

Single-nucleus RNA-seq of frozen hepatoblastoma tissues

Amélie ROEHRIG, 2nd year PhD student

« Functional Genomics of Solid Tumors » (J. Zucman-Rossi). *Supervisor: Eric Letouzé*

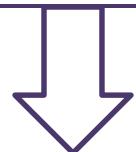
CELL seminar, 13th of April 2021

HEPATOBLASTOMA (HB)

HEPATOBLASTOMA

[3,4]

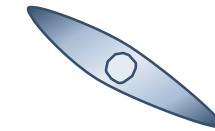
- Most frequent hepatic tumor occurring in children (< 5 years old)
 - No underlying hepatic pathology
 - Treatment: **cisplatin + surgery**



Epithelial histotype



Mesenchymal histotype



Fetal

~ Late liver development
Low proliferation
Good prognosis

Embryonal

~ Early liver development
Strong proliferation
Relapses, metastases
Bad prognosis

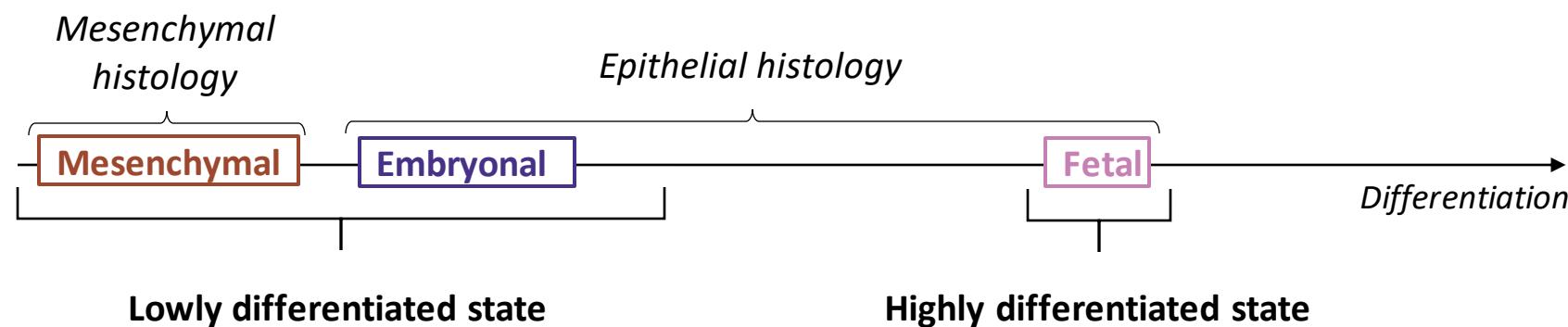
HEPATOBLASTOMA (HB)

HEPATOBLASTOMA

[3,4]

- Most frequent hepatic tumor occurring in children (< 5 years old)
 - No underlying hepatic pathology
 - Treatment: **cisplatin + surgery**

3 main histotypes related to **differentiation levels**



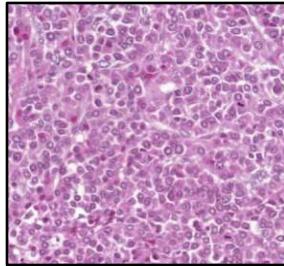
HEPATOBLASTOMA (HB)

HEPATOBLASTOMA

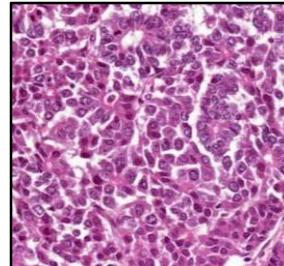
[3]

- Most frequent hepatic tumor occurring in children (< 5 years old)
 - No underlying hepatic pathology
 - Treatment: **cisplatin + surgery**

Strong phenotypic **heterogeneity**



70% fetal
30% embryonal



45% fetal
50% embryonal
+ fibrosis

HEPATOBLASTOMA (HB)

What is known at the molecular level?

Molecular alterations: ^[5,6,7]

- Activation of **Wnt-signaling pathway (CTNNB1)**
- Some (rare) mutations: NFE2L2, TERT

} *Simple genomic landscape*

Transcriptomic/epigenetic classifications:

- Cairo *et al* (2008): 2 transcriptomic clusters **C1 (fetal), C2 (embryonal)**
- Carrillo-Reixach *et al* (2020): 2 epigenomic clusters **Epi-CA (fetal), Epi-CB (embryonal)**

} *Lack of integrated genomic analysis on a broad cohort*

PROBLEMATICS

Main goal: understand the molecular heterogeneity of HB cells

- Molecular heterogeneity across the **tumor** and along **disease progression**
- Why do some HB cells **resist to chemotherapy?**

Hypothesis: embryonal HB resist to chemotherapy → relapses + metastases

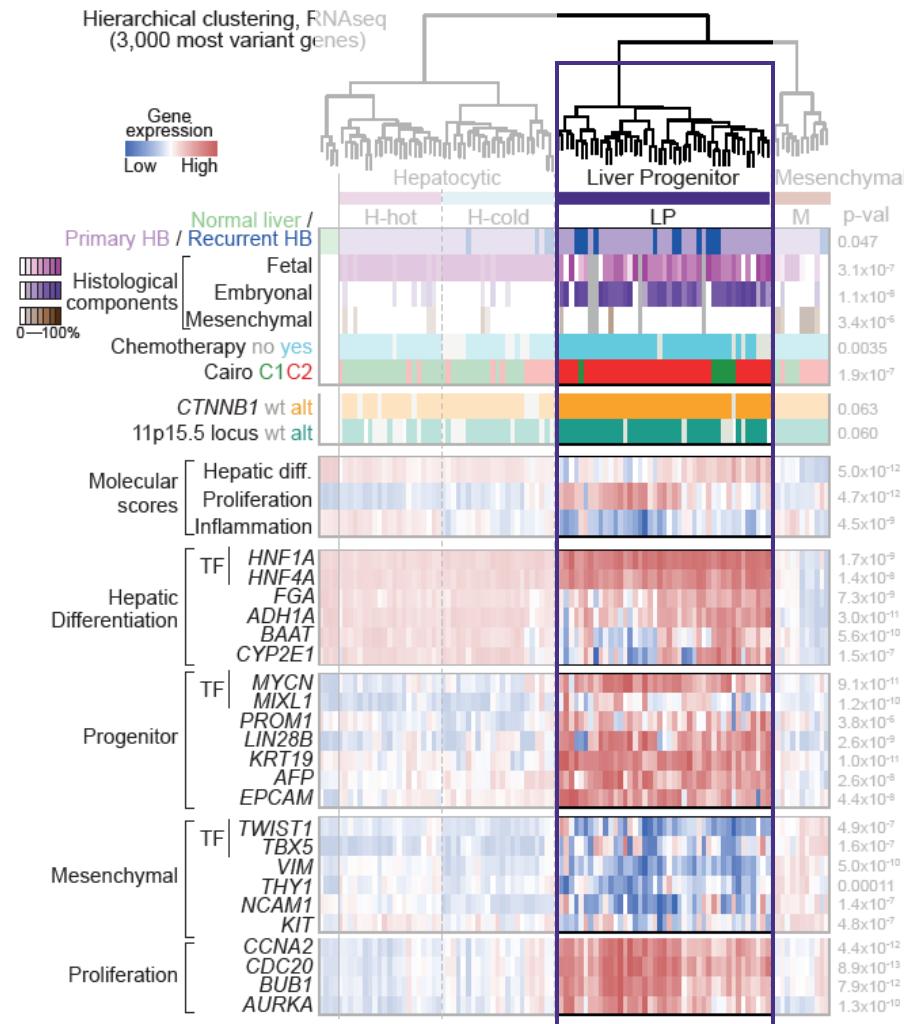


Integrated analysis of a large HB cohort:

Whole-Genome Sequencing (WGS)	n=48
Whole-Exome Sequencing (WES)	n=52
RNA-sequencing (RNA-seq)	n=100
Reduced Representation Bisulfite Sequencing (RRBS)	n=84
Single-nucleus RNA-sequencing (snRNA-seq)	n=4 matched samples

Hirsch et al., Cancer Discov, in press

BULK TRANSCRIPTOMIC ANALYSES OF HB



RNA-seq of 100 HB and 4 non-tumor livers → identification of **3 transcriptomic clusters** that resemble the histological HB subtypes



Mesenchymal HB

Mesenchymal histology

Differentiation: -

Mesenchymal TF* (VIM, TWIST1)

Liver Progenitor HB

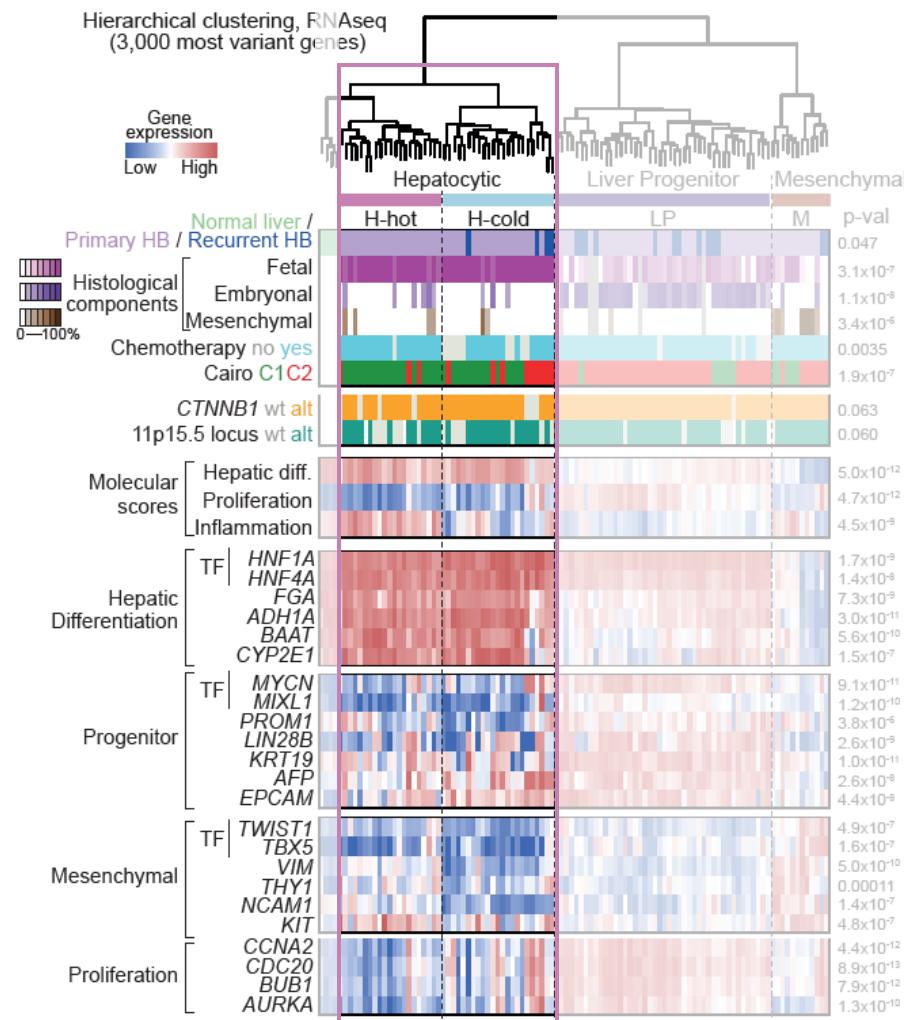
Embryonal histology

Differentiation: -/+

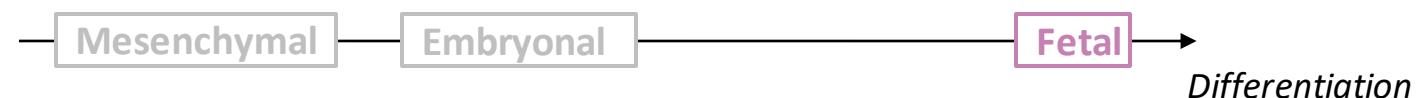
Proliferation ++
Progenitor TF (MYCN, PROM1)

= the HB subtype that resists to chemotherapy

BULK TRANSCRIPTOMIC ANALYSES OF HB



RNA-seq of 100 HB and 4 non-tumor livers → identification of **3 transcriptomic clusters** that resemble the histological HB subtypes



Mesenchymal HB

Mesenchymal histology

Differentiation: -

Mesenchymal TF* (VIM, TWIST1)

Liver Progenitor HB

Embryonal histology

Differentiation: -/+

Progenitor TF (MYCN, PROM1)

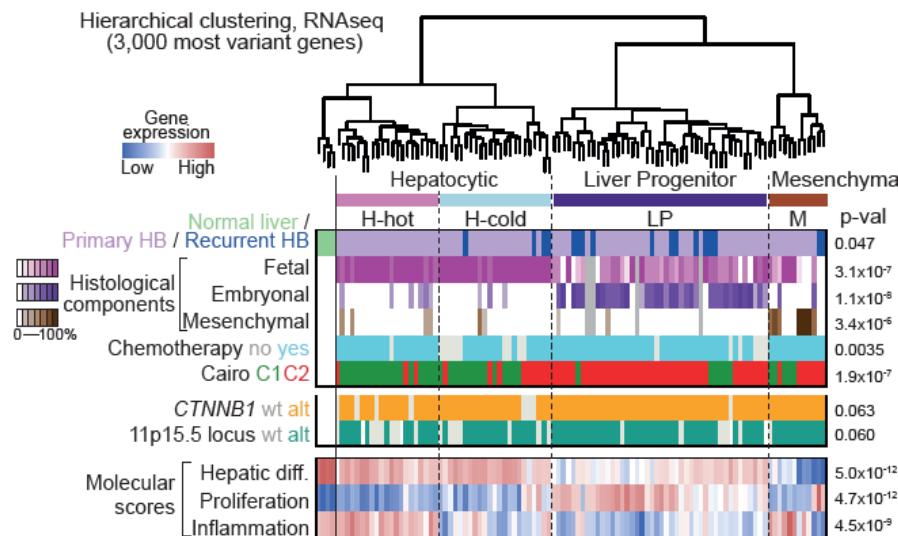
Hepatocytic HB

Fetal histology

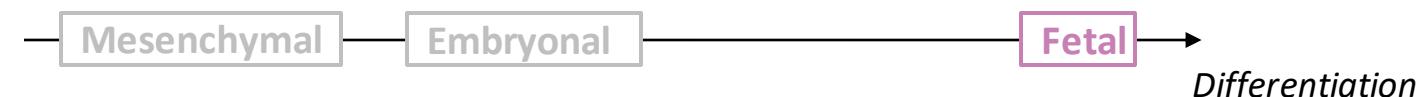
Normal-like

Differentiation: +

BULK TRANSCRIPTOMIC ANALYSES OF HB



RNA-seq of 100 HB and 4 non-tumor livers → identification of **3 transcriptomic clusters** that resemble the histological HB subtypes



Mesenchymal HB

Mesenchymal histology
Differentiation: -
Mesenchymal TF* (VIM, TWIST1)

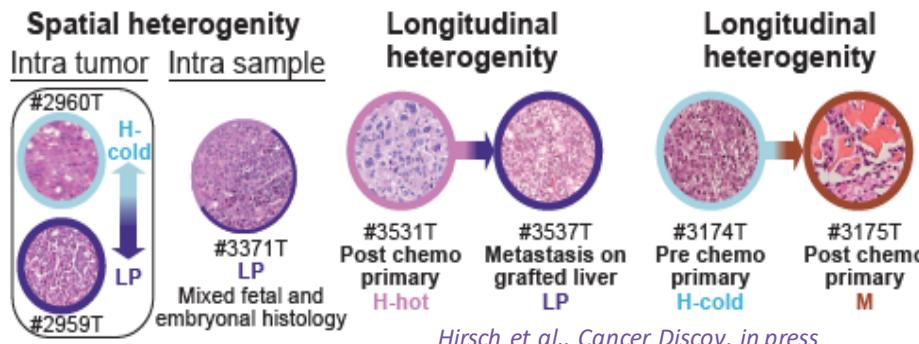
Liver Progenitor HB

Embryonal histology
Differentiation: -/+
Progenitor TF (MYCN, PROM1)

Hepatocytic HB

Fetal histology
Normal-like
Differentiation: +
Differentiation TF (HNF4A, CYP2E1)

Main problem to characterize and cure HB:
their strong heterogeneity and plasticity



BULK TRANSCRIPTOMIC ANALYSES OF HB

Goal: investigate HB heterogeneity and plasticity

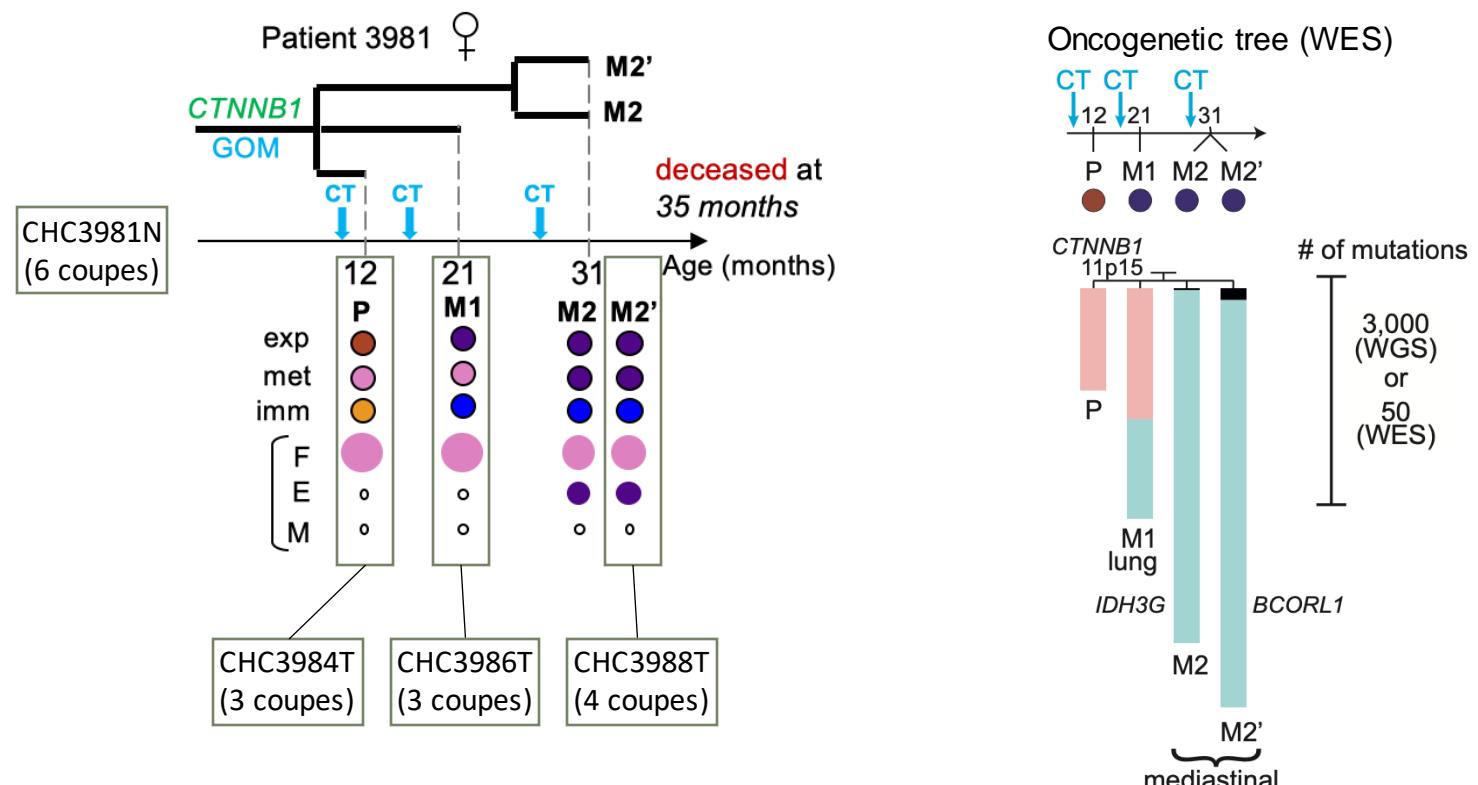
- ➔ Study the transitions between hepatocytic / liver progenitor / mesenchymal components
- ➔ Characterize HB tumor evolution

Solution: single-nucleus RNA-seq (snRNAseq) of frozen HB tumors to investigate plasticity at the single-cell level

SN-RNASEQ OF FROZEN HB

→ 1 patient (which combines all three histologies/clusters), 4 samples: 1 **non-tumor** sample, 1 **primary**, 2 **metastases**

Sel scRNAseq	CHC-ID	Nb.coups	Patient.ident	Type	Gender	Age.at.surgery	Age_at_diag	Histological.Diagn	Serie 2020	RNAseq	sei.WGS	se.WES	se.RBBS	RNAseq	clust	mol.WISP.E	mol.WISP.F1	mol.WISP.F2	mol.WISP.M	immune	HB	CTNNB1	detz locus 11p15 merged	cisplat.sig.nm cisplat.sig.grp
1	CHC3981N	6	3981	N	F	12.2	9.13347023	HB	NGS_Validation	no	no	yes	no									Small_delet	wt	
	CHC3982T	2	3981	T	F	12.2	9.13347023	HB	Validation	no	no	no	no									Small_delet	wt	
	CHC3983T	3	3981	T	F	12.2	9.13347023	HB	Validation_exc	no	no	no	no									Small_delet	wt	22 no
2	CHC3984T	3	3981	T	F	12.2	9.13347023	HB	NGS	yes	no	yes	yes	M	0.53	0.44	0.00	0.02	hot		Small_delet	wt		
	CHC3985T	4	3981	T	F	12.2	9.13347023	HB	Validation_red	no	no	no	no	E	0.80	0.18	0.00	0.03	cold		Small_delet	wt	27 yes	
3	CHC3986T	3	3981	M	F	21.2	9.13347023	HB.META	NGS	yes	no	yes	yes	E	0.72	0.28	0.00	0.00	cold		Small_delet	wt	152 yes	
	CHC3987T	2	3981	M	F	31	9.13347023	HB.META	NGS	yes	no	yes	yes	E	0.74	0.24	0.02	0.00	cold		Small_delet	wt	168 yes	
4	CHC3988T	4	3981	M	F	31	9.13347023	HB.META	NGS_redondar	yes	no	yes	yes								Small_delet	wt		



SN-RNASEQ OF FROZEN HB: QUALITY CONTROL

FastQ files (Integragen)

CellRanger count (include introns)

UMI counts matrix

	AAACCTGAGCCAGTTT	AAACCTGAGTGGAGTC	AAACCTGCAGATCTGT	AAACCTGGTCCCCACA	AAACGGGAGCGAGAAA
FTL	1203	305	653	2286	1295
MT-CO2	674	541	354	749	1204
MT-CO1	834	797	429	696	953
MT-CO3	618	285	341	523	867
FTH1	553	301	301	818	489
MT-ND4	472	285	227	384	660
RPS2	425	243	296	496	558
RPL13	491	388	215	352	511
MT-ND2	434	218	152	274	463
MT-ATP6	250	167	135	291	412
MT-ND1	379	225	165	281	454
MT-CYB	201	92	119	212	319
EEF1A1	459	134	159	125	223
APOA2	268	158	79	287	398
RPLP1	275	122	154	297	261
RPL13A	297	174	148	182	296
RPS18	224	225	117	208	292
MALAT1	342	271	146	152	161
ALB	512	179	97	156	101
RPS19	195	131	152	201	262
APOA1	761	189	168	241	68

