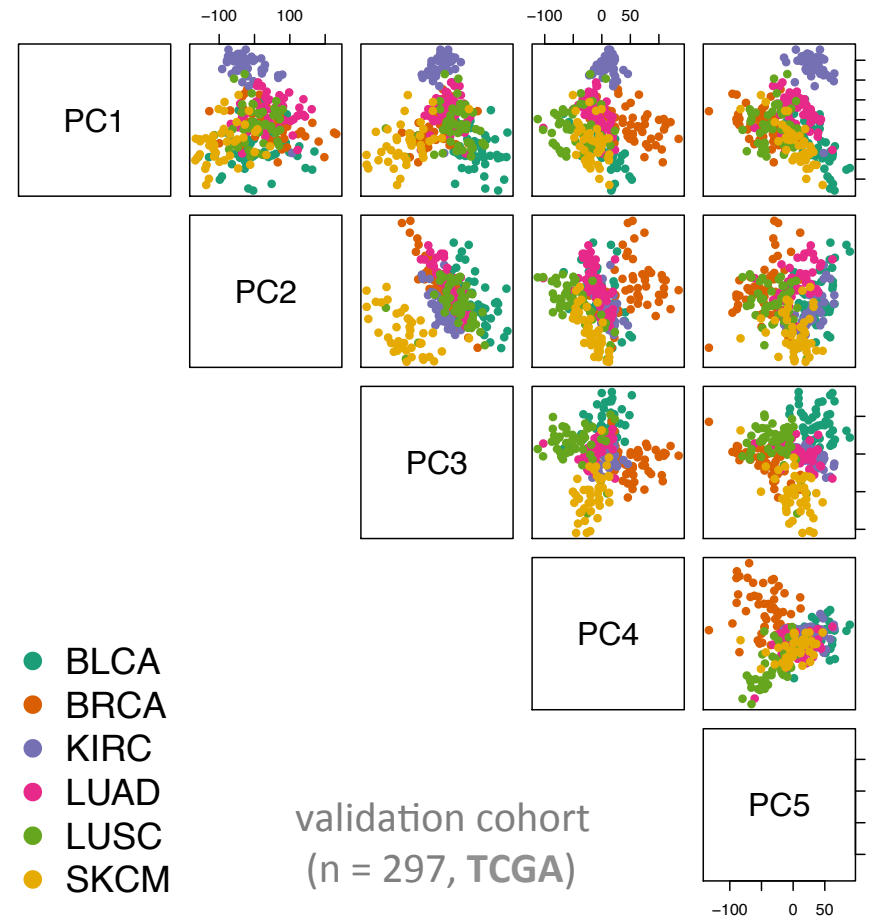
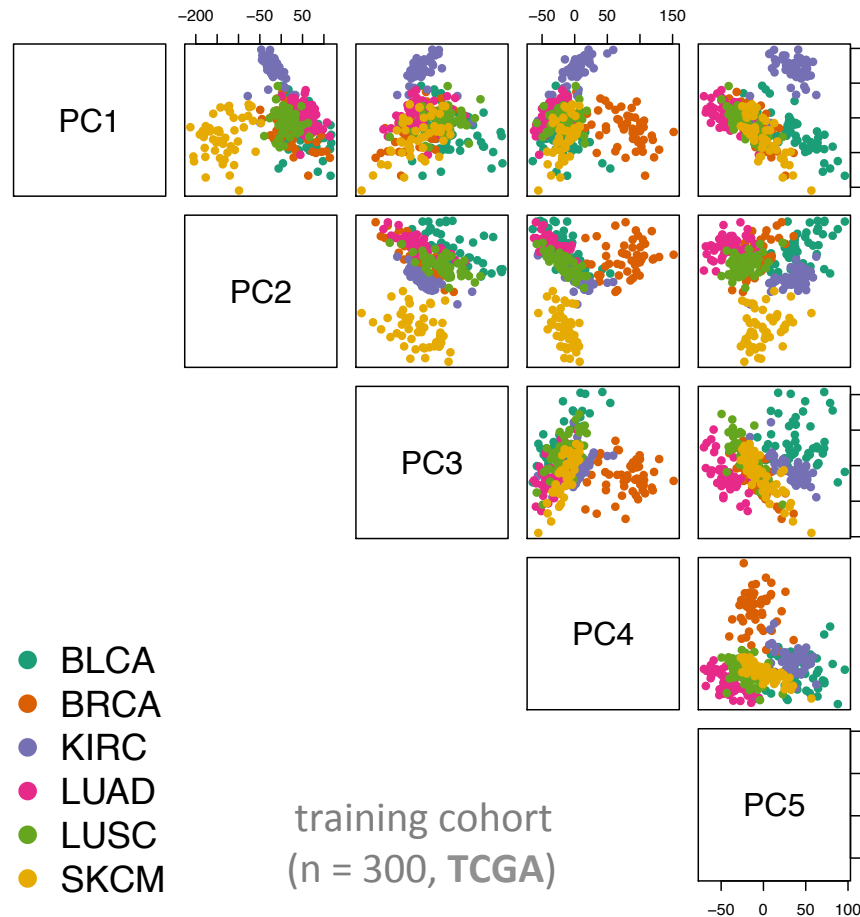
# Methylation changes depend on cancer type