SN-RNASEQ OF FROZEN HB: QUALITY CONTROL

FastQ files (Integragen)

CellRanger count (include introns)

UMI counts matrix

Keep cells with:

> 500 genes detected

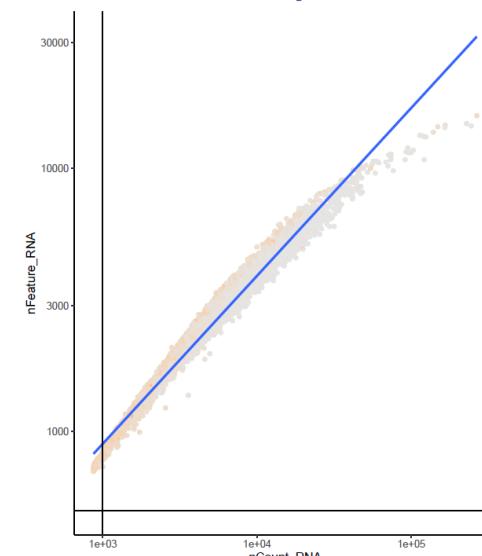
> 1000 UMI counts

< 5% mitochondrial reads

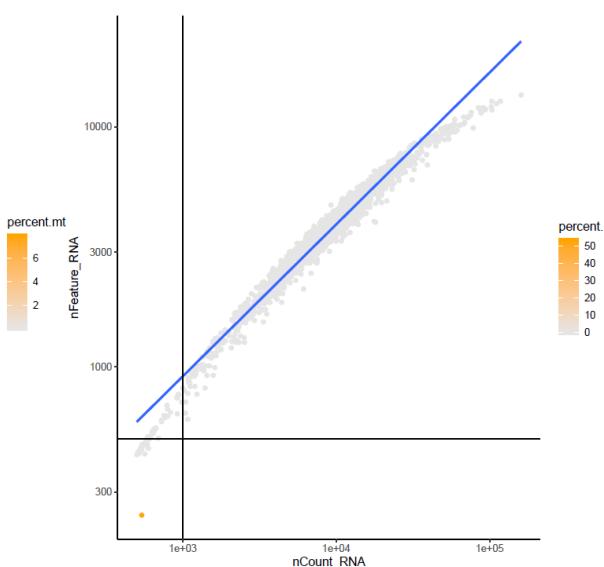
Keep genes with > 1 transcript in
> 3 non outlier cells

Remove mitochondrial + ERCC genes

Primary tumor

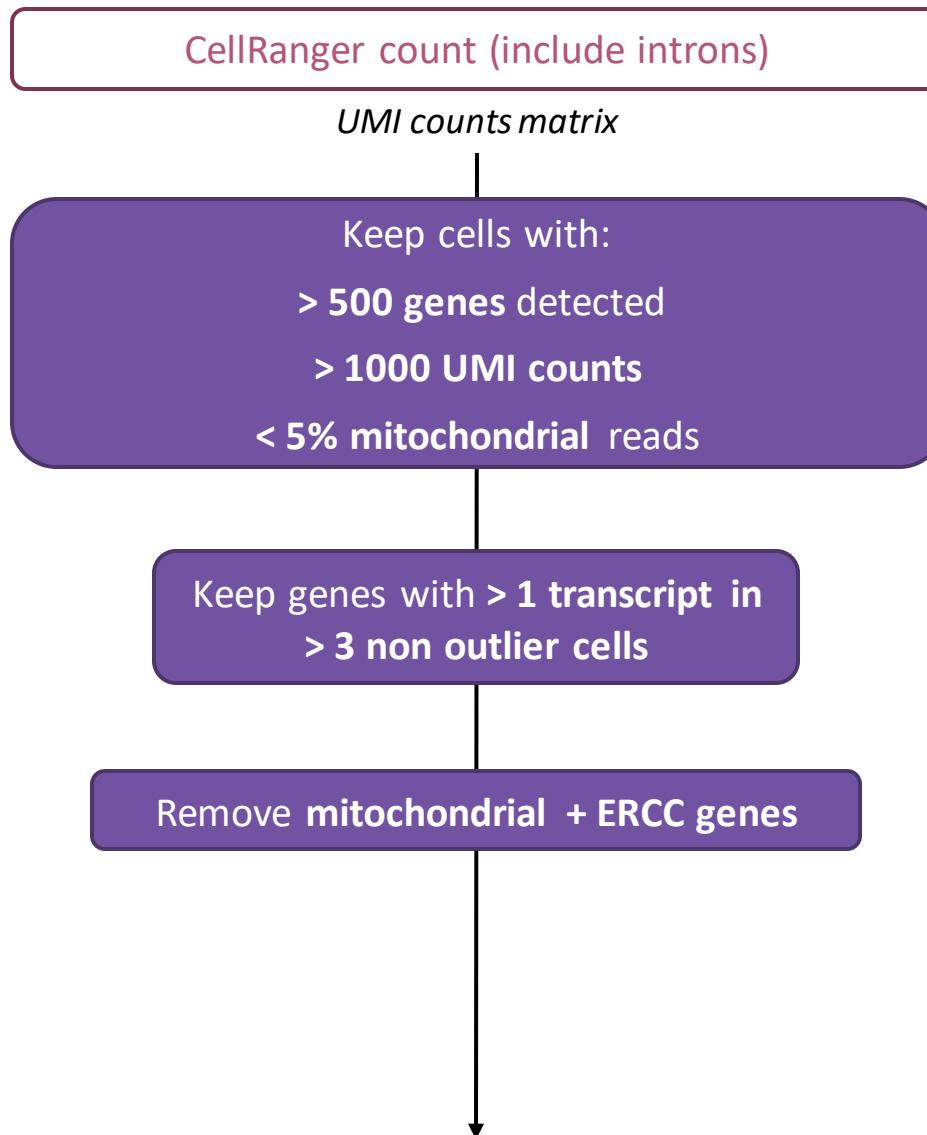


Metastasis #2



SN-RNASEQ OF FROZEN HB: QUALITY CONTROL

FastQ files (Integragen)



Before QC:

Sample	Nb of cells	Total nb of genes	Mean reads per cell	Median genes per cell
NT	9142	36601	31923	2042
Primary	5869	36601	54300	2334
Metastasis #1	3652	36601	88099	1052
Metastasis #2	2875	36601	75910	3446

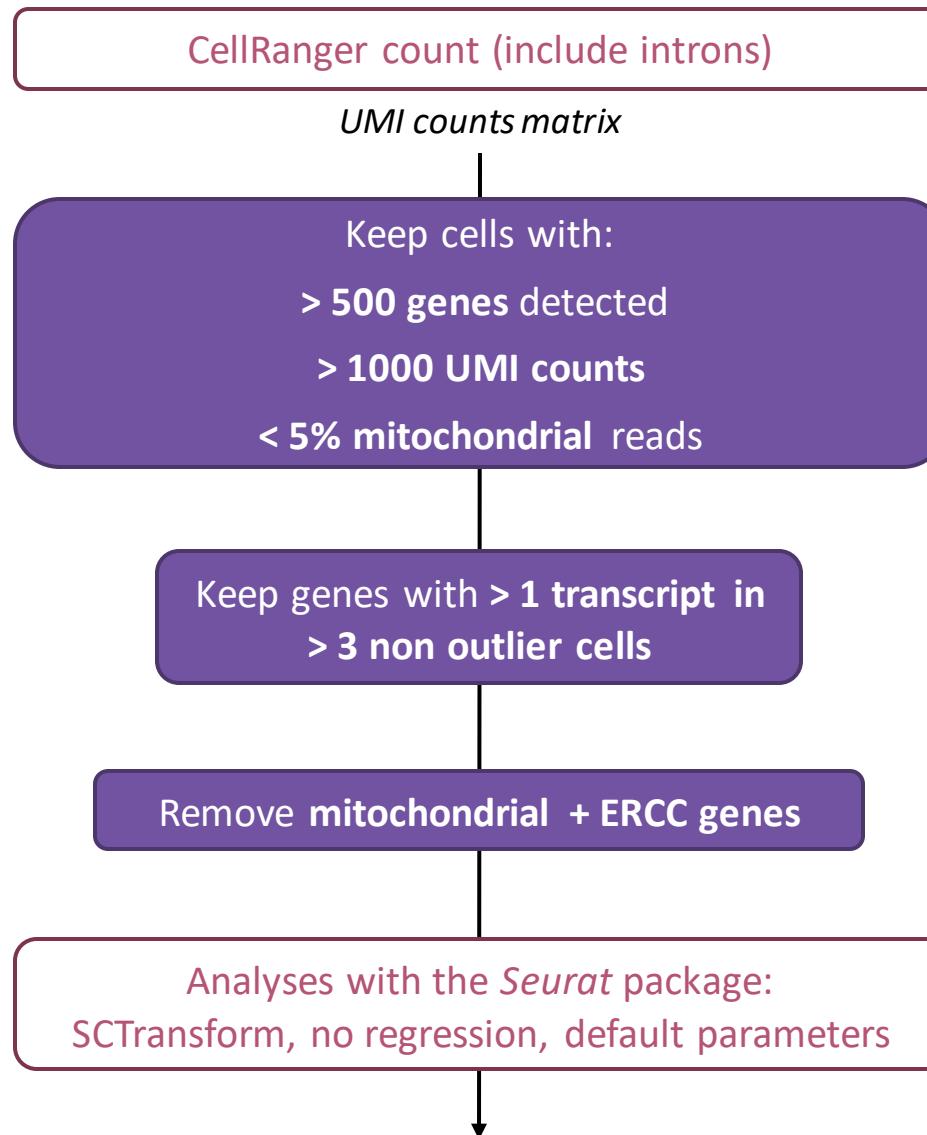
After QC:

Sample	Nb of cells	Total nb of genes
NT	9140	16470
Primary	5685	18384
Metastasis #1	2943	11401
Metastasis #2	2826	18294

Pb during experiments: more PCR cycles required
QC metrics not as good as the other samples

SN-RNASEQ OF FROZEN HB: QUALITY CONTROL

FastQ files (Integragen)



Before QC:

Sample	Nb of cells	Total nb of genes	Mean reads per cell	Median genes per cell
NT	9142	36601	31923	2042
Primary	5869	36601	54300	2334
Metastasis #1	3652	36601	88099	1052
Metastasis #2	2875	36601	75910	3446

After QC:

Sample	Nb of cells	Total nb of genes
NT	9140	16470
Primary	5685	18384
Metastasis #1	2943	11401
Metastasis #2	2826	18294

SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

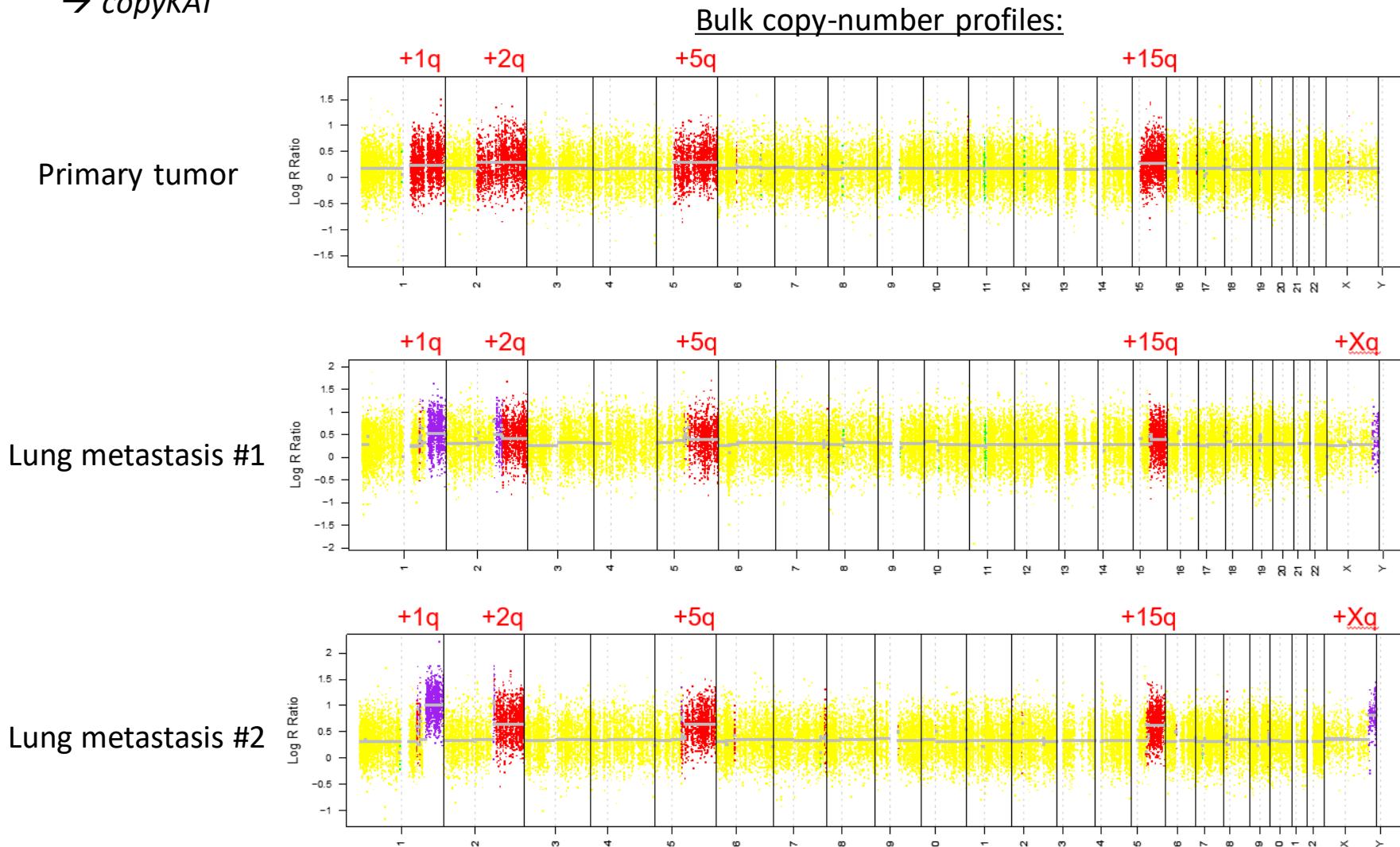
Goal: restrict to the tumor cells only using virtual copy-number alterations profiles

- *inferCNV*
- *copyKAT*

SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

Goal: restrict to the tumor cells only using virtual copy-number alterations profiles

- *inferCNV*
- *copyKAT*

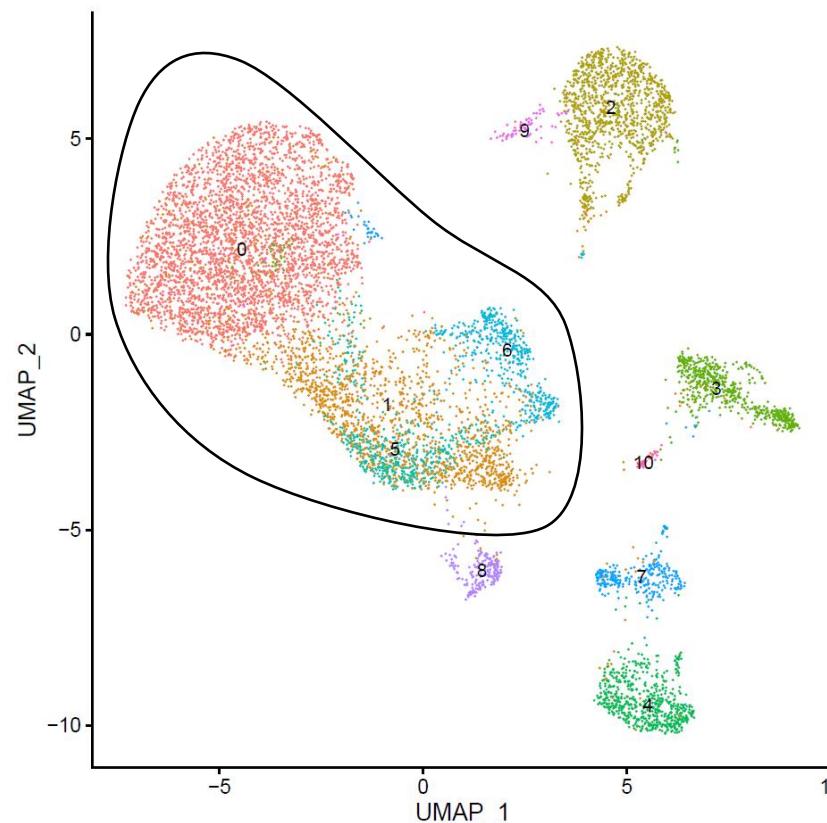


SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

Goal: restrict to the tumor cells only using virtual copy-number alterations profiles

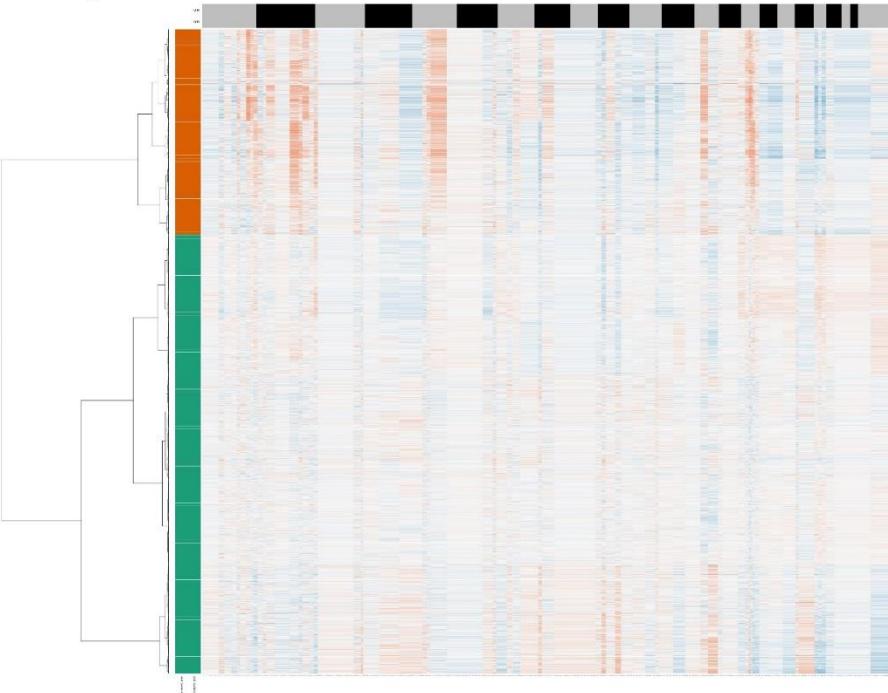
- *inferCNV*
- *copyKAT*

Define a set of **reference cells** → normal hepatocytes in the non-tumor sample



SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

CopyKAT (default parameters)

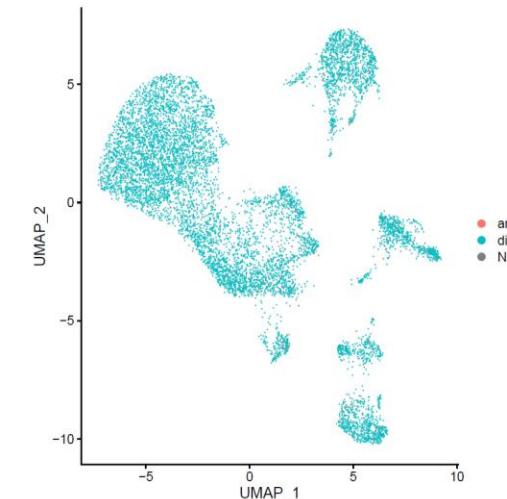


Too stringent... → inferCNV

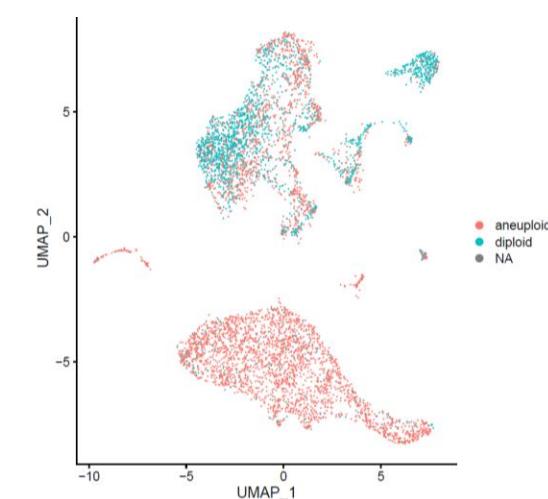
Aneuploid cells (tumoral)

Diploid cells (normal)

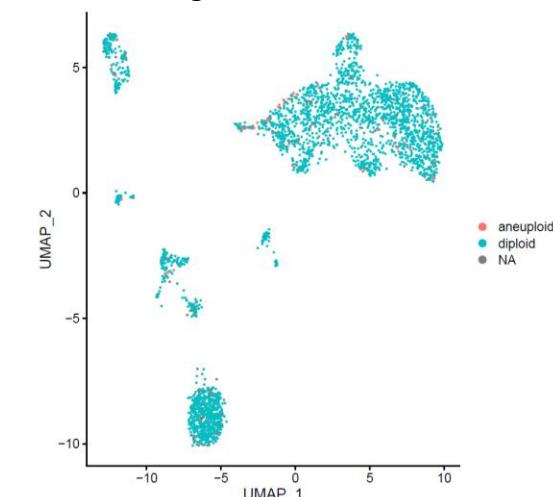
Non-tumor sample



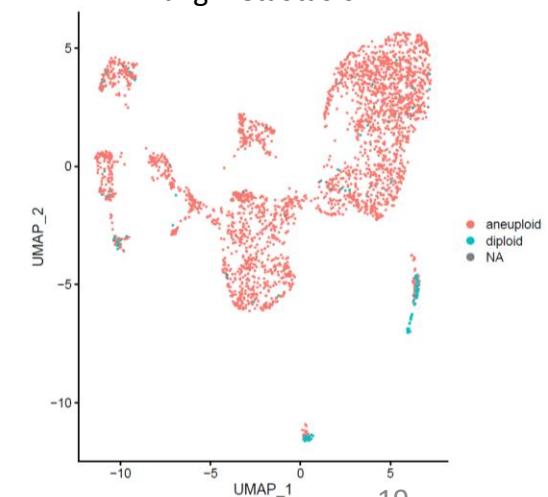
Primary tumor



Lung metastasis #1

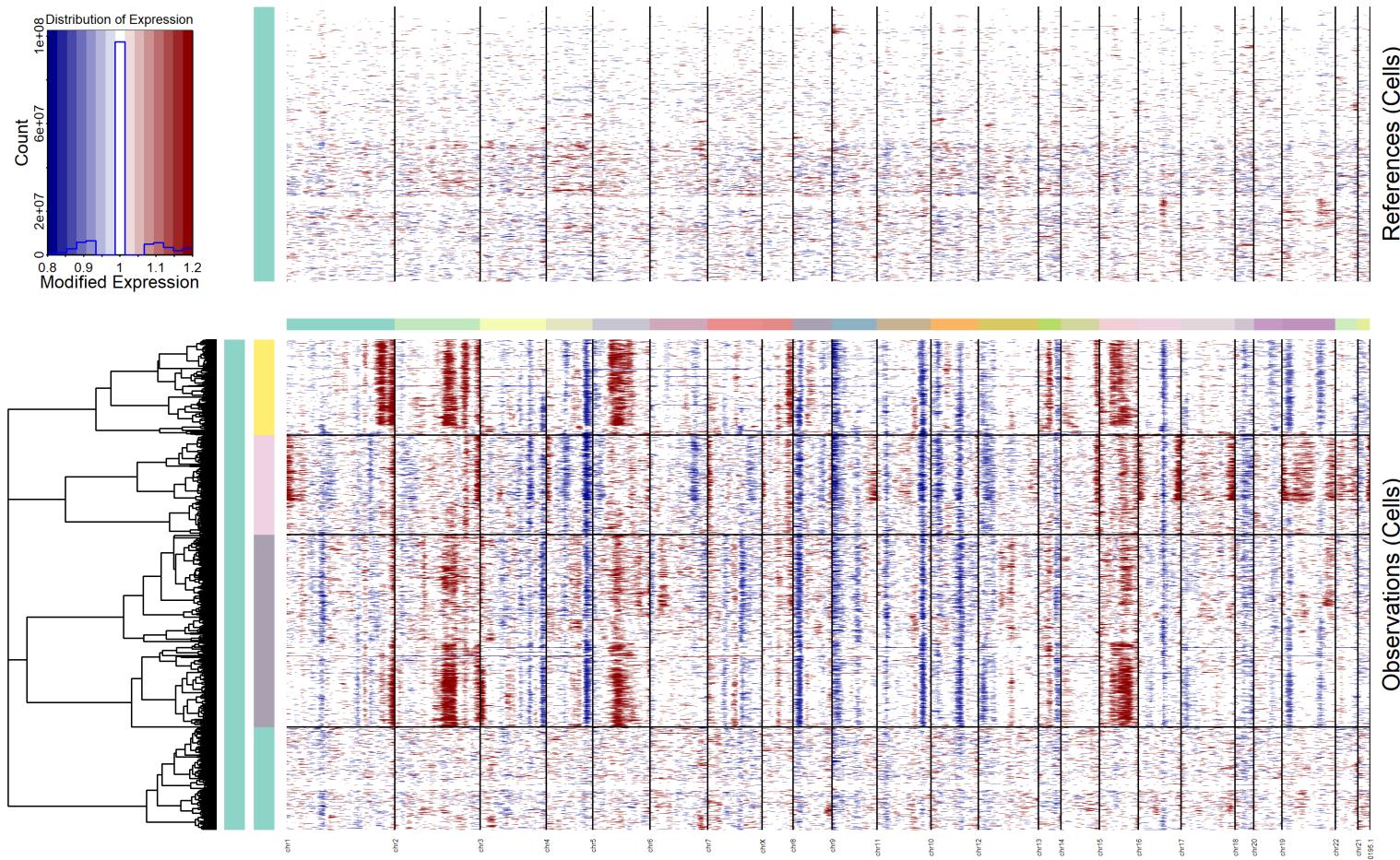


Lung metastasis #2



SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

inferCNV (default parameters for 10x genomics data)

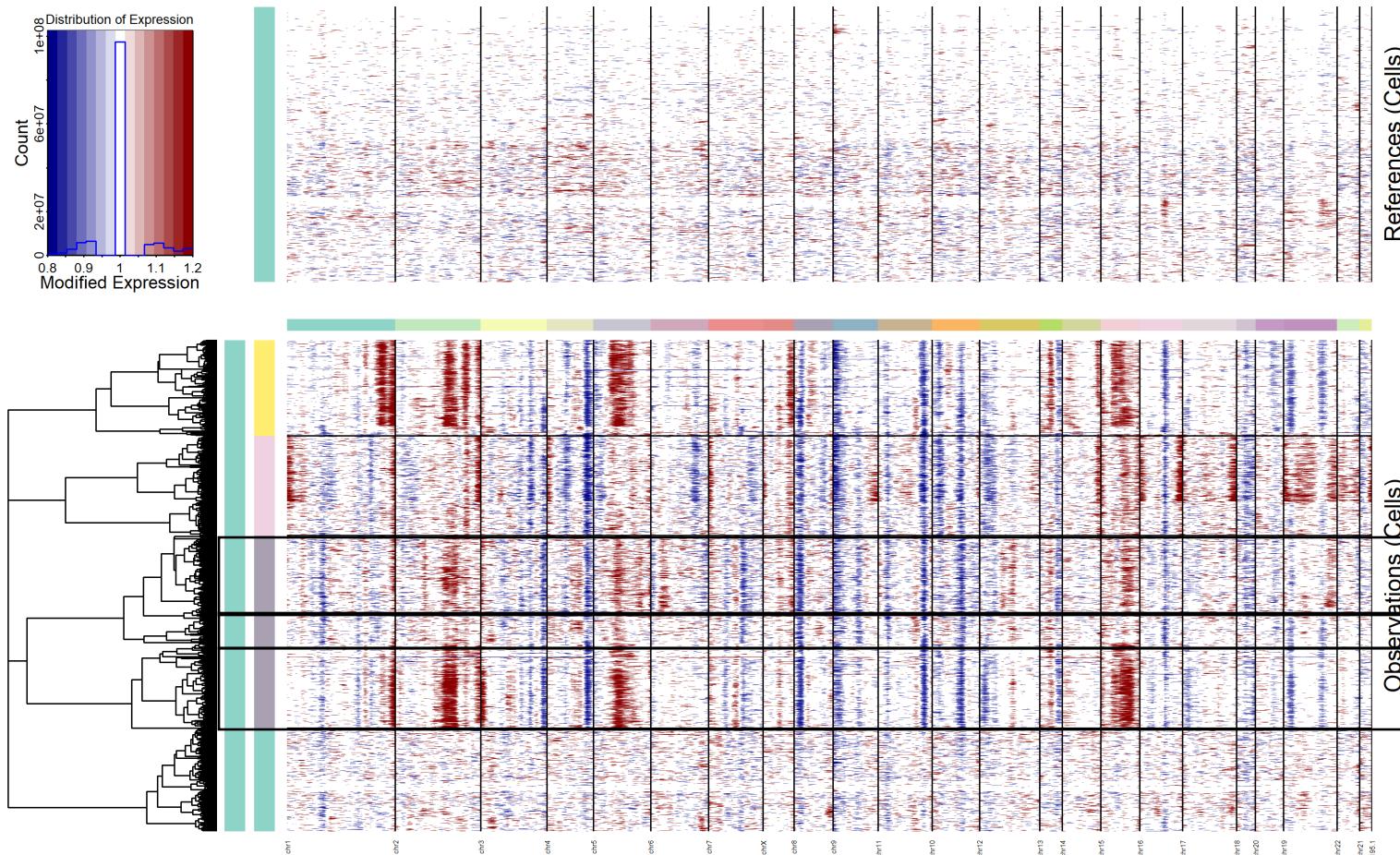


Non-tumor sample
Primary tumor
Lung metastasis #1
Lung metastasis #2

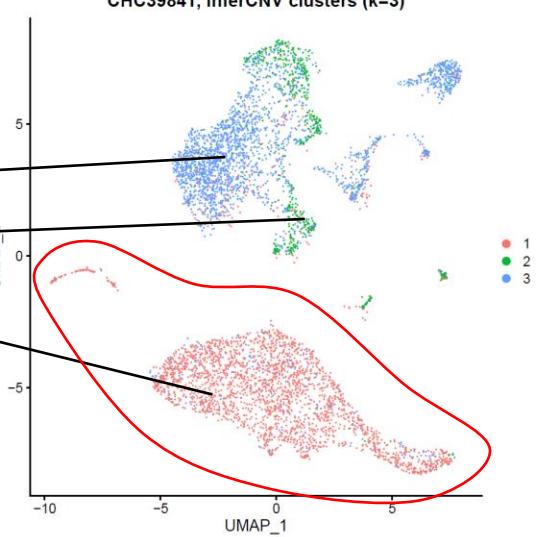
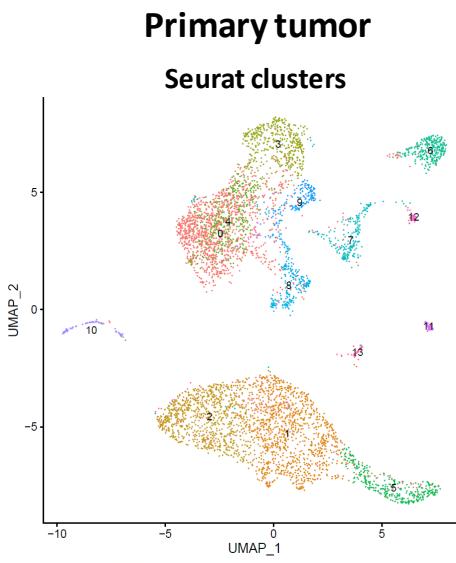
Observations (Cells) References (Cells)

SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

inferCNV (default parameters for 10x genomics data)



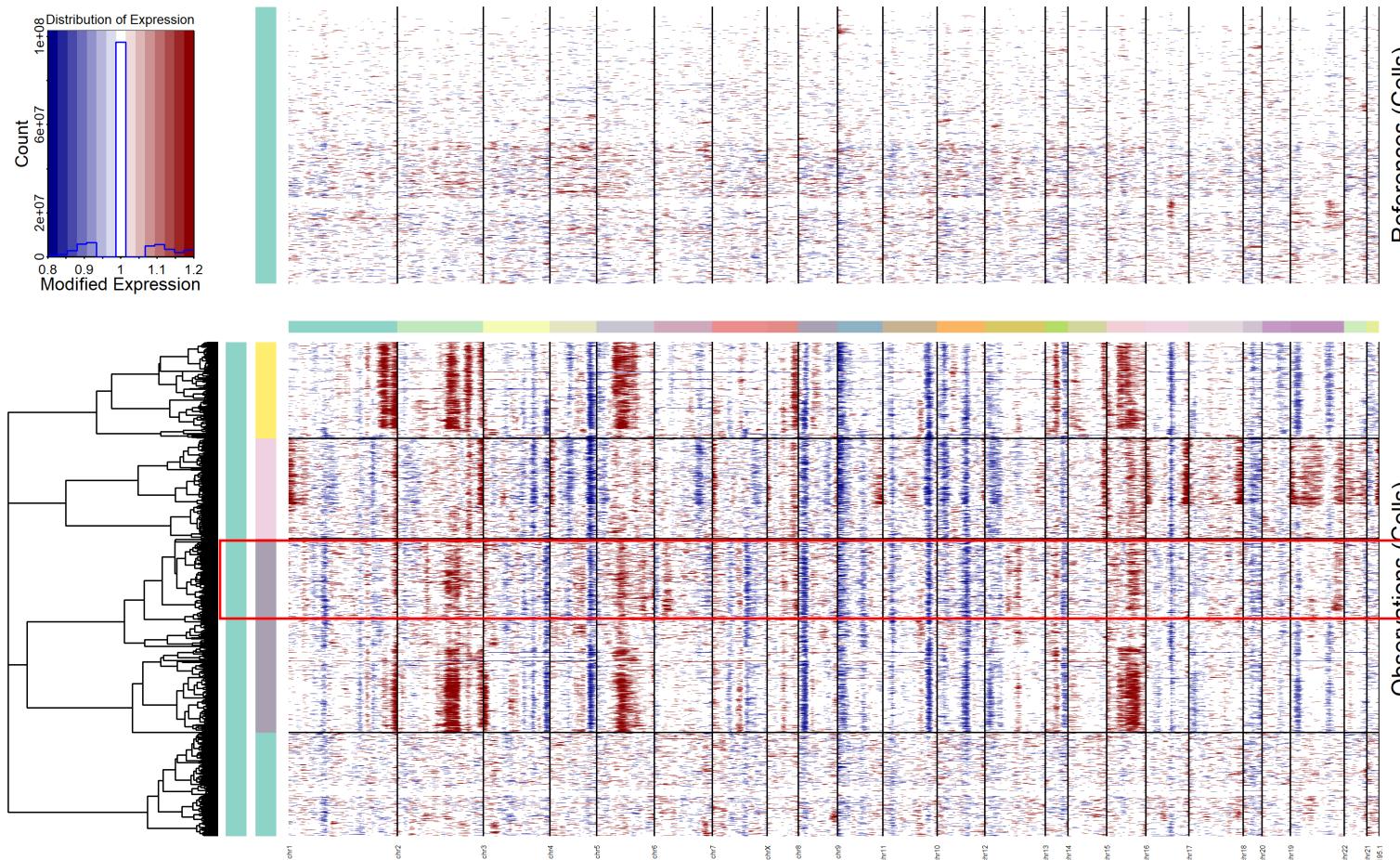
Non-tumor sample
Primary tumor
Lung metastasis #1
Lung metastasis #2



Tumor cells according to inferCNV

SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

inferCNV (default parameters for 10x genomics data)

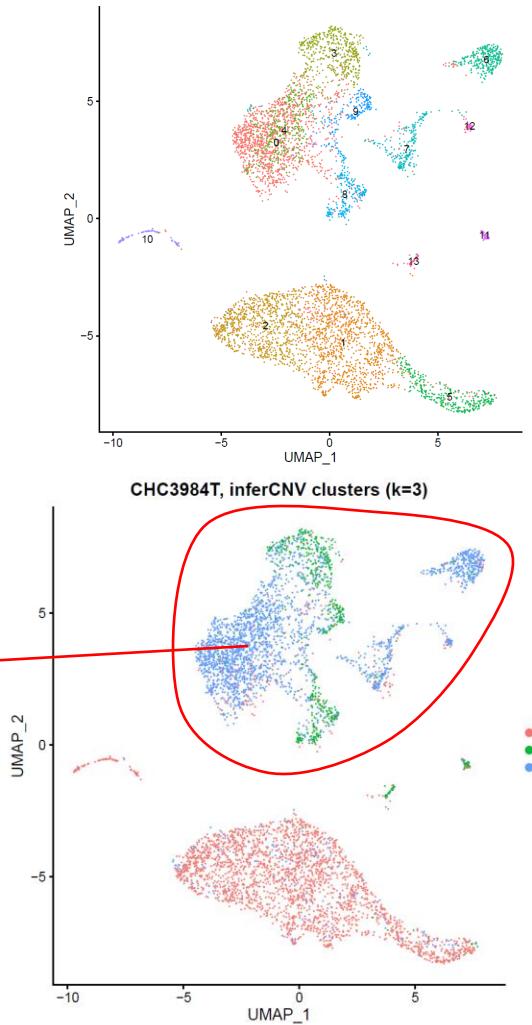


Non-tumor sample
Primary tumor
Lung metastasis #1
Lung metastasis #2

Question: Normal cells that are contaminated by tumor RNA?

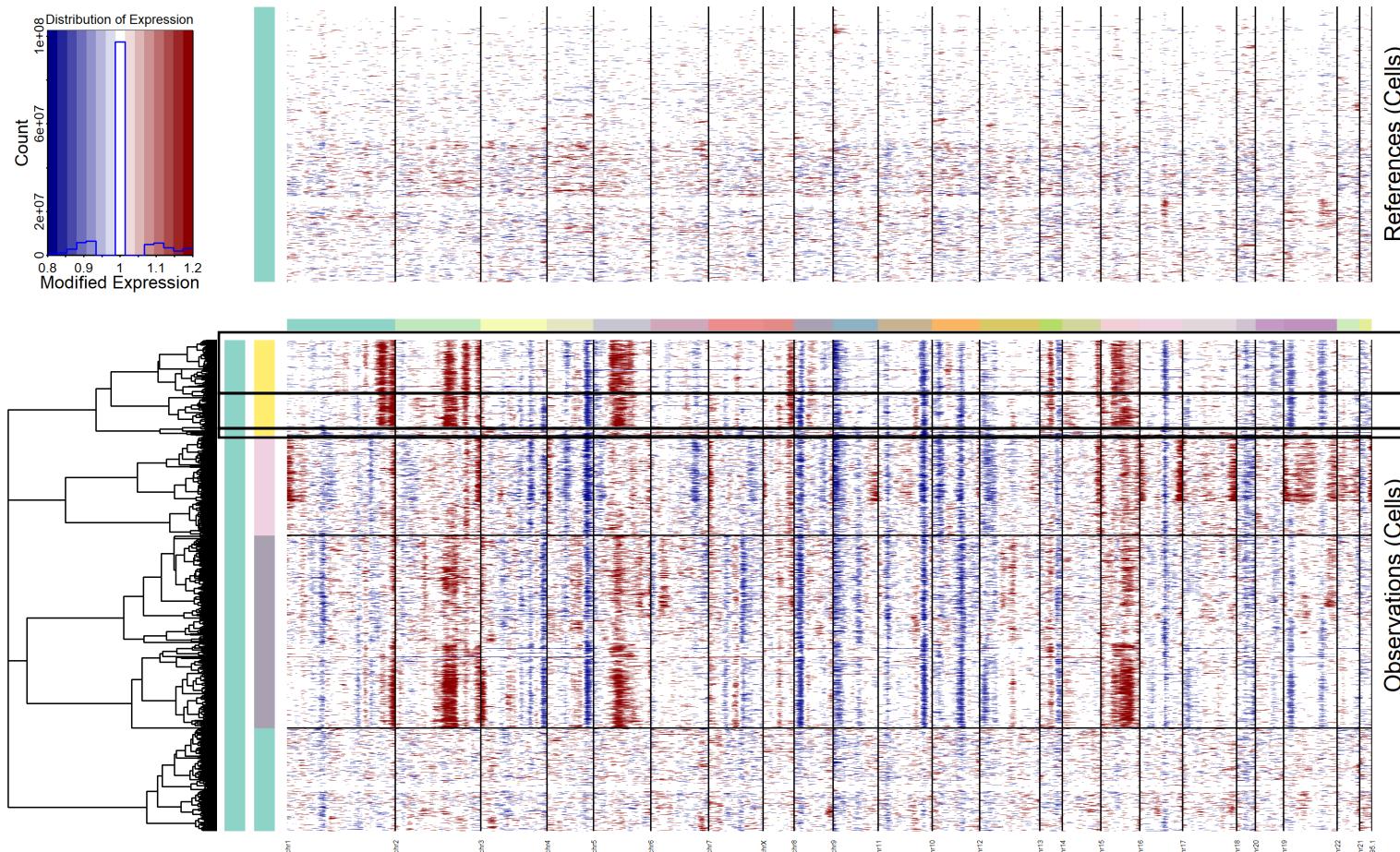
Observations (Cells) → References (Cells)

Primary tumor
Seurat clusters



SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

inferCNV (default parameters for 10x genomics data)

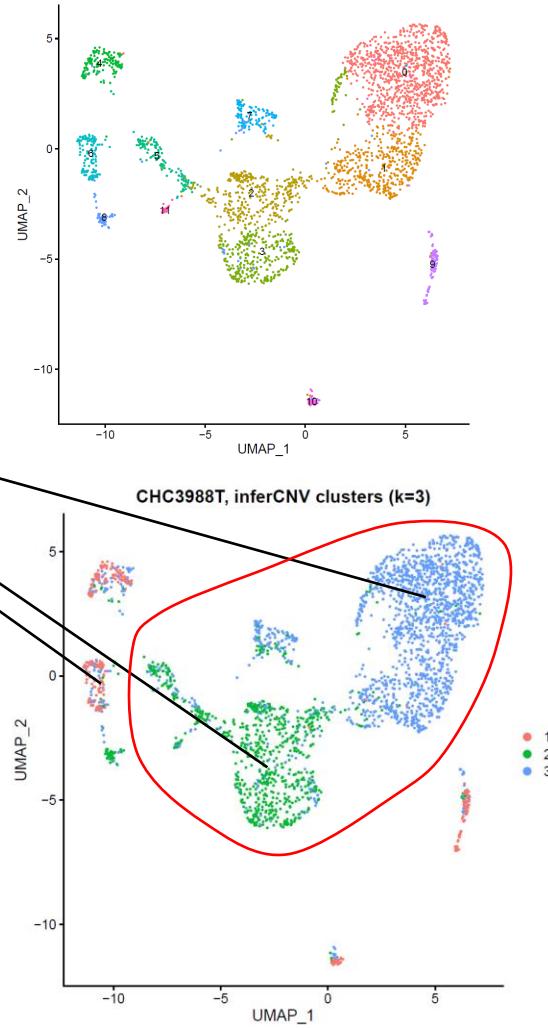


Non-tumor sample
Primary tumor
Lung metastasis #1
Lung metastasis #2

Observations (Cells)

References (Cells)

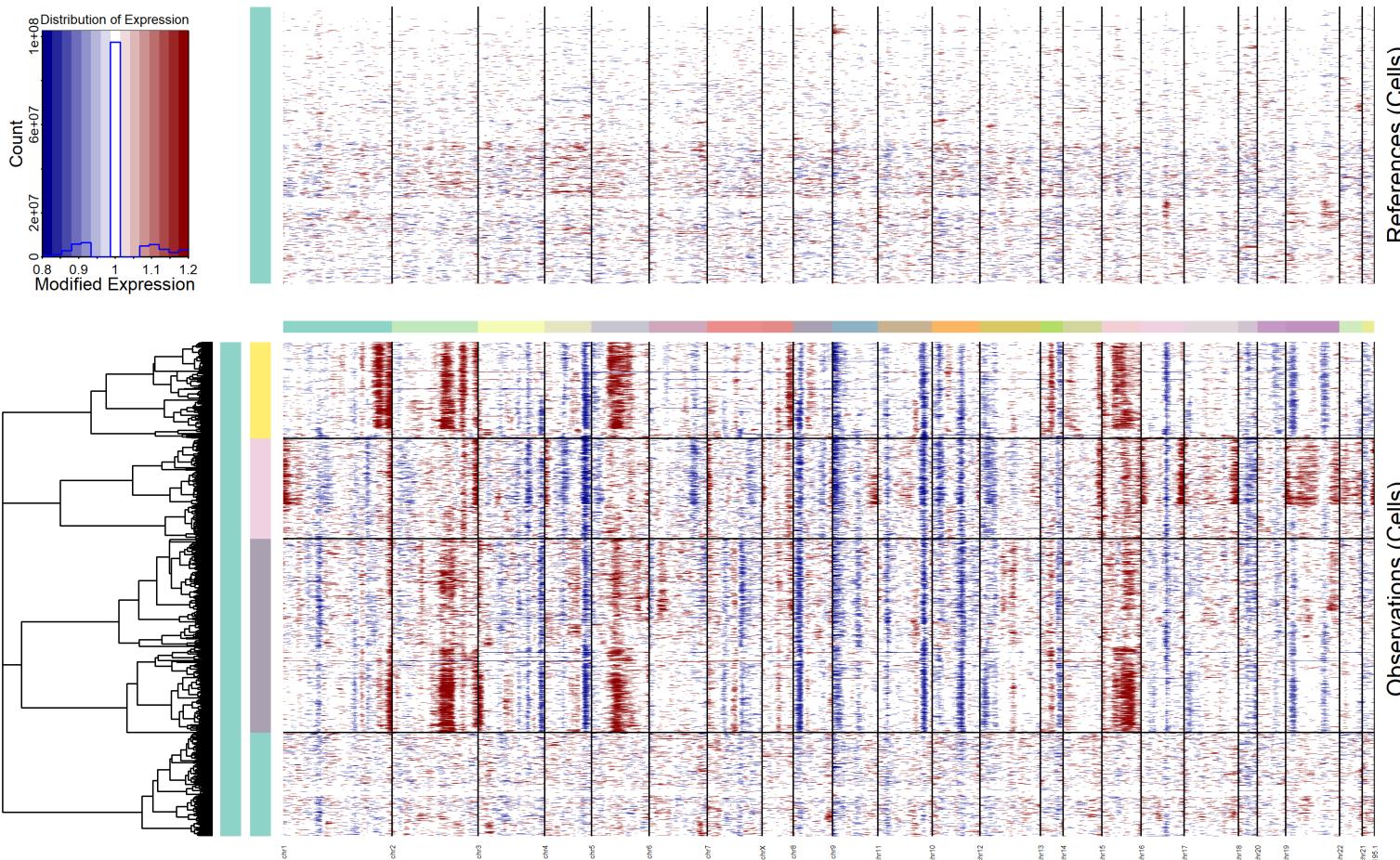
Metastasis #2
Seurat clusters



Tumor cells according to inferCNV

SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

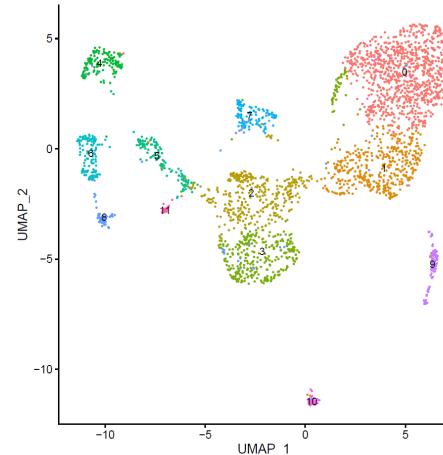
inferCNV (default parameters for 10x genomics data)



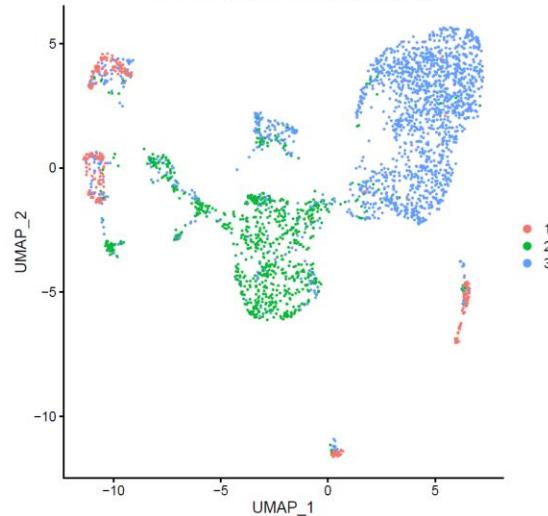
Non-tumor sample
Primary tumor
Lung metastasis #1
Lung metastasis #2

Question: Unexpected deletions in the single-nucleus profiles compared to the bulk expected profiles?