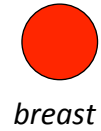
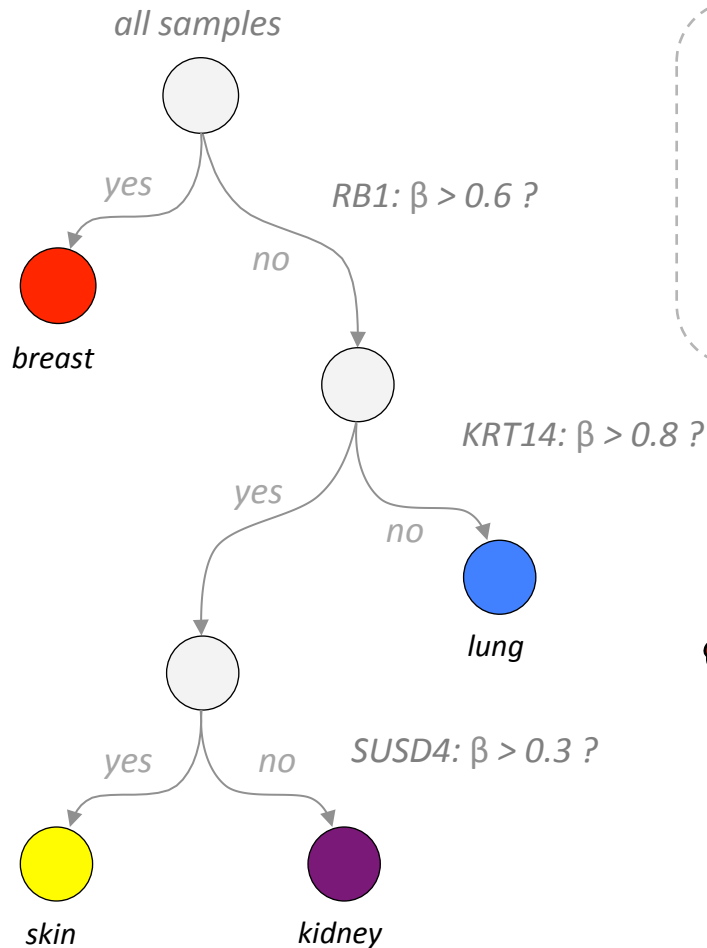
# Methylation changes depend on cancer type