Metastasis #2
Seurat clusters



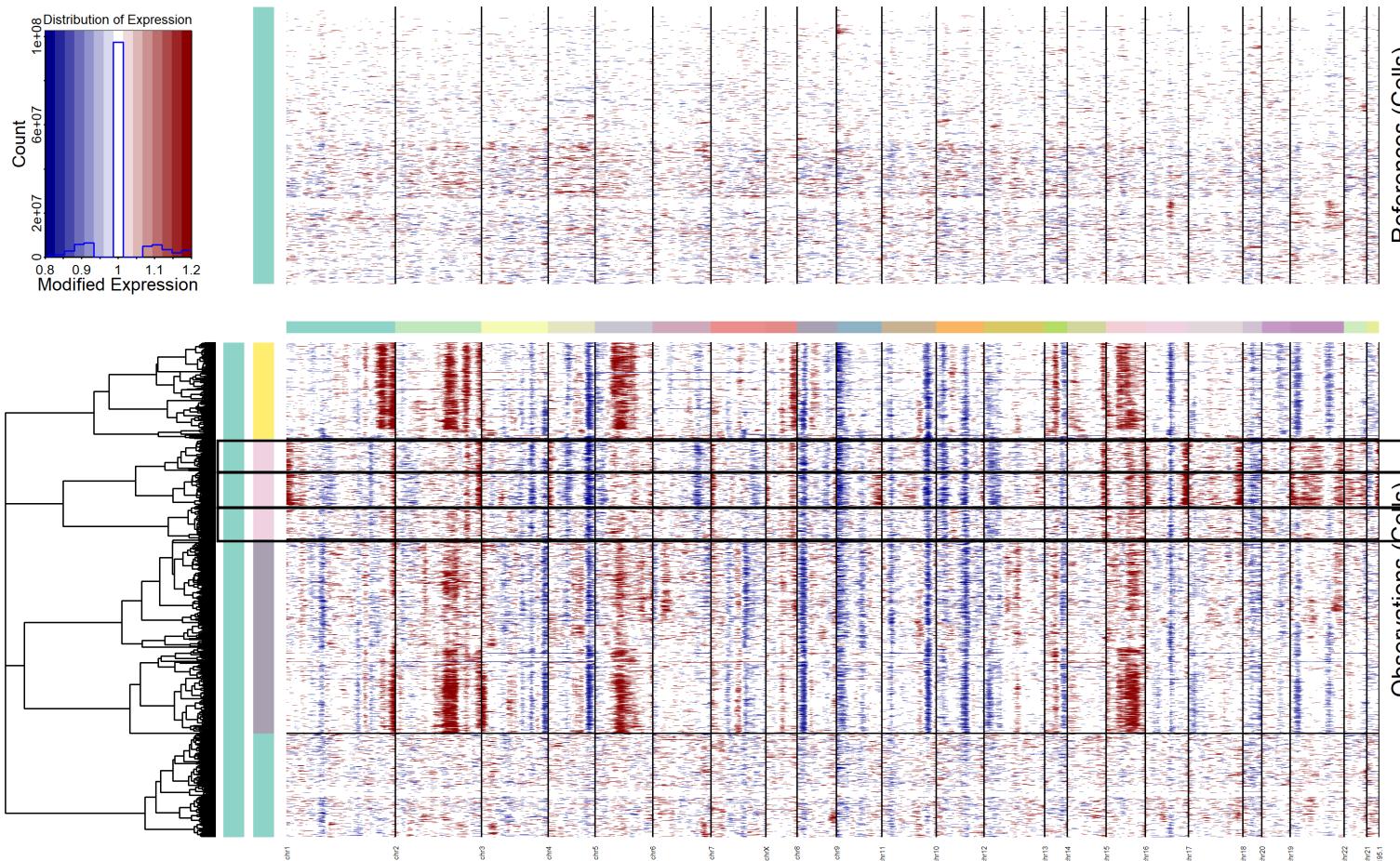
CHC3988T, inferCNV clusters (k=3)



Tumor cells according to inferCNV

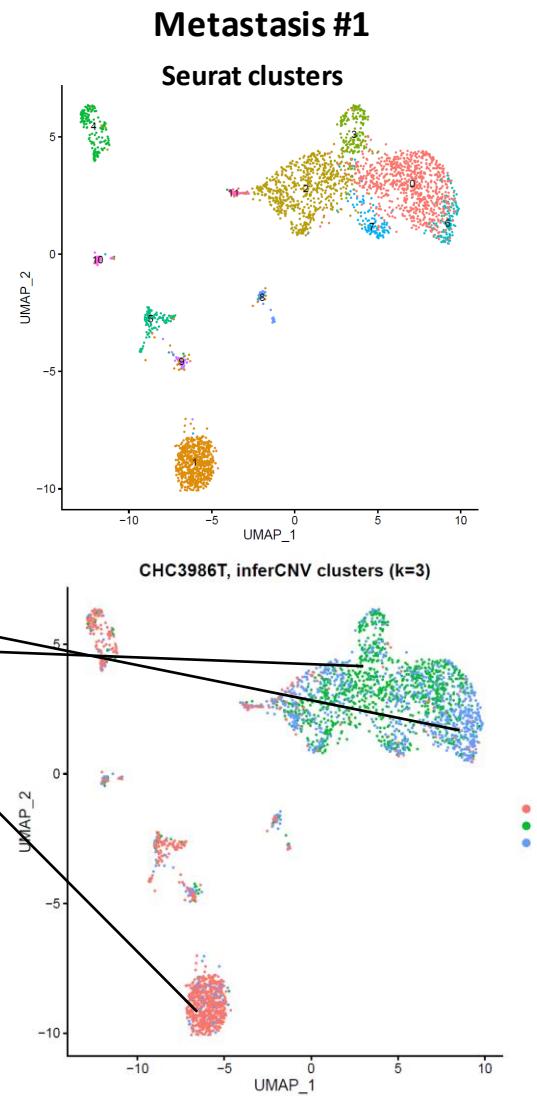
SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

inferCNV (default parameters for 10x genomics data)



Non-tumor sample
Primary tumor
Lung metastasis #1
Lung metastasis #2

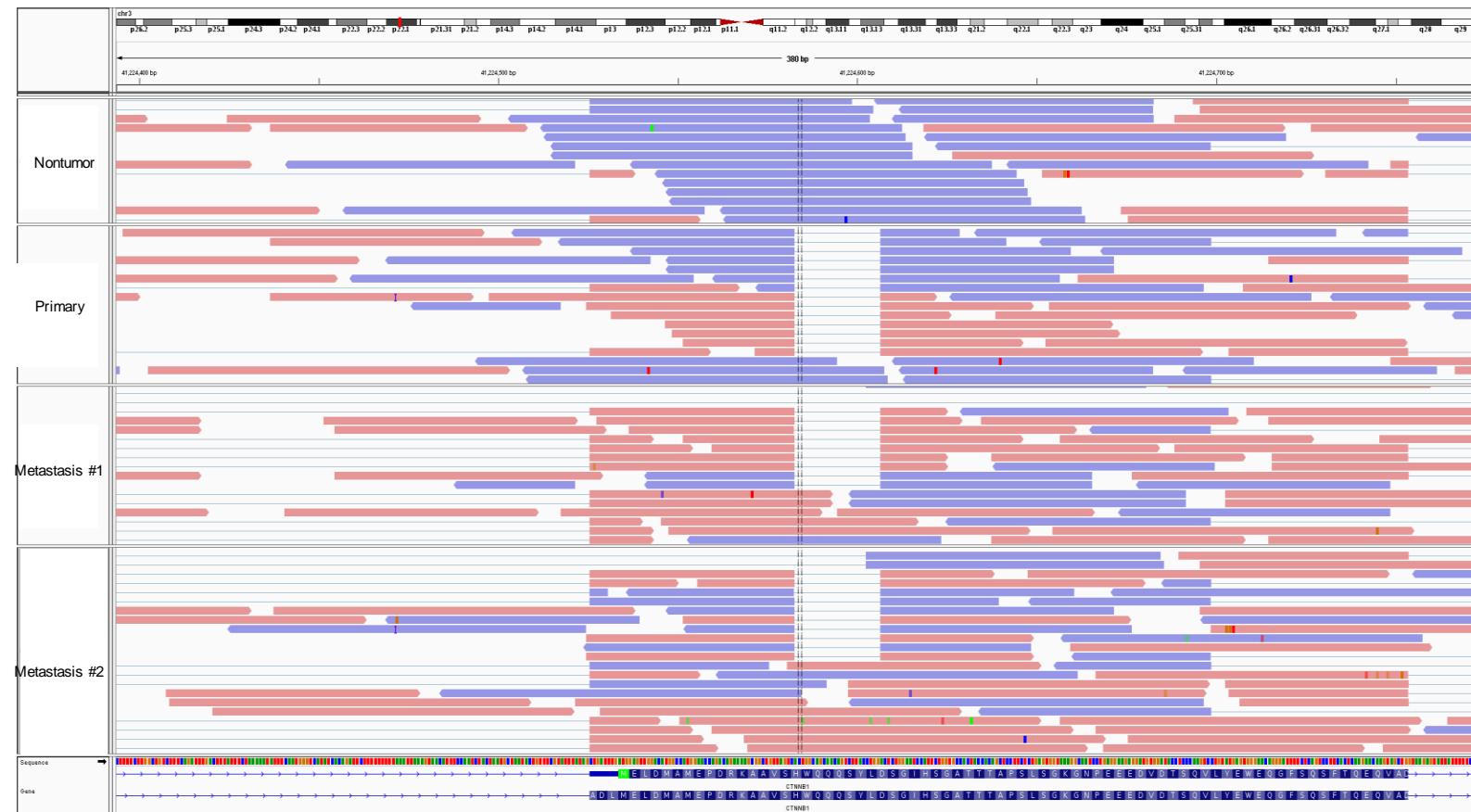
Question: Why this unexpected copy-number profile? Is it because of the additional PCR cycles?



SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

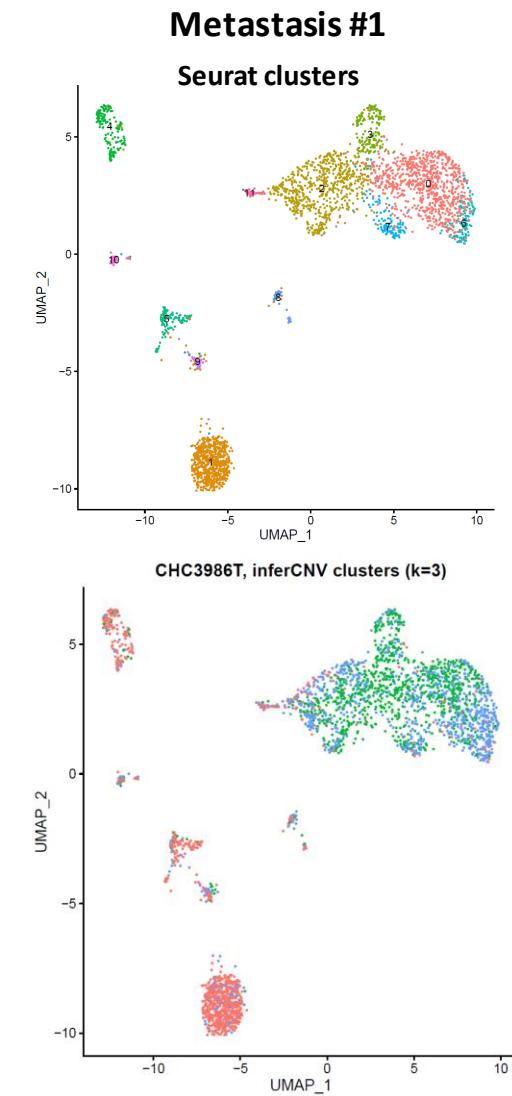
inferCNV (default parameters for 10x genomics data)

IGV screenshot of the BAM files



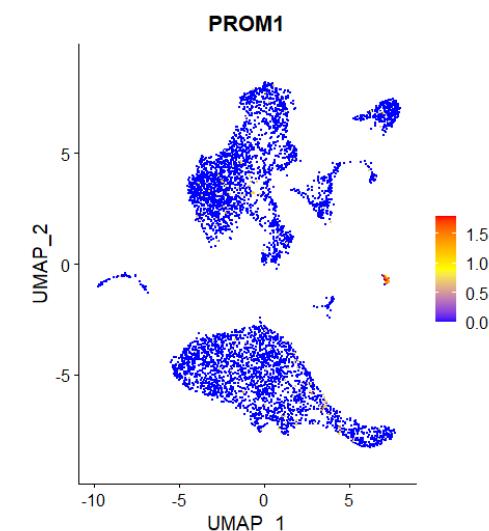
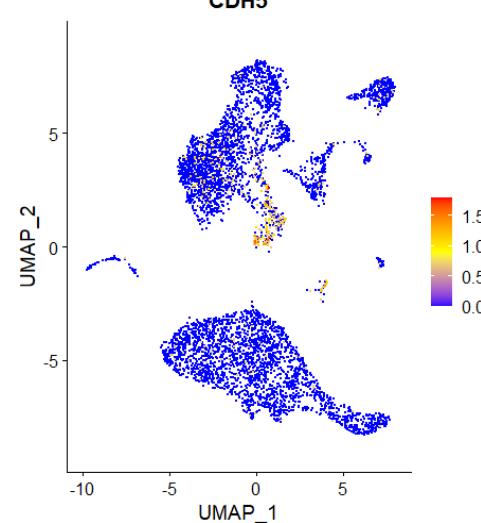
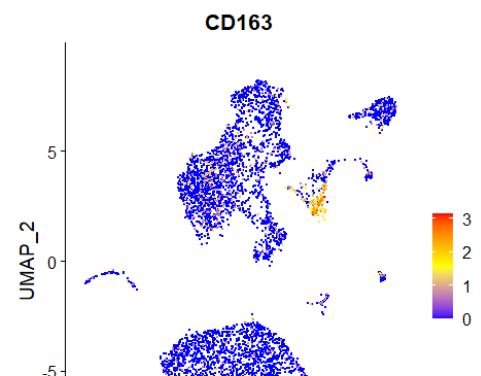
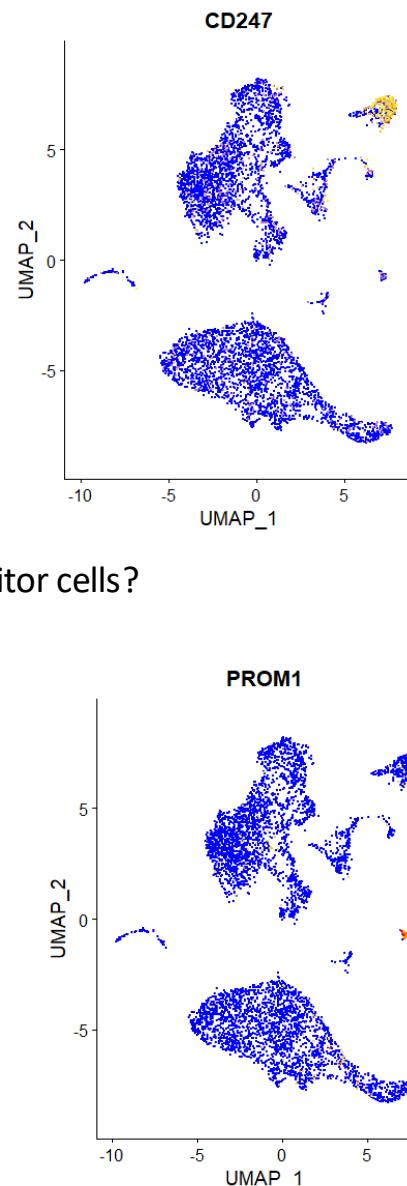
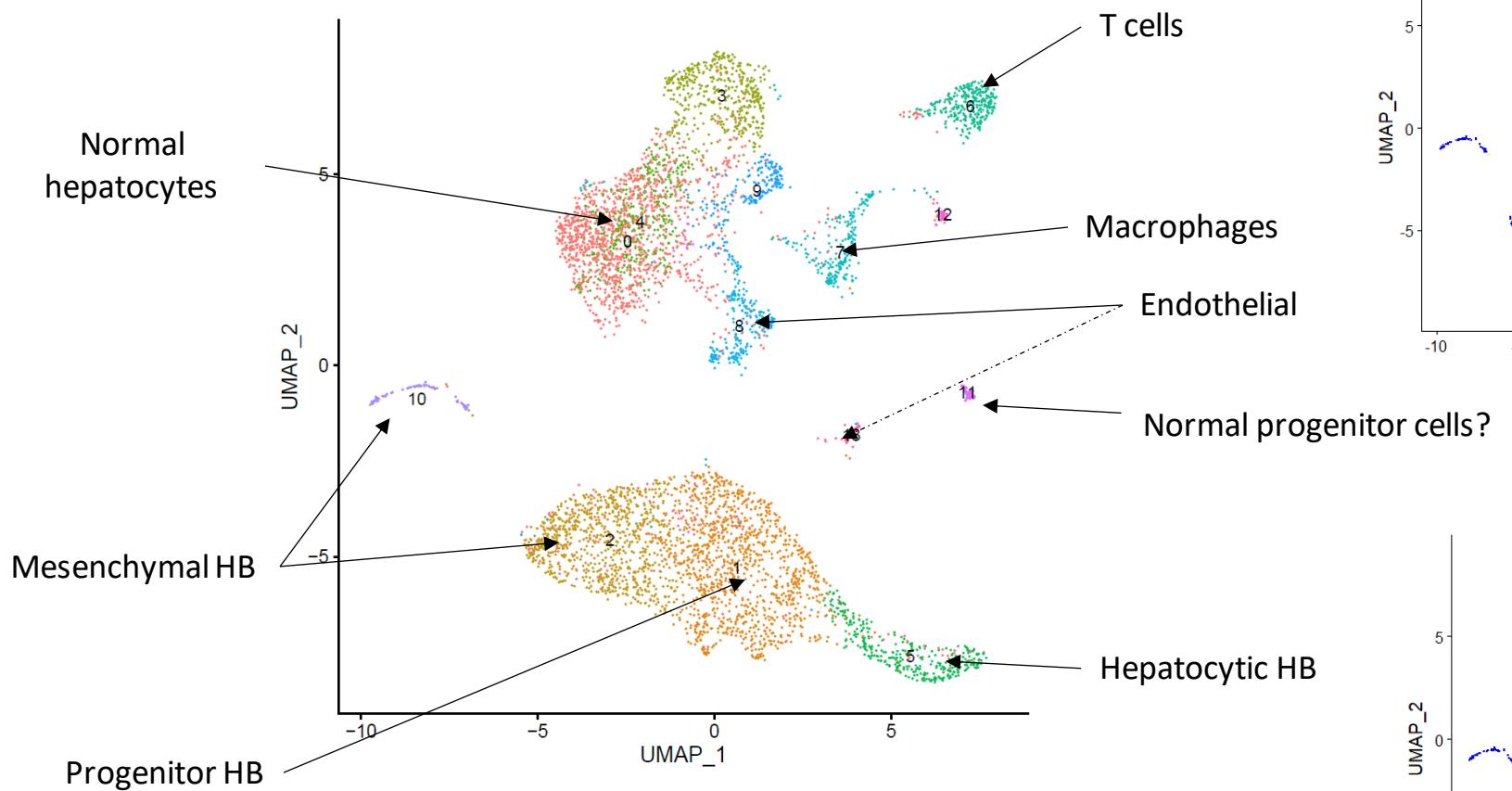
Non-tumor sample
Primary tumor
Lung metastasis #1
Lung metastasis #2

Question: Why this unexpected copy-number profile? Is it because of the additional PCR cycles?