# Tissue identity dictates methylation pattern



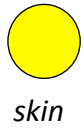
# Random forests



breast



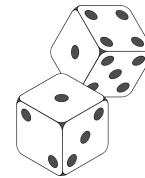
kidney



skin



lung

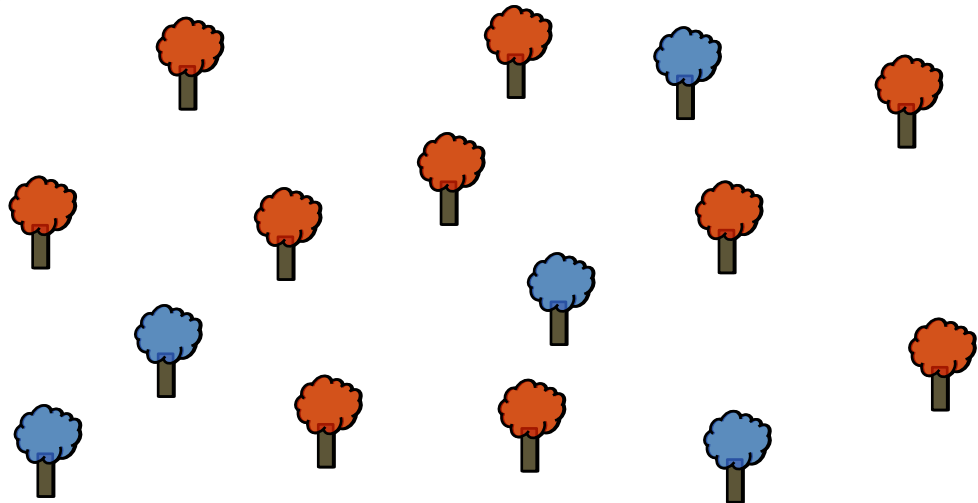


## samples

- once per tree
- w/ replacement

## probes

- every node
- $\sqrt{n}$  ( $\approx 700$ )

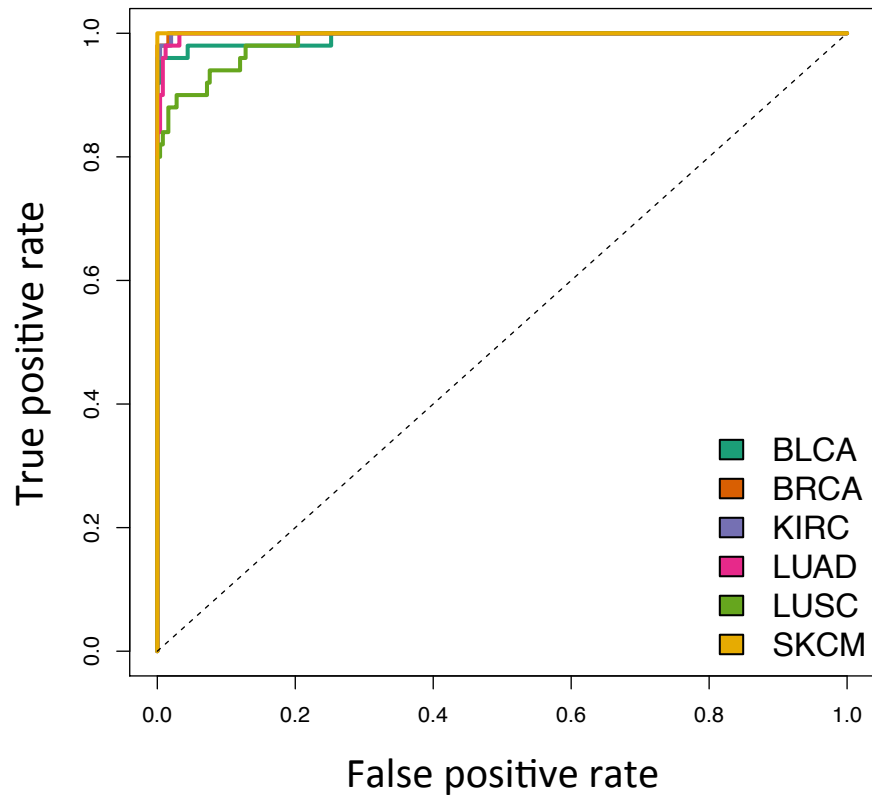


# Classification works well in training cohort

Predicted	True class					
	BLCA	BRCA	KIRC	LUAD	LUSC	SKCM
BLCA	45	0	0	1	1	0
BRCA	1	50	0	0	1	0
KIRC	1	0	49	0	0	0
LUAD	0	0	0	49	2	0
LUSC	3	0	1	0	46	1
SKCM	0	0	0	0	0	49