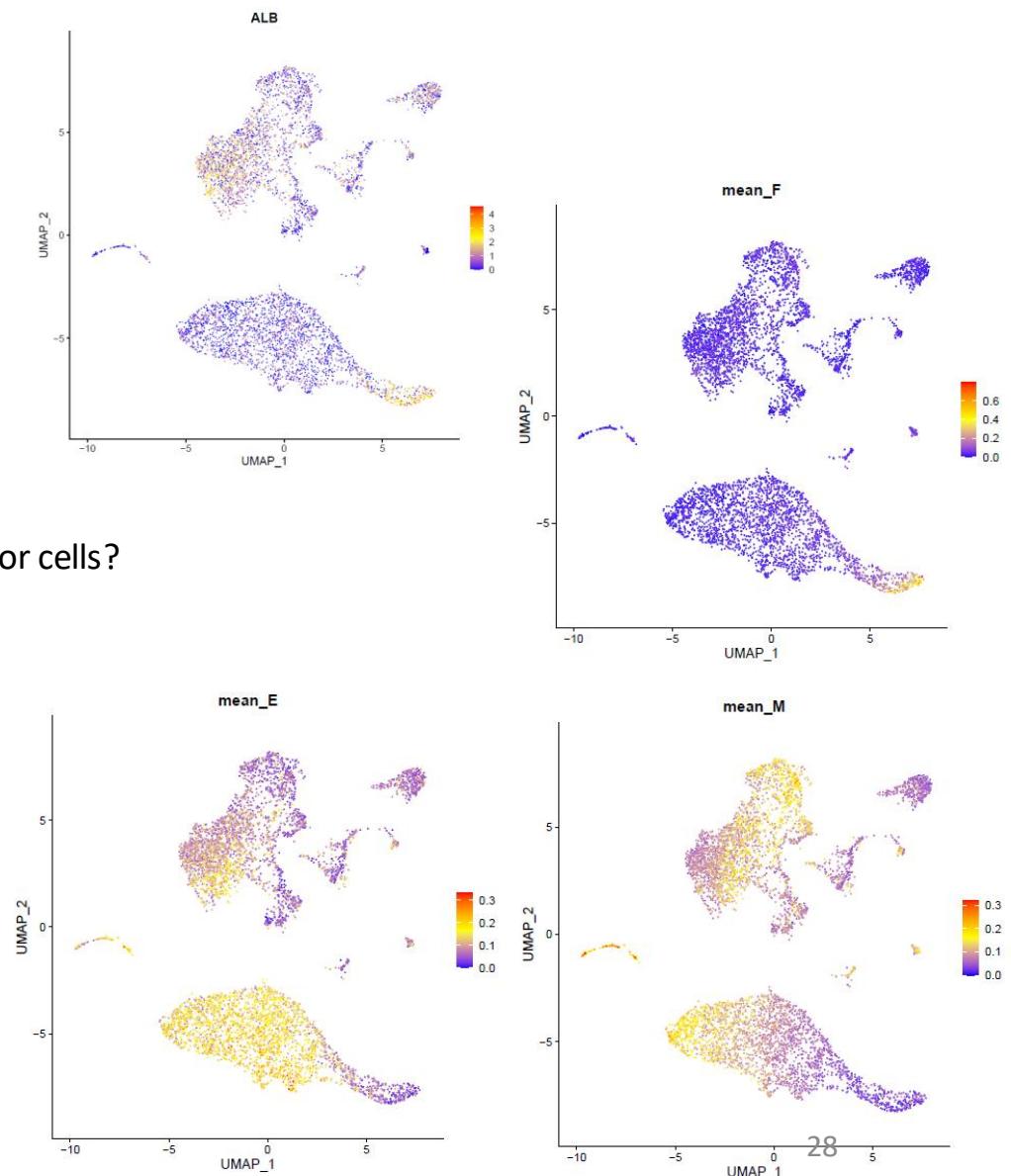
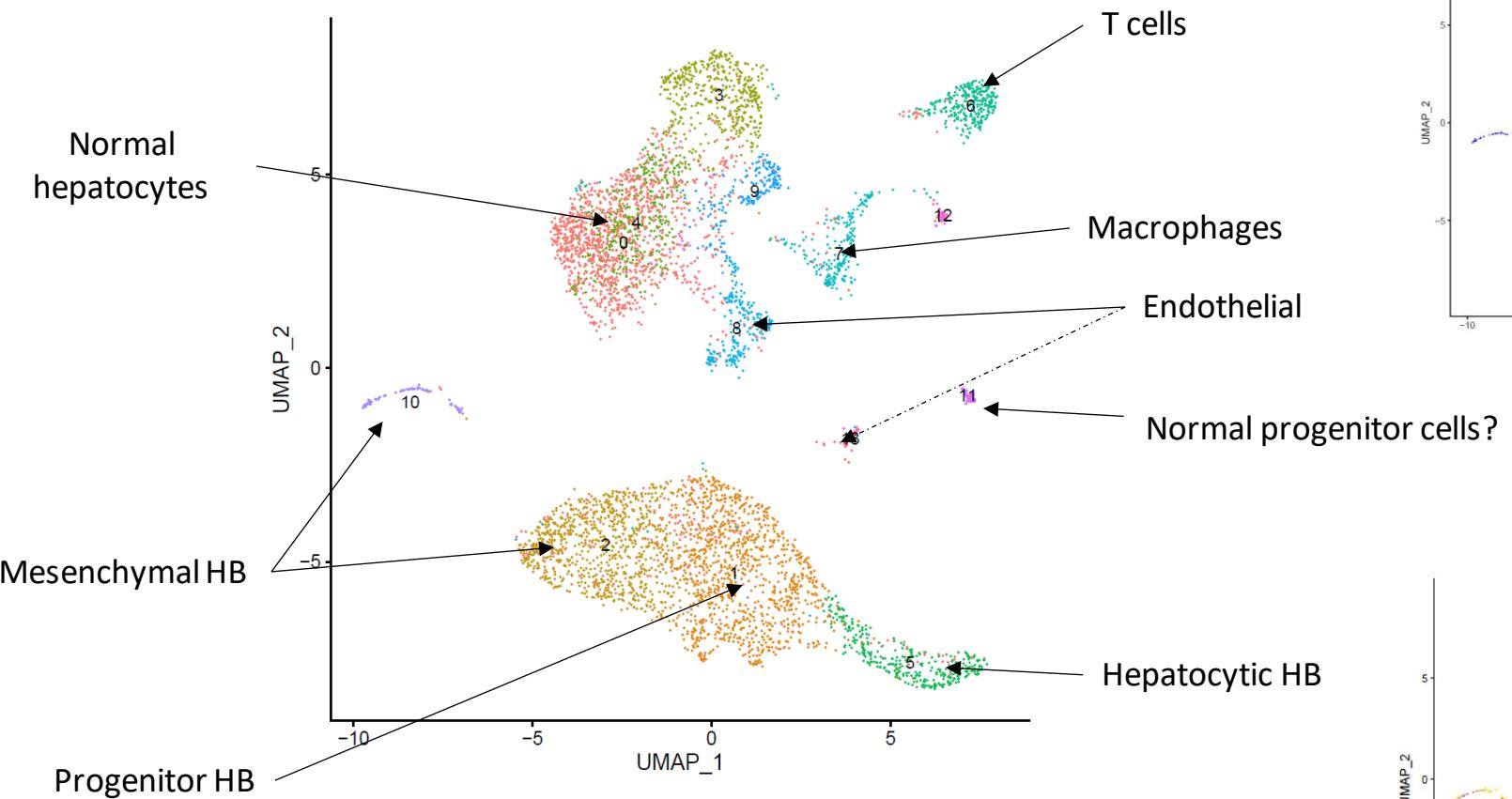
SN-RNASEQ OF FROZEN HB: PRELIMINARY ANNOTATIONS

Primary tumor



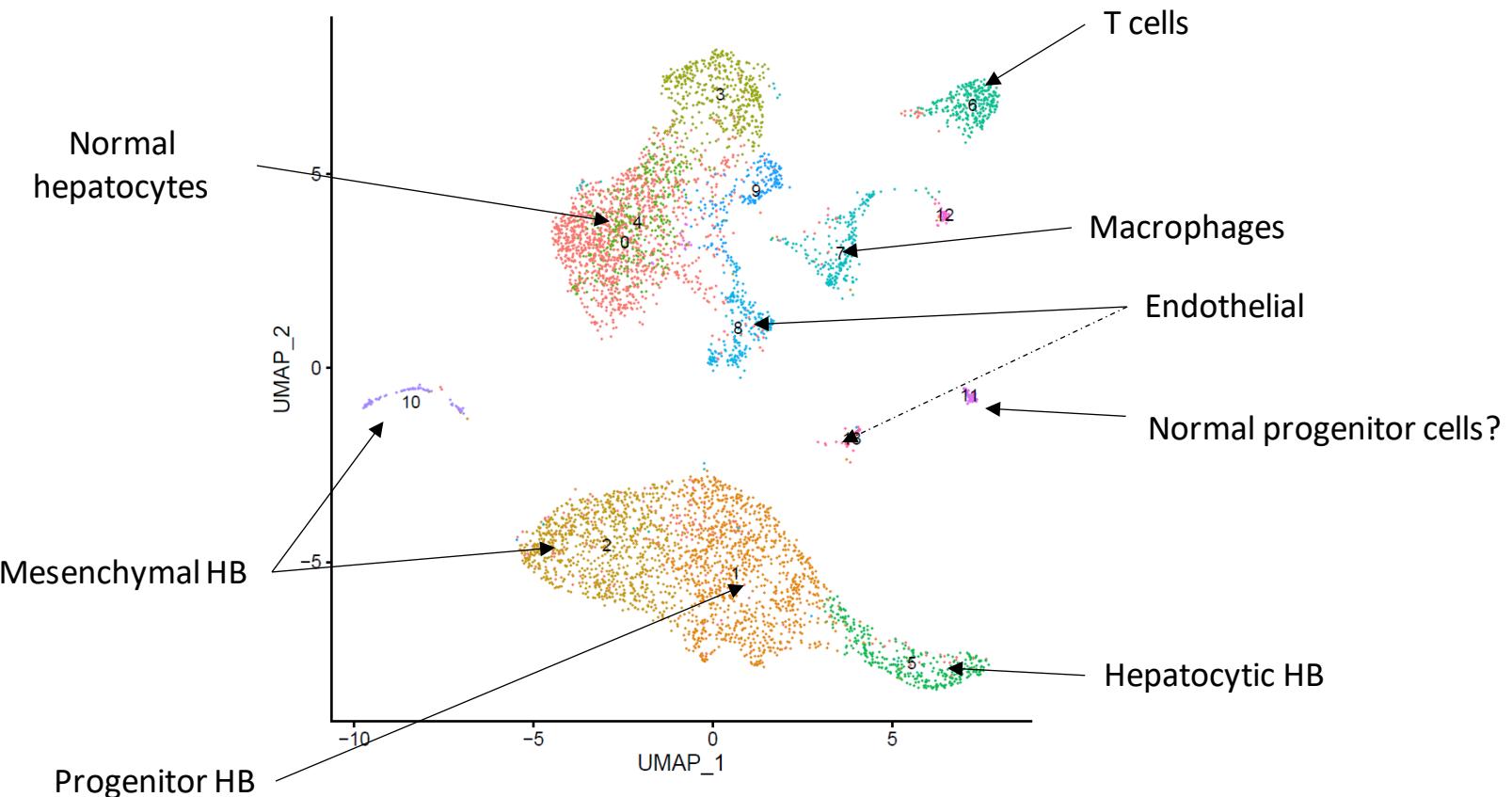
SN-RNASEQ OF FROZEN HB: PRELIMINARY ANNOTATIONS

Primary tumor

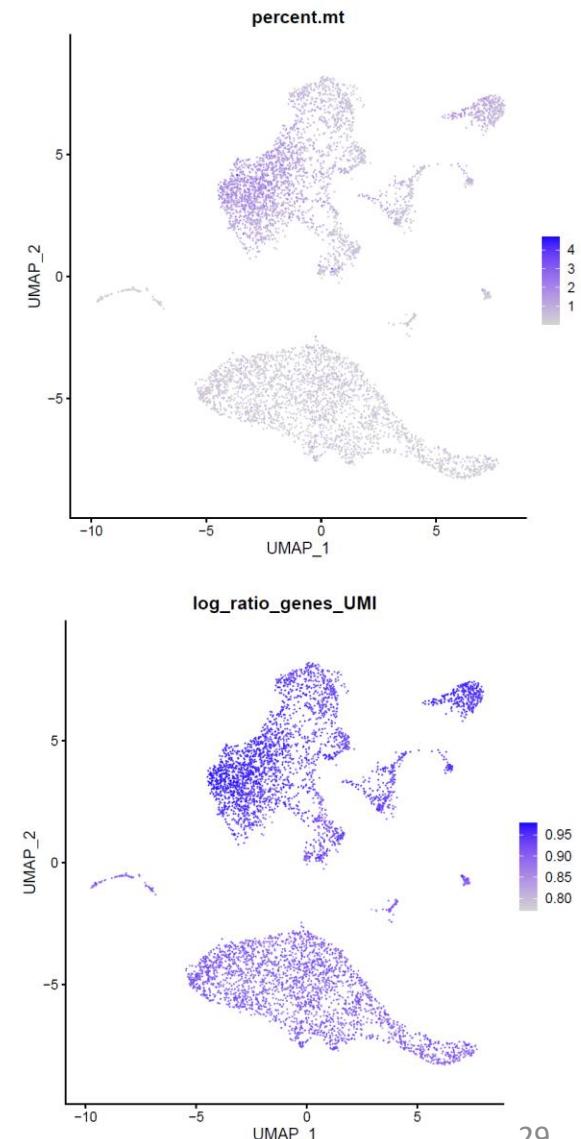


SN-RNASEQ OF FROZEN HB: PRELIMINARY ANNOTATIONS

Primary tumor

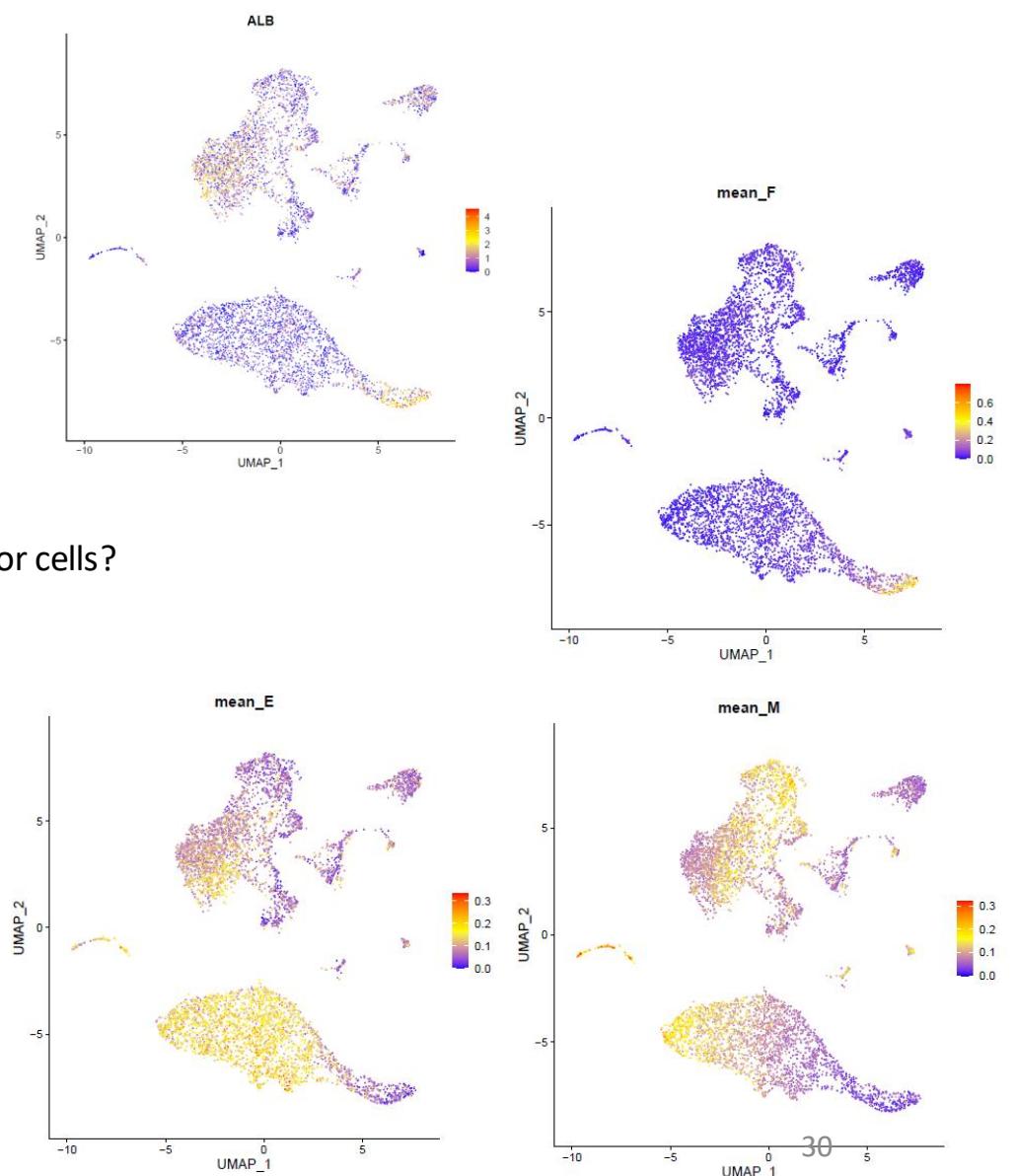
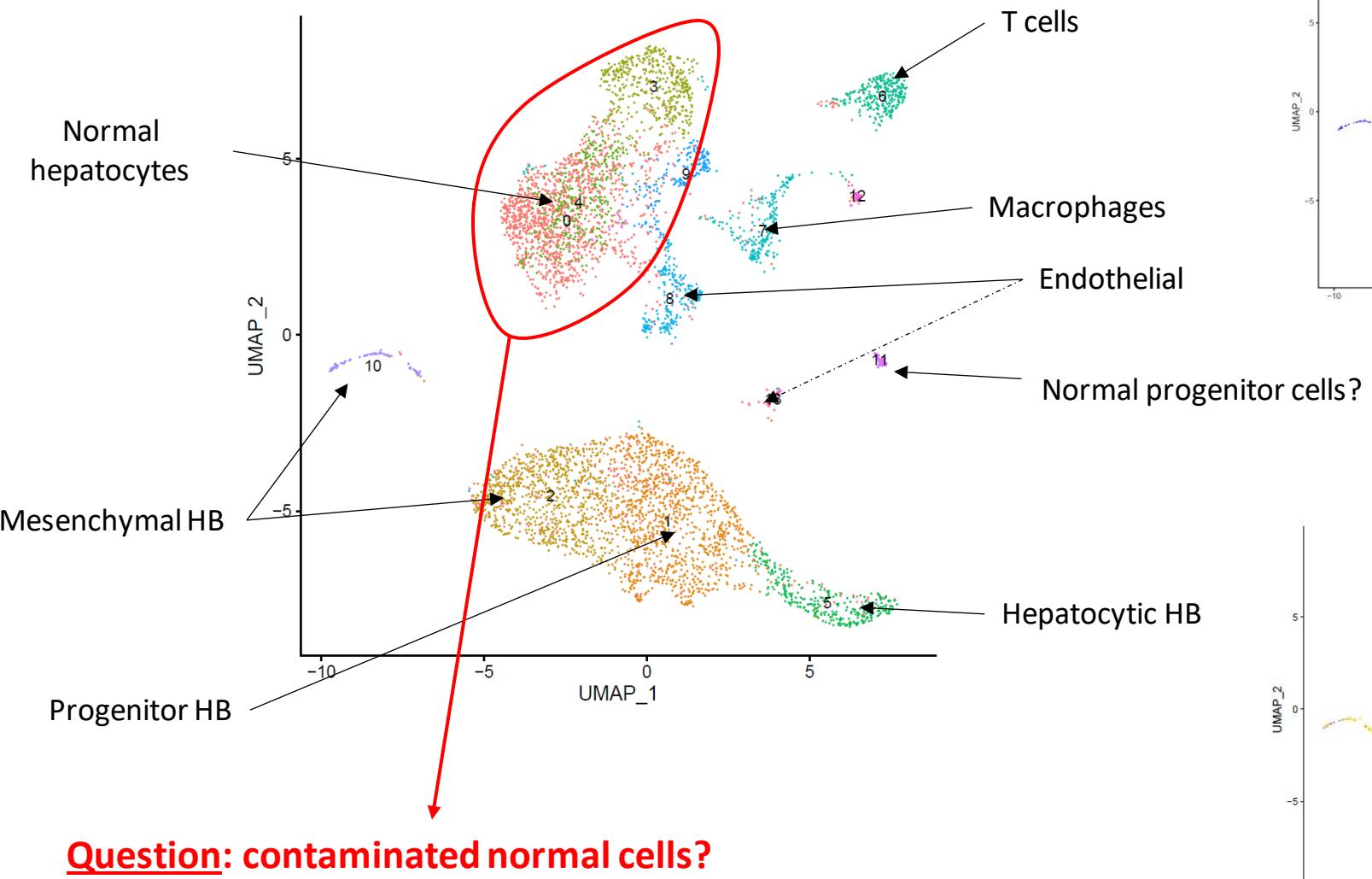


Link with mitochondrial %?