Class	Sens.	Spec.	PPV	NPV
BLCA	0.9	0.99	0.96	0.98
BRCA	1	0.99	0.96	1
KIRC	0.98	1	0.98	1
LUAD	0.98	0.99	0.96	1
LUSC	0.92	0.98	0.9	0.98
SKCM	0.98	1	1	1

# Classification works well in training cohort



Class	Sens.	Spec.	PPV	NPV
BLCA	0.9	0.99	0.96	0.98
BRCA	1	0.99	0.96	1
KIRC	0.98	1	0.98	1
LUAD	0.98	0.99	0.96	1
LUSC	0.92	0.98	0.9	0.98
SKCM	0.98	1	1	1

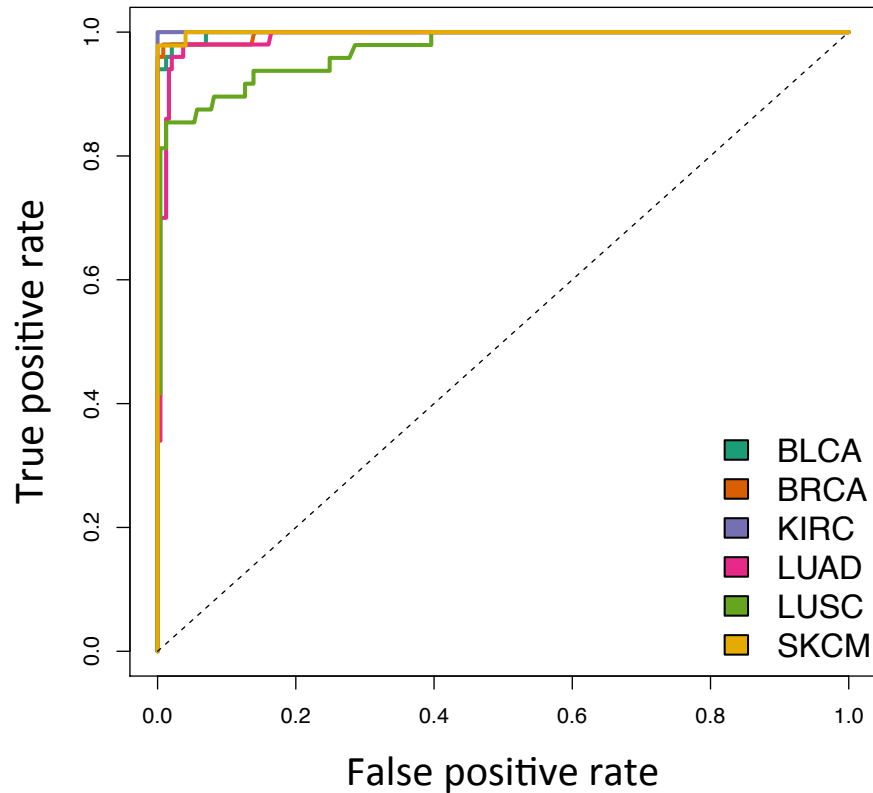
# Performance is similar in validation cohort

Predicted	True class					
	BLCA	BRCA	KIRC	LUAD	LUSC	SKCM
BLCA	47	0	0	0	1	0
BRCA	0	48	0	0	0	0
KIRC	0	0	49	0	0	0
LUAD	0	0	0	44	4	0
LUSC	3	2	0	6	43	1
SKCM	0	0	0	0	0	45