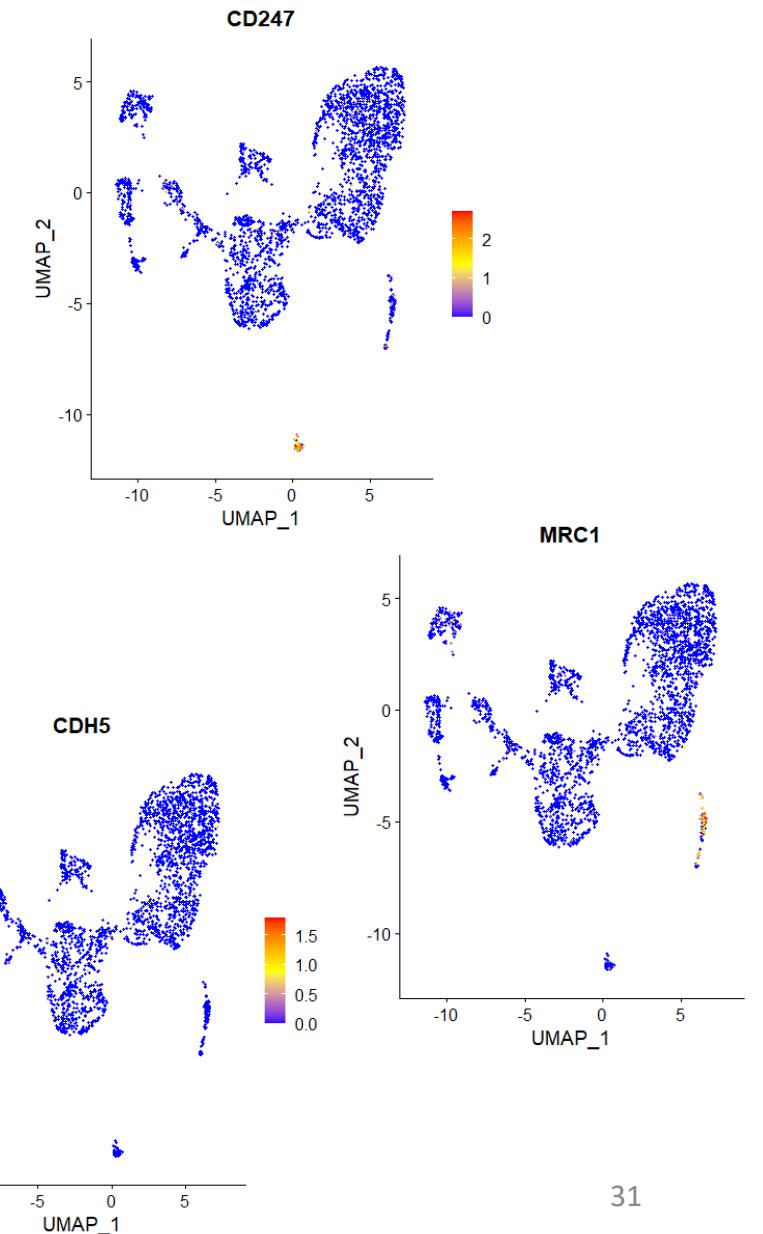
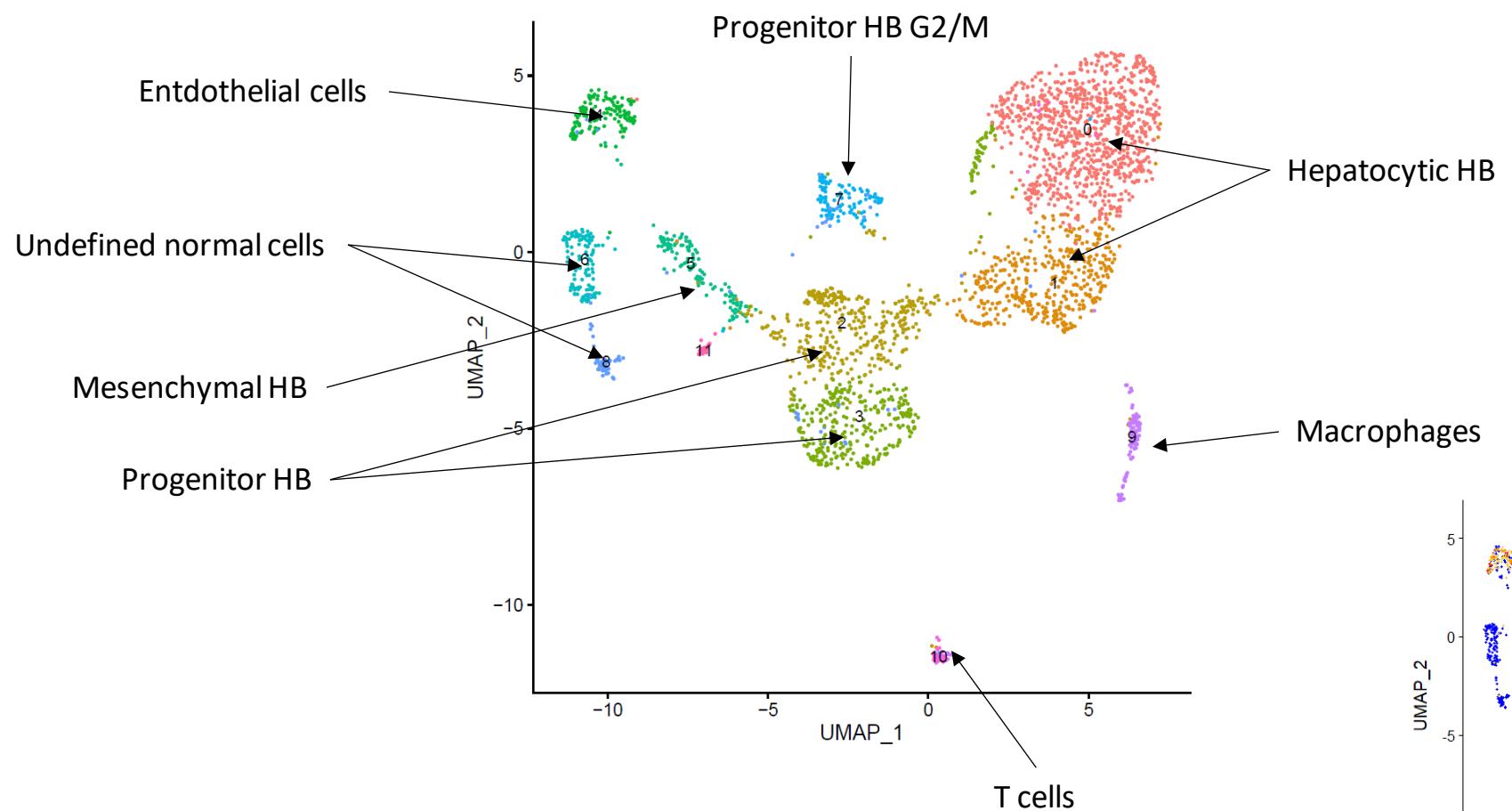
SN-RNASEQ OF FROZEN HB: PRELIMINARY ANNOTATIONS

Primary tumor



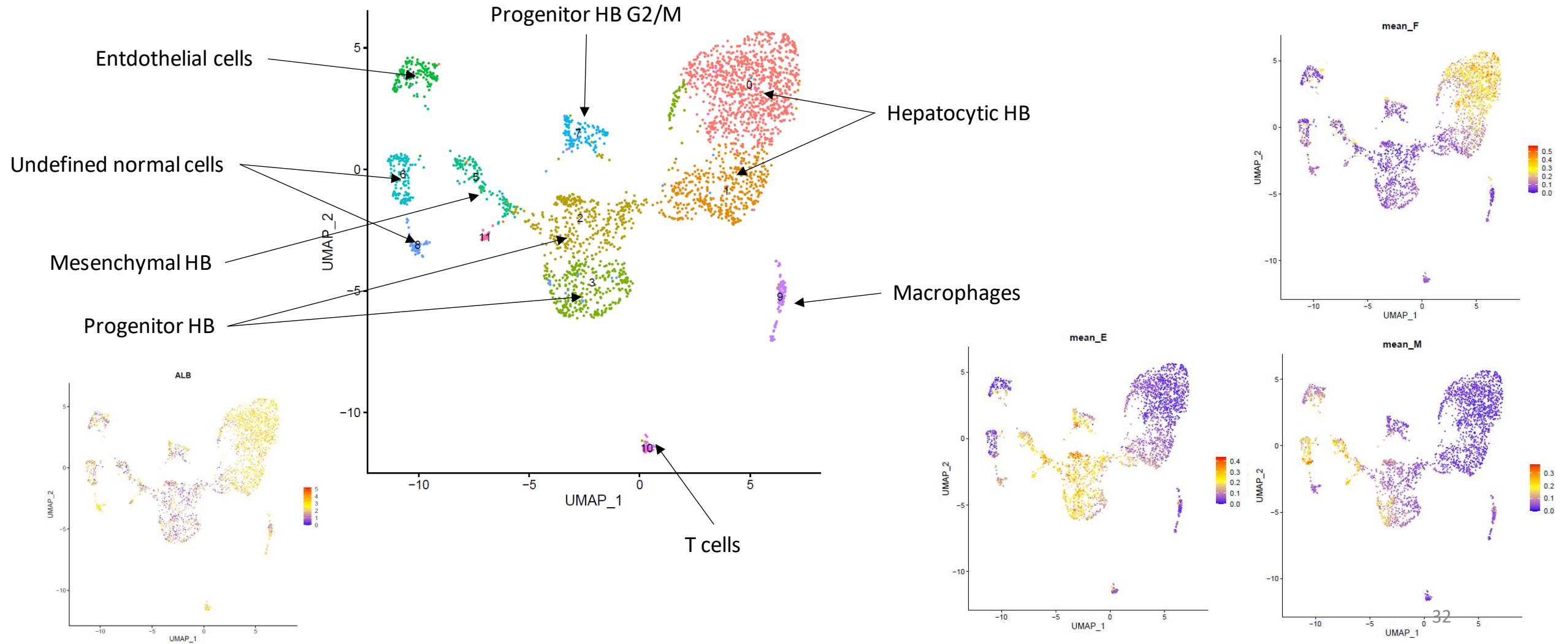
SN-RNASEQ OF FROZEN HB: PRELIMINARY ANNOTATIONS

Lung metastasis #2



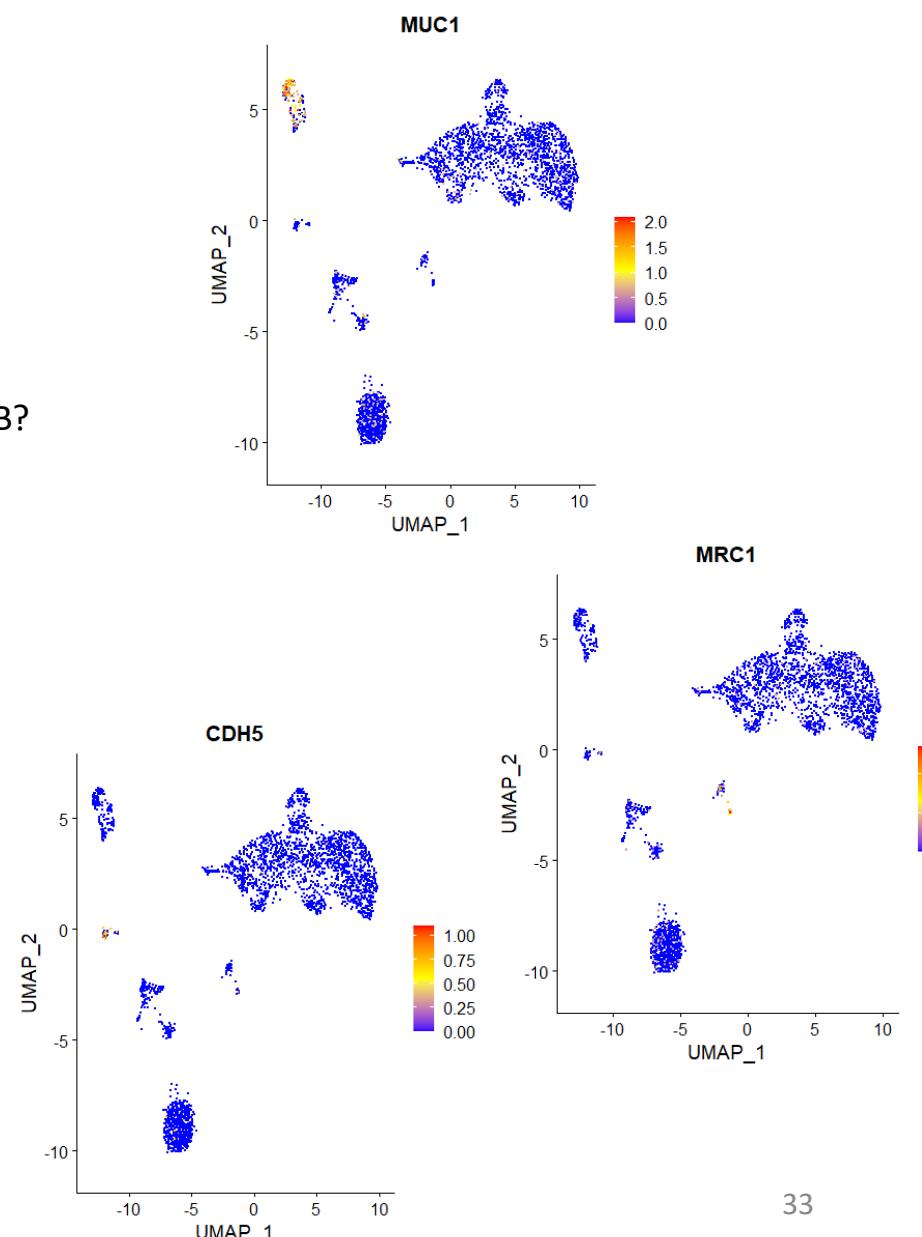
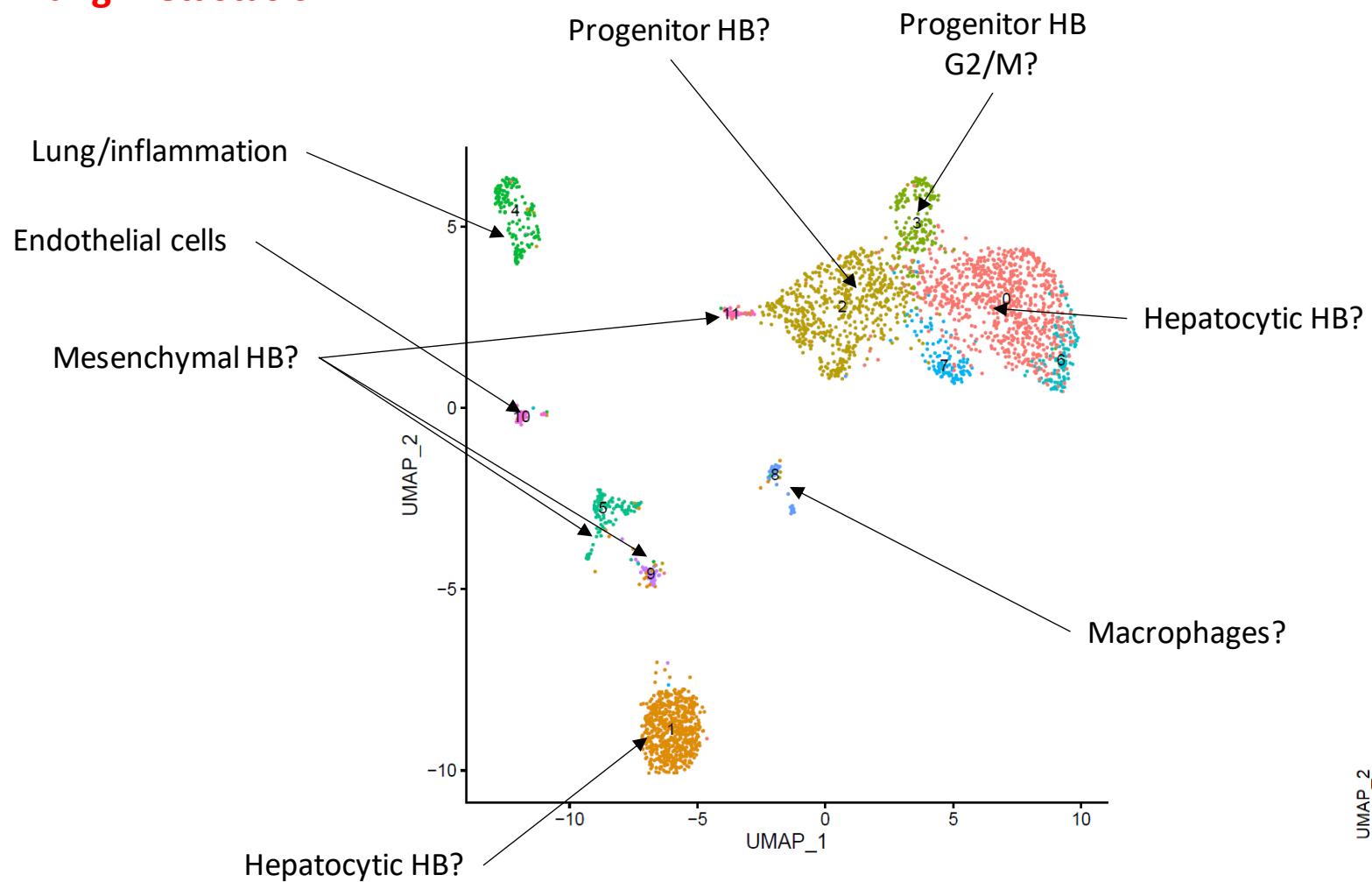
SN-RNASEQ OF FROZEN HB: PRELIMINARY ANNOTATIONS

Lung metastasis #2



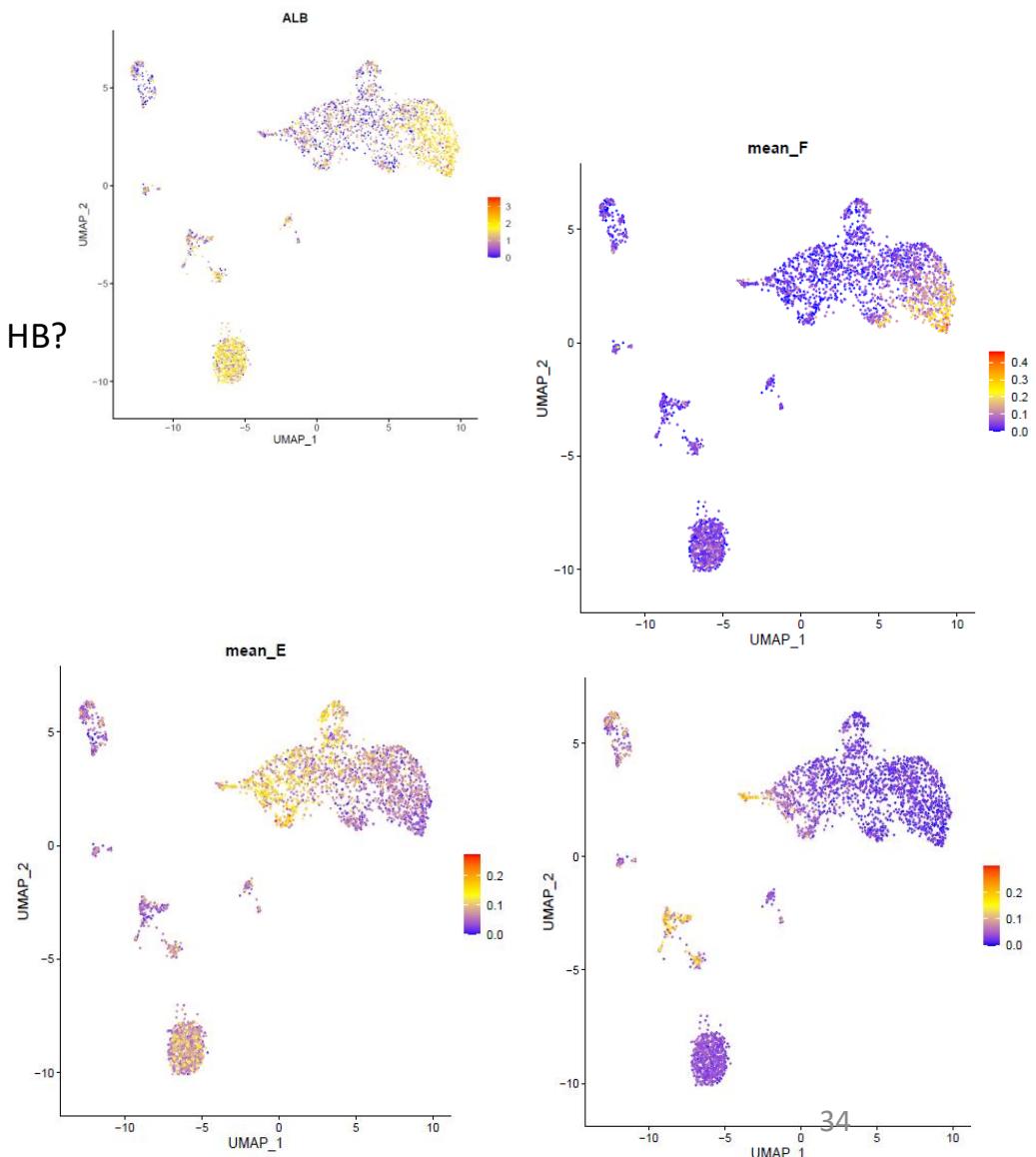
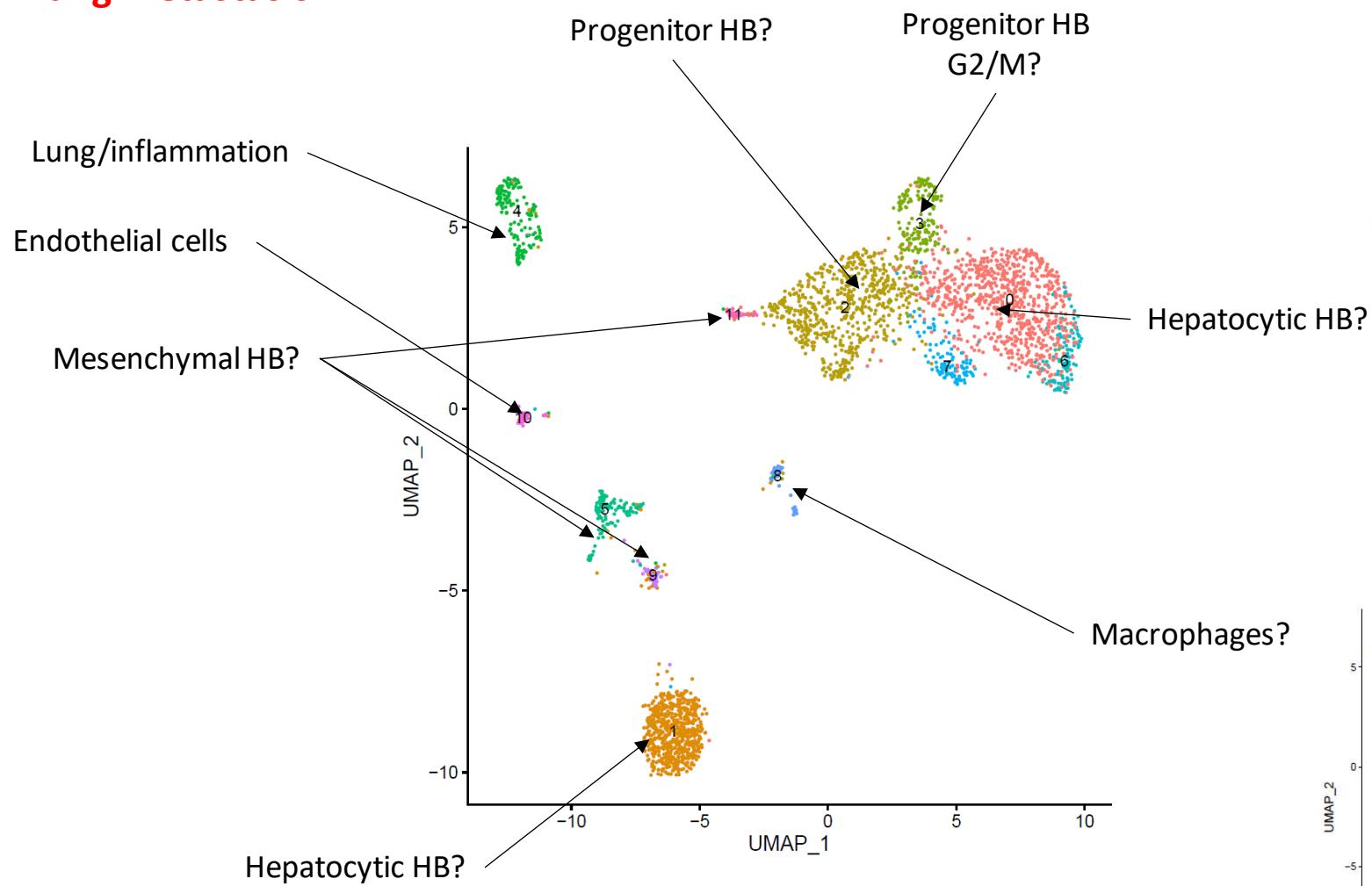
SN-RNASEQ OF FROZEN HB: PRELIMINARY ANNOTATIONS