Class	Sens.	Spec.	PPV	NPV
BLCA	0.94	1	0.98	0.99
BRCA	0.96	1	1	0.99
KIRC	1	1	1	1
LUAD	0.88	0.98	0.92	0.98
LUSC	0.9	0.95	0.78	0.98
SKCM	0.98	1	1	1



# Performance is similar in validation cohort

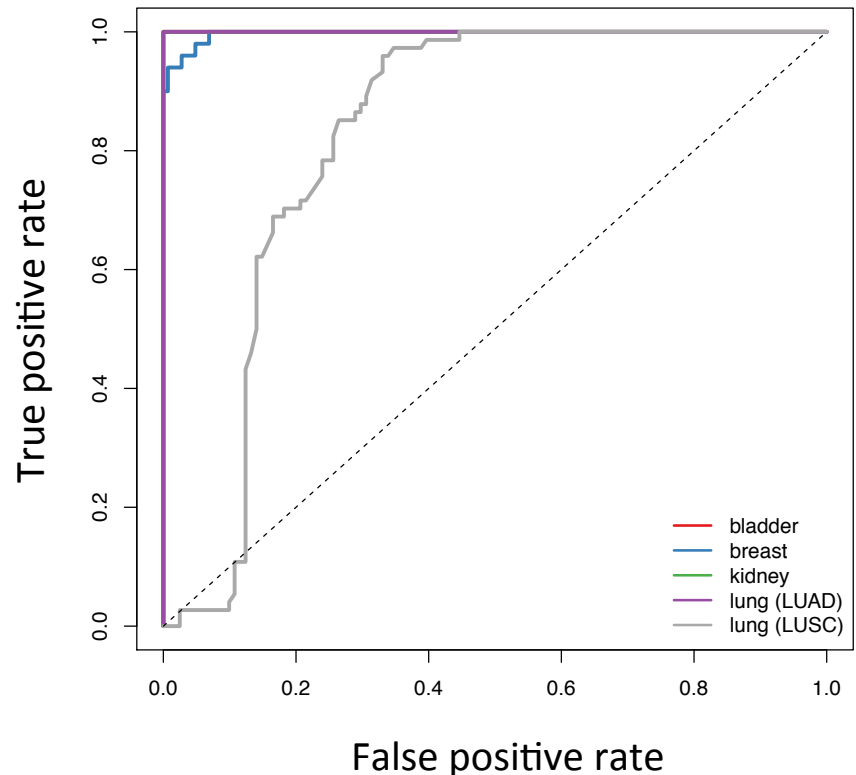


Class	Sens.	Spec.	PPV	NPV
BLCA	0.94	1	0.98	0.99
BRCA	0.96	1	1	0.99
KIRC	1	1	1	1
LUAD	0.88	0.98	0.92	0.98
LUSC	0.9	0.95	0.78	0.98
SKCM	0.98	1	1	1

# The method is also applicable to normal tissue

Predicted	True class			
	bladder	breast	kidney	lung
bladder	12	0	0	0
breast	0	43	0	0
kidney	6	7	50	0
lung	3	0	0	74

Class	Sens.	Spec.	PPV	NPV
bladder	0.57	1	1	0.95
breast	0.86	1	1	0.95
kidney	1	0.91	0.79	1
lung	1	0.98	0.96	1



→ LUAD more reliably identifies normal lung

Training cohort (TCGA)

6 entities  
300 primary tumors

**(1) train classifier**



Random  
forest  
classifier

**(3) diagnostic precision**

Validation cohorts (TCGA)

Validation cohort 1

6 entities  
293 primaries

Validation cohort 2

5 organs  
193 normal tissue controls

**(2) performance evaluation**

Brain metastases data set

**Neuropathology**

23 primaries  
45 metastases (brain)

**Division of Neurosurgery  
& Thoraxklinik**

17 primaries, 20 metastases  
(19x brain, 1x adrenal)

# Classification is applicable to brain metastases

Predicted	True class			
	bladder	breast	lung	skin
bladder	2	0	0	0
breast	0	19	2	0
lung	0	4	63	0
skin	0	0	0	15

Class	Sens.	Spec.	PPV	NPV
bladder	1	1	1	1
breast	0.83	0.98	0.9	0.95
lung	0.97	0.9	0.94	0.95
skin	1	1	1	1

Predicted	True class			
	bladder	breast	lung	skin
bladder	1	0	0	0
breast	0	13	2	0
lung	0	3	32	0
skin	0	0	0	13

Class	Sens.	Spec.	PPV	NPV
bladder	1	1	1	1
breast	0.81	0.96	0.87	0.94
lung	0.94	0.9	0.91	0.93
skin	1	1	1	1



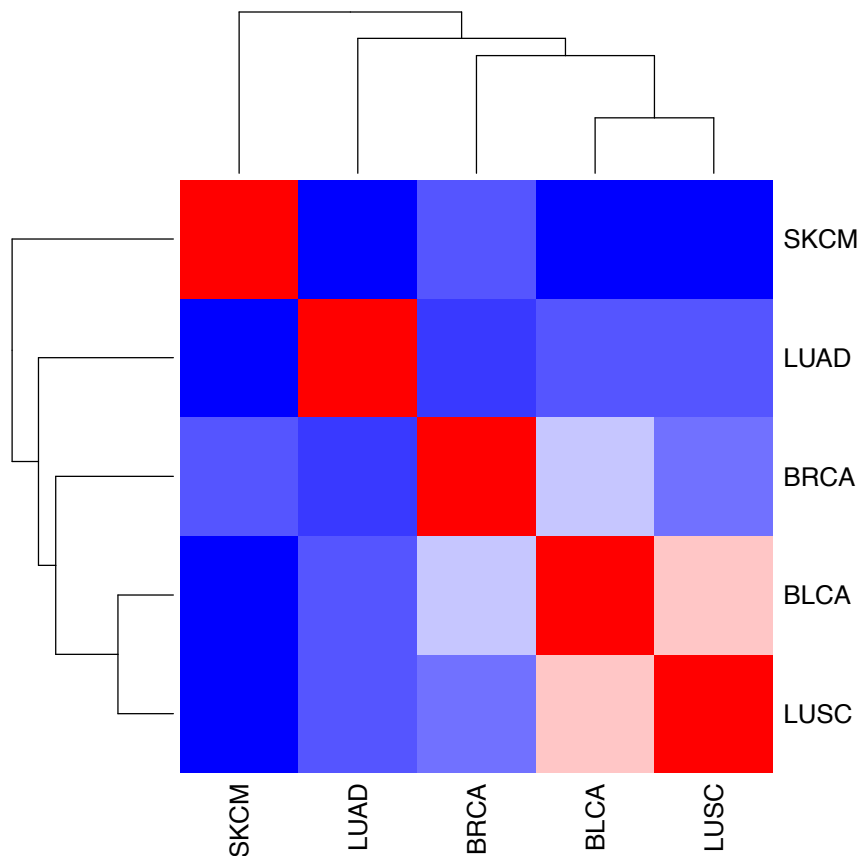
incl. primaries



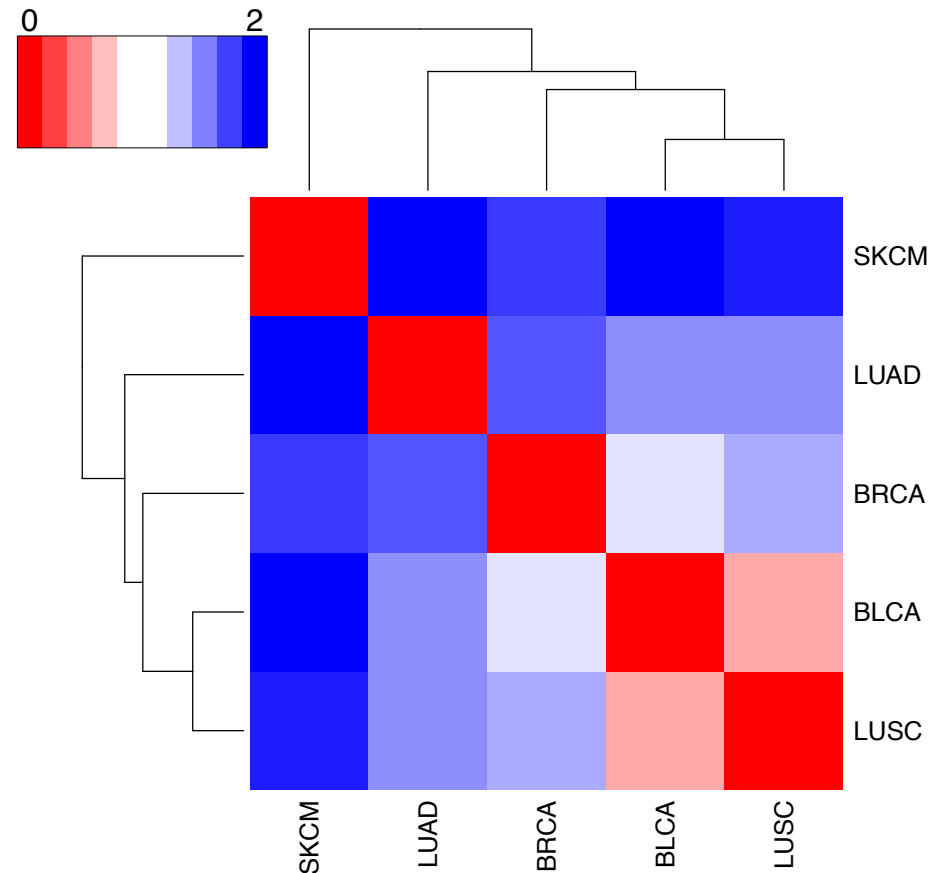
brain metastases

# LUSC and BLCA score similar

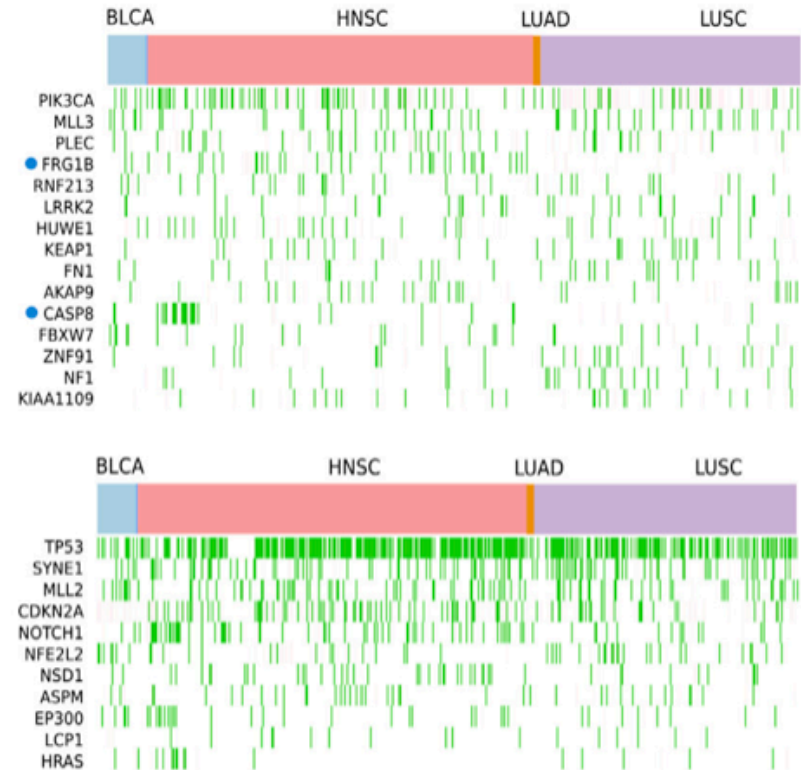
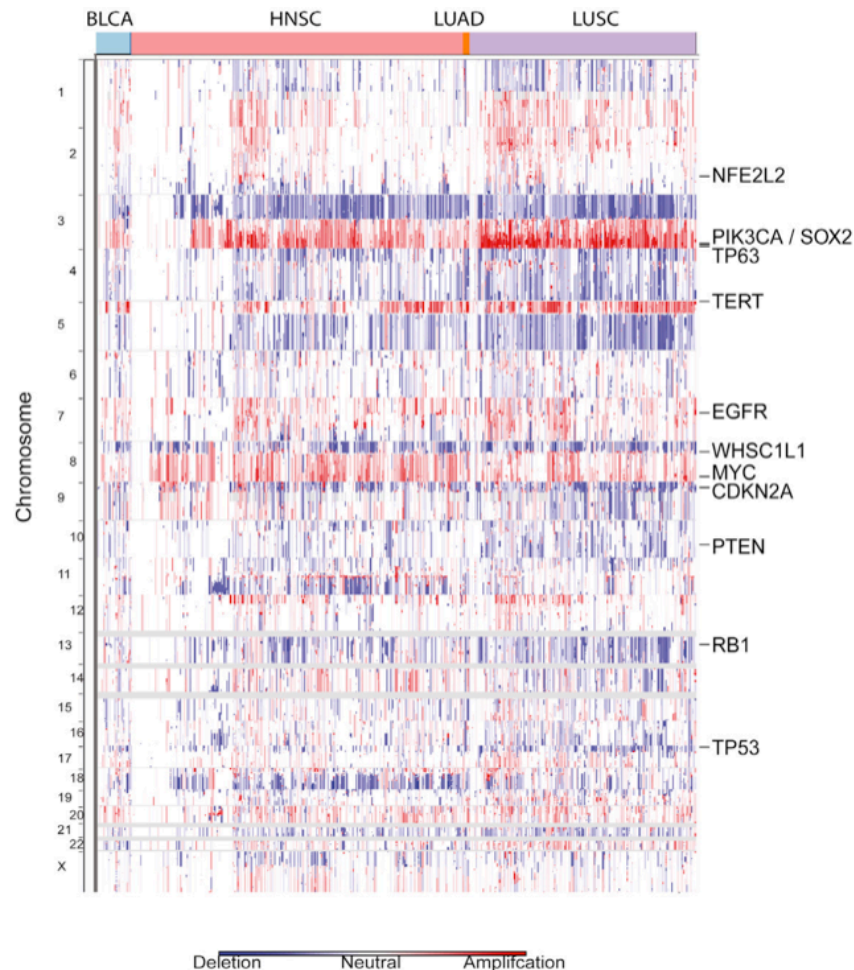
all samples incl. primaries



brain metastases



# This was also noted in a pan-cancer analysis



*Hoadley et al., Cancer Cell, 2014*

---

# Summary and outlook

---

- differences in methylation
  - might depend on cancer type
  - still, many CpGs do *not* change
  - in-depth analysis of paired samples
- random forest classifier can reliably identify CUP cases
  - more entities & samples for brain mets
  - extend to extracranial mets
  - make available as online tool
- data for primary tumors is sufficient for classifier training
  - more entities & samples for primaries
  - substratification within entities

# Acknowledgements

## **Neurosurgery, Heidelberg, Germany**

- Prof. Christel Herold-Mende
- Steffen Dettling, Andreas Mock, Carmen Rapp, Saskia Rösch, Rolf Warta
- Melanie Greibich, Hilde Göltzer, Farzaneh Kashfi

## **Neuropathology, Heidelberg, Germany**

- David Capper
- Prof. Andreas von Deimling

## **Thoraxklinik, Heidelberg, Germany**

- Michael Meister & colleagues

## **DKFZ, Heidelberg, Germany**

- David Jones

## **University of Vienna, Austria**

- Matthias Preusser
- Anna Sophie Berghoff

## **University of Toronto, Canada**

- Paul Kongkham