Lung metastasis #1



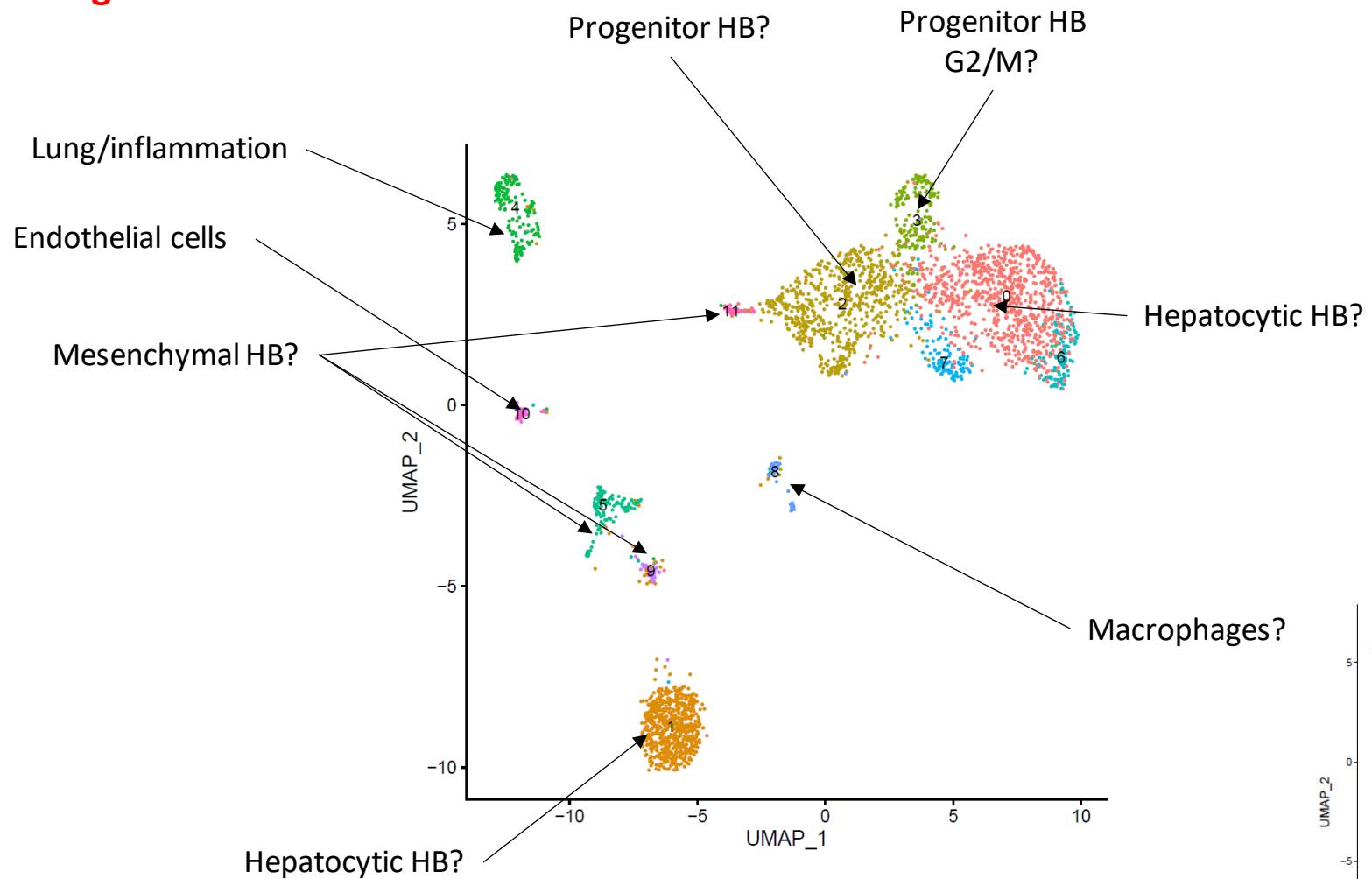
SN-RNASEQ OF FROZEN HB: PRELIMINARY ANNOTATIONS

Lung metastasis #1

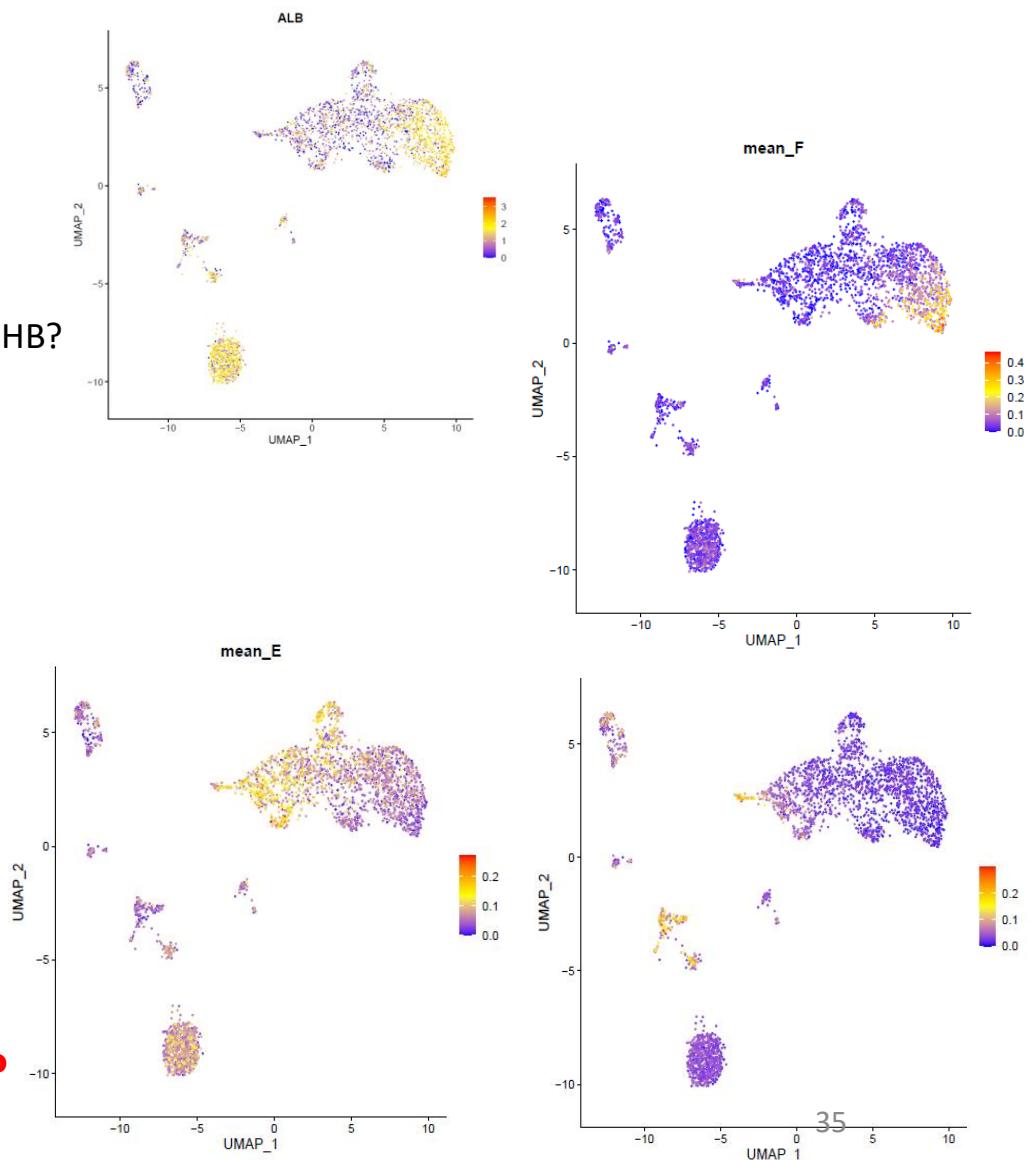


SN-RNASEQ OF FROZEN HB: PRELIMINARY ANNOTATIONS

Lung metastasis #1



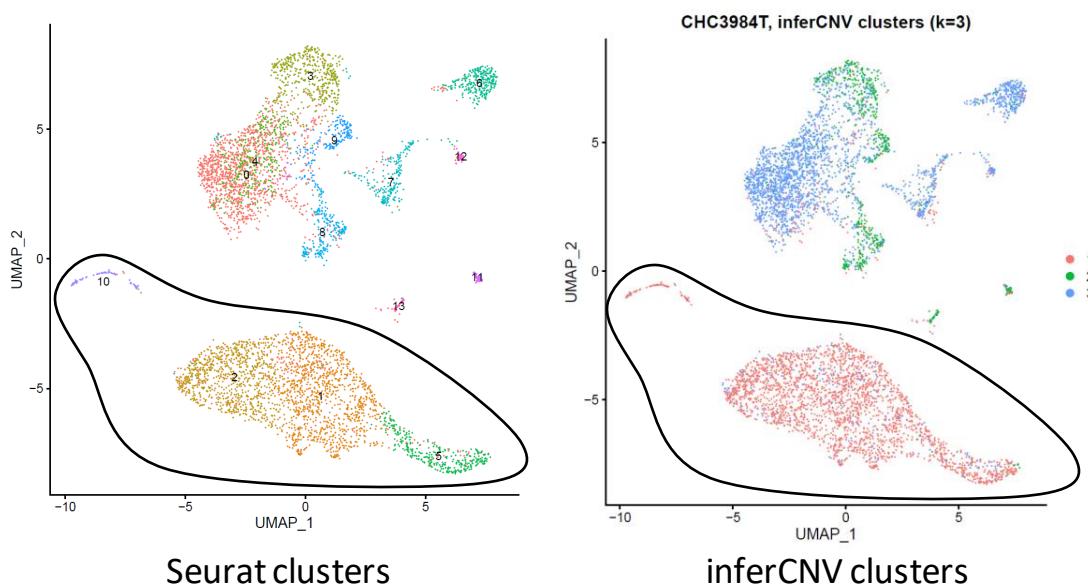
**Question: why would there be normal hepatic cells in lung metastasis?
+ blurred gene expression**



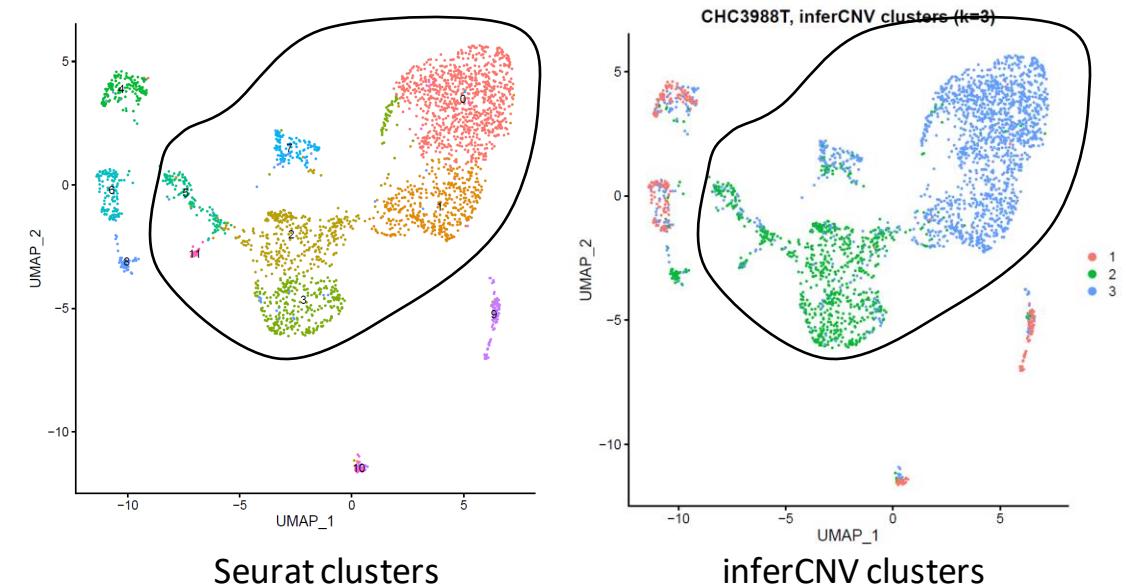
SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

Intersection between Seurat inferred tumor clusters and inferCNV results → keep cells which are **100% tumoral**

Primary tumor



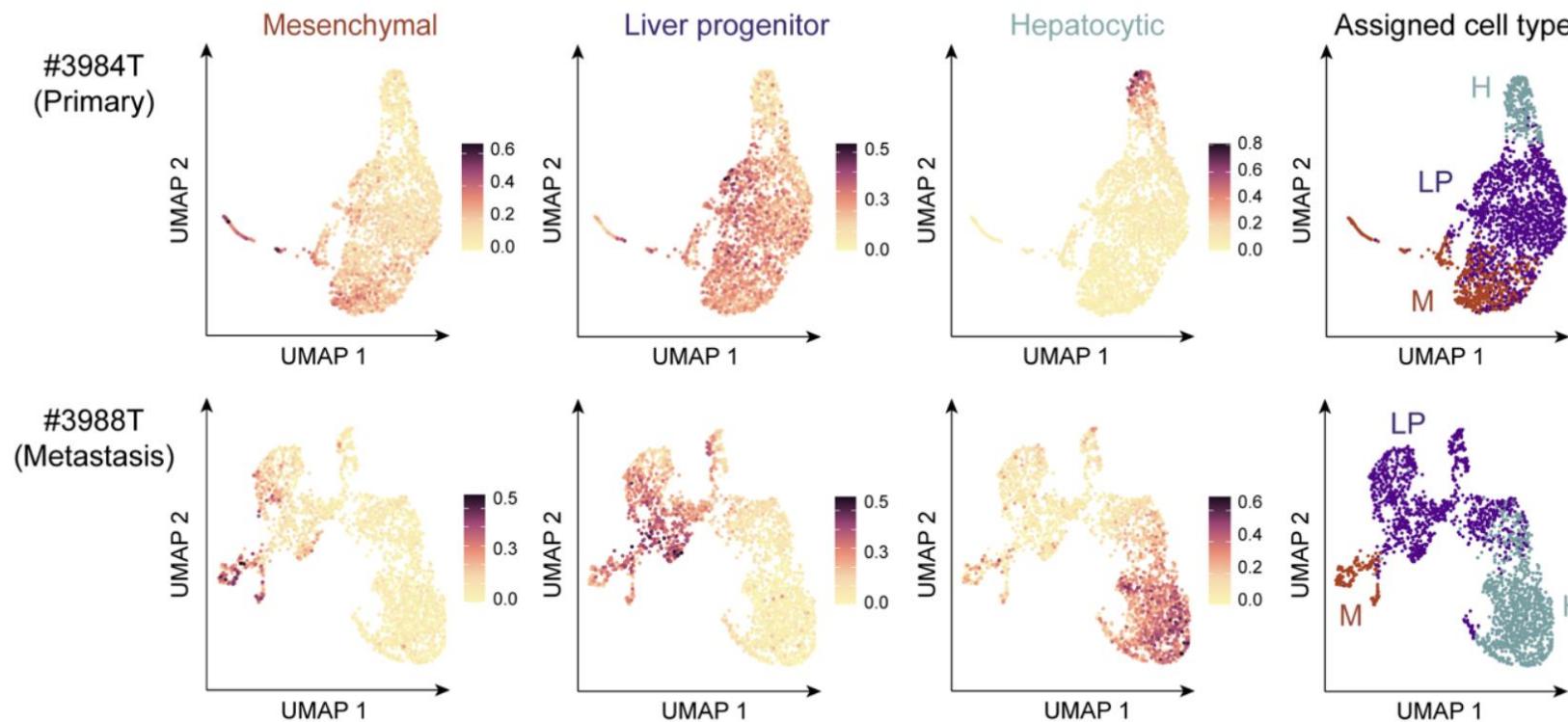
Metastasis #2



→ Metastasis #1 left out for the moment

SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF H/LP/M SUBTYPES

→ Gene signatures associated with the bulk transcriptomic clusters

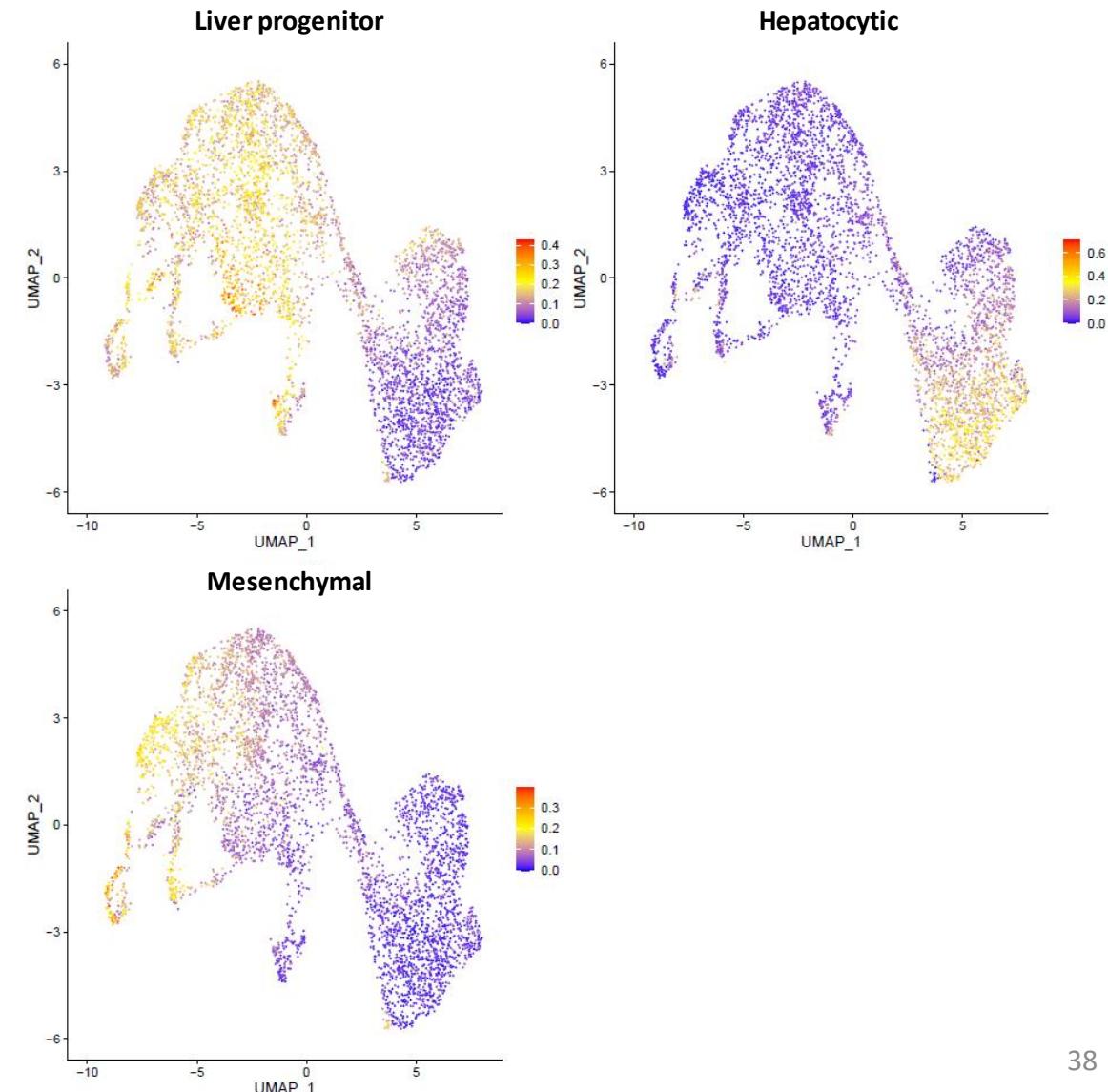
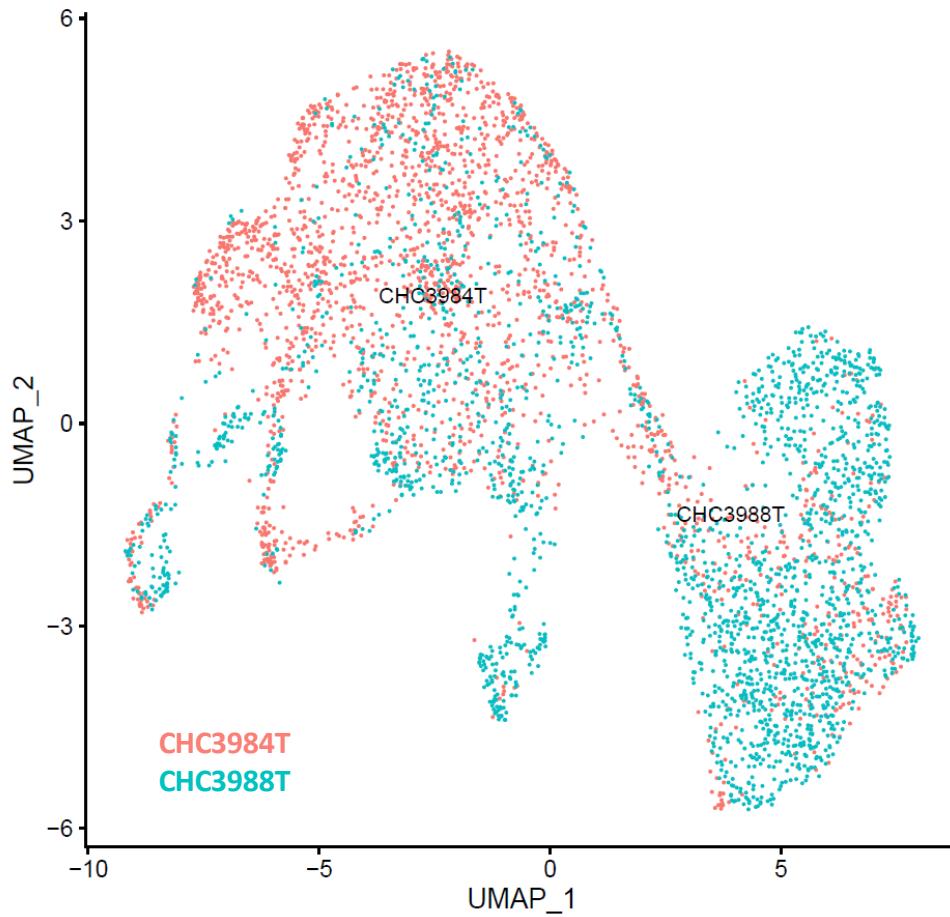


Hirsch et al., *Cancer Discov*, in press

- Identification in snRNA-seq data of 3 areas corresponding respectively to:
 - **Mesenchymal**
 - **Liver Progenitor**
 - **Hepatocytic** behaviours

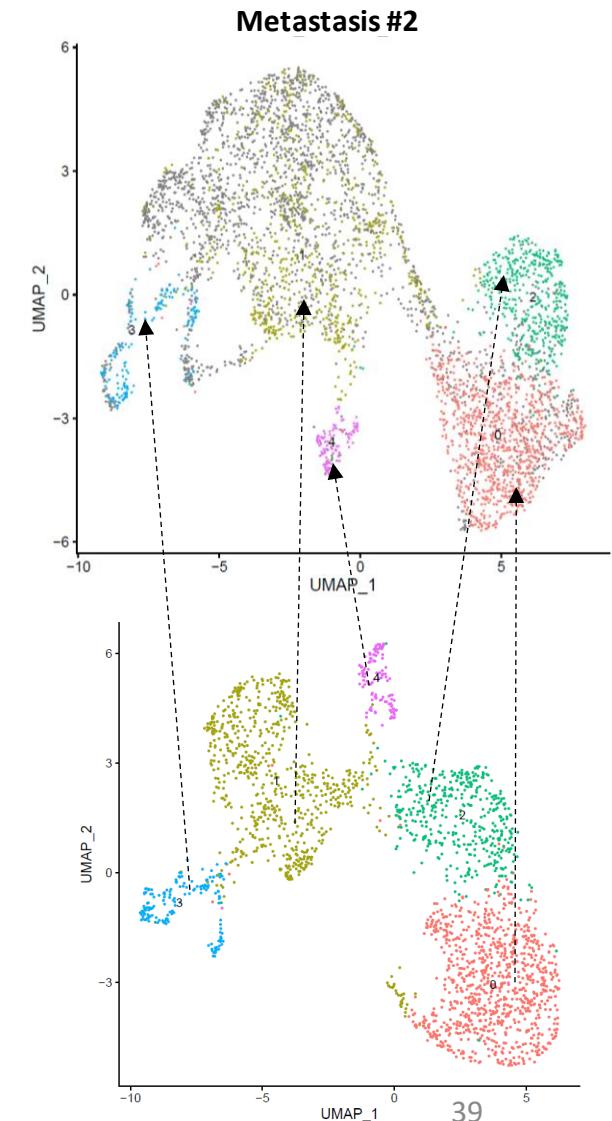
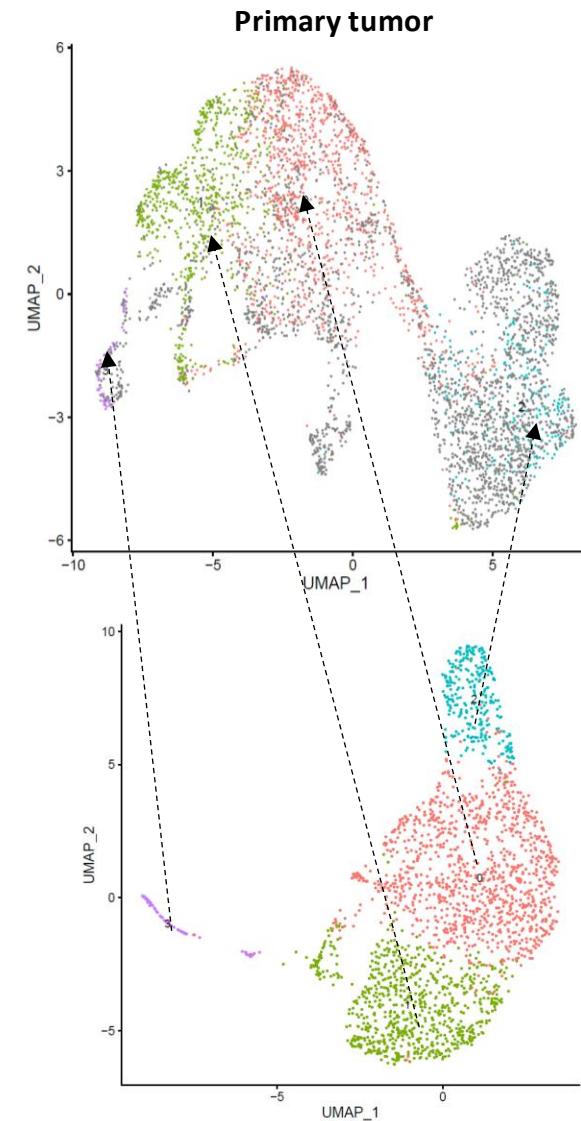
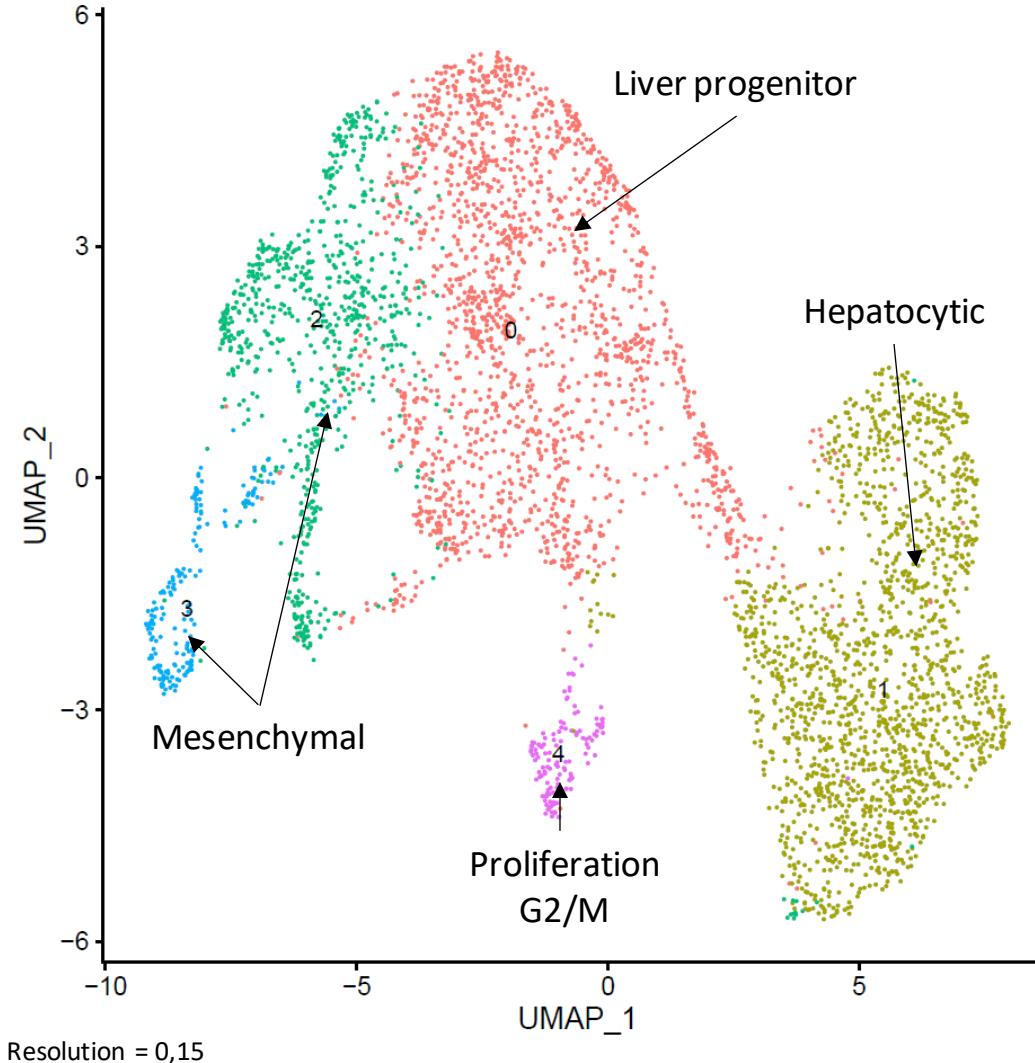
SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF H/LP/M SUBTYPES

→ Integration of the primary tumor and metastasis #2 (default Seurat integration pipeline)



SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF H/LP/M SUBTYPES

→ Integration of the primary tumor and metastasis #2 (default Seurat integration pipeline)

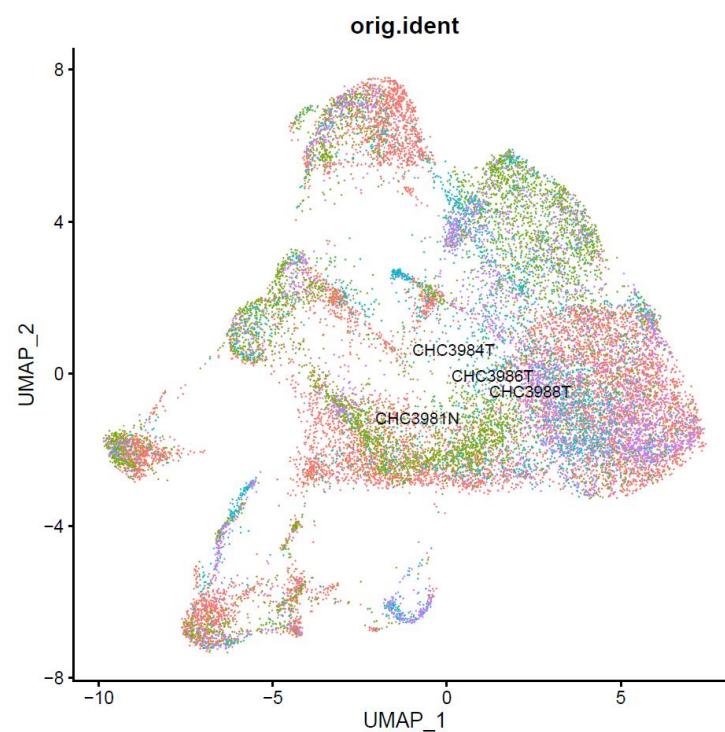


SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF H/LP/M SUBTYPES

→ Integration of the primary tumor and metastasis #2 (default Seurat integration pipeline)

Question: integration pipeline Seurat vs Harmony?

→ Integration of all 4 samples

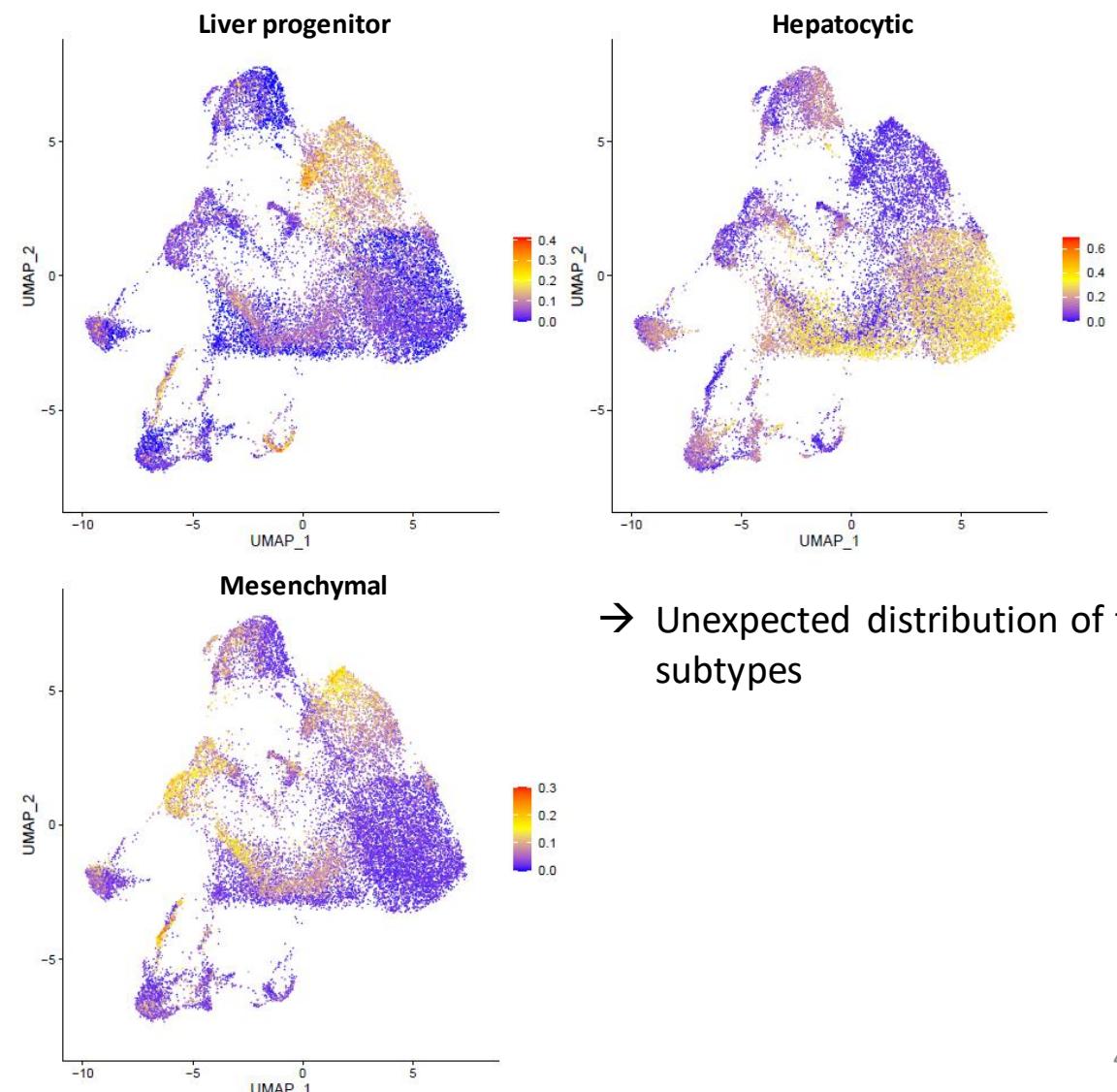


Non tumor sample

Primary tumor

Meatstasis #1

Metastasis #2



→ Unexpected distribution of the subtypes

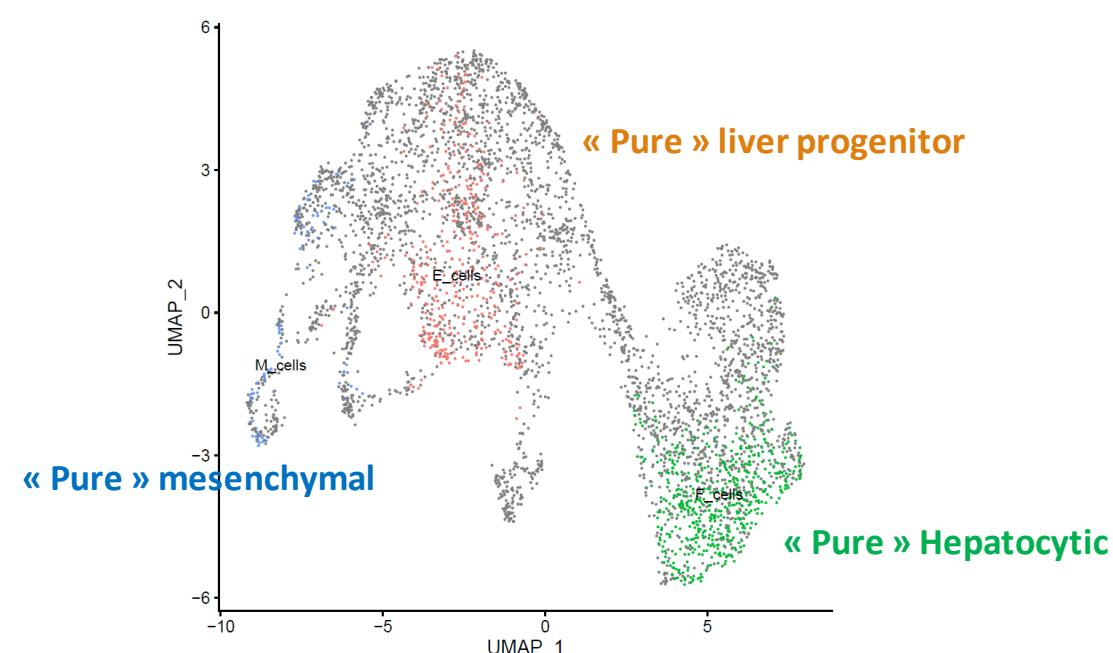
SN-RNASEQ OF FROZEN HB: DETERMINATION OF SPECIFIC MARKERS

Problematic to find relevant markers: some typical markers expressed in the whole cluster, some in one subgroup of cells only
→ probably cell populations under transition between the subtypes

Strategy:

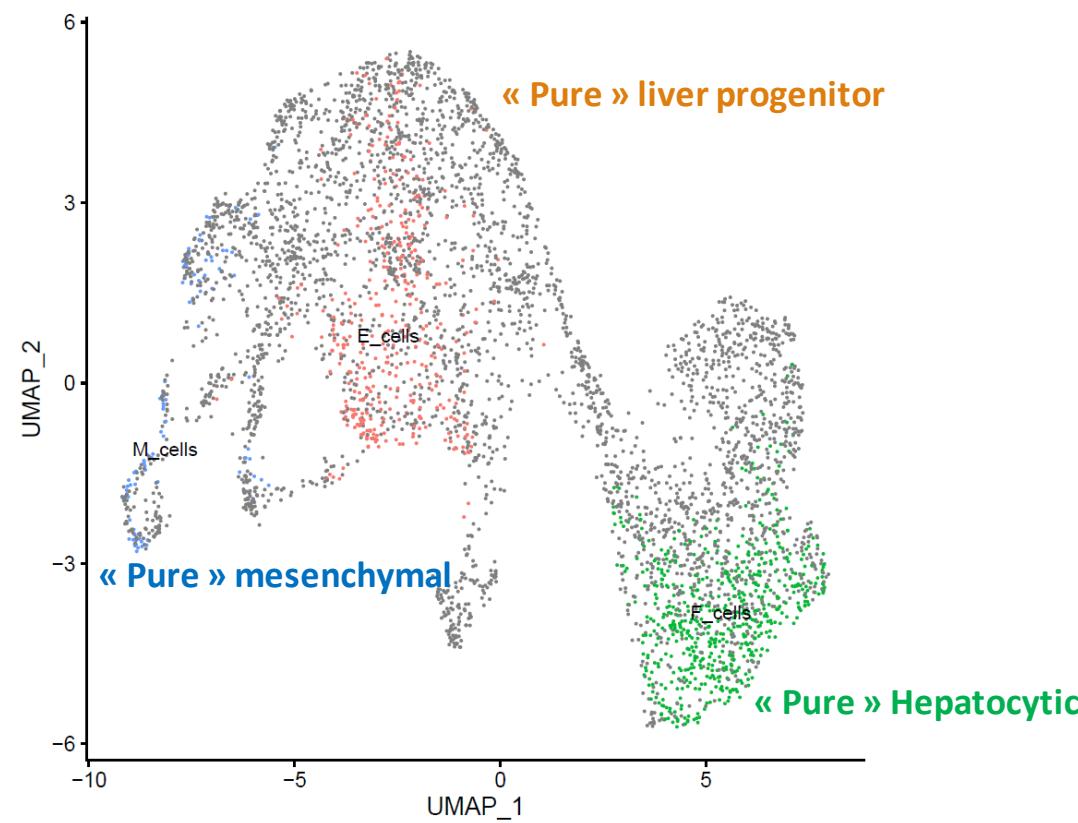
- Identify the **most likely « pure » cells** in each hepatocytic/liver progenitor/mesenchymal subgroup = thresholds on the average expression of the previous bulk markers
- **Differential expression** (Seurat, Wilcoxon test) between those « pure » cell subclusters
- Identification of relevant markers = **good detection in the « pure » subcluster** of interest, bad detection in other « pure » cells
- Restriction to markers that are also **differentially expressed in the bulk dataset**

→ 3/4 markers per subtype



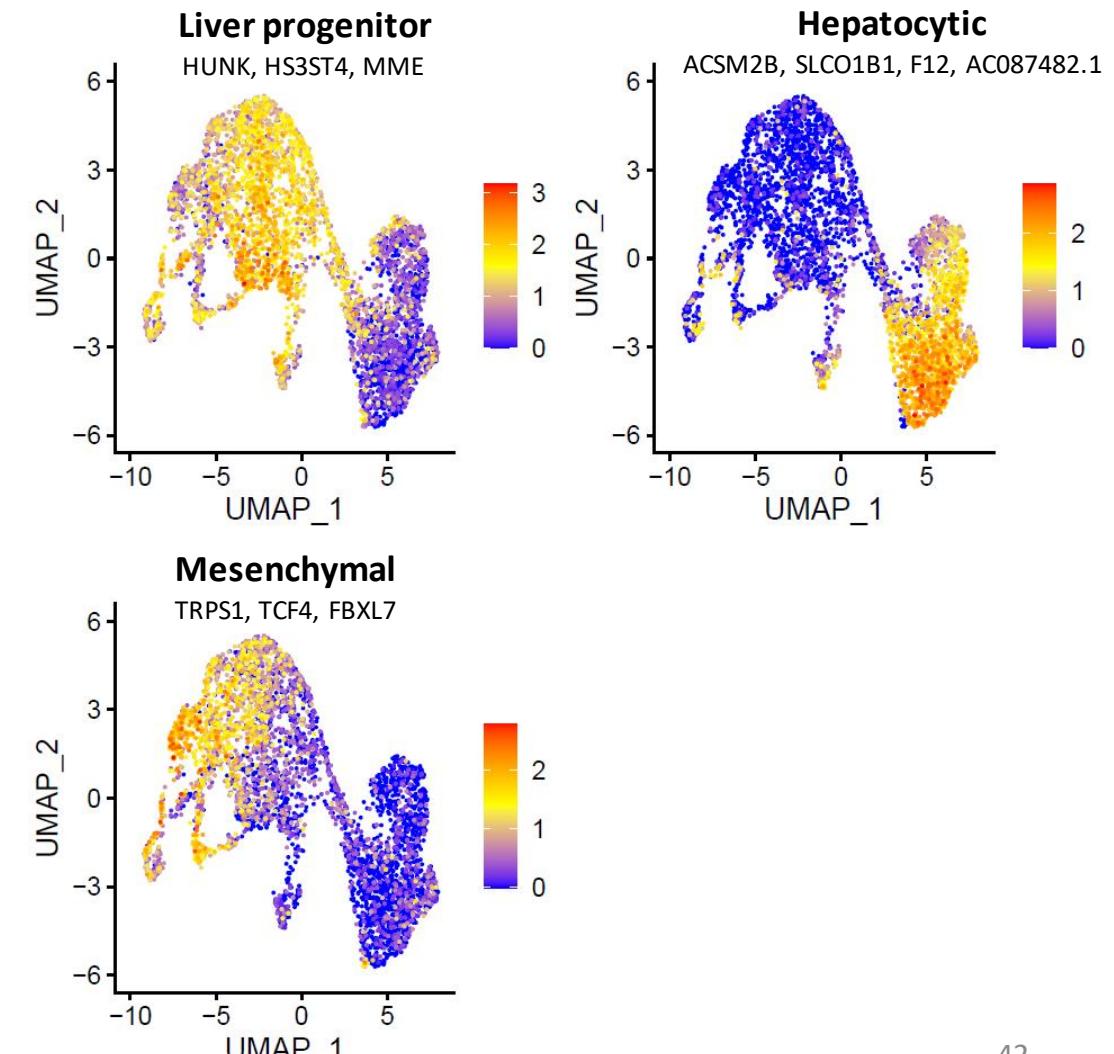
SN-RNASEQ OF FROZEN HB: DETERMINATION OF SPECIFIC MARKERS

Problematic to find relevant markers: some typical markers expressed in the whole cluster, some in one subgroup of cells only
→ probably cell populations under transition between the subtypes



Questions:

- How to find relevant genes for single-cell data?
- Mixed populations/transitions
- Liver progenitor/mesenchymal cells share similar markers



SN-RNASEQ OF FROZEN HB: UNDERSTAND TUMOR EVOLUTION

Cell plasticity → how can cells undergo transitions from one subtype to another?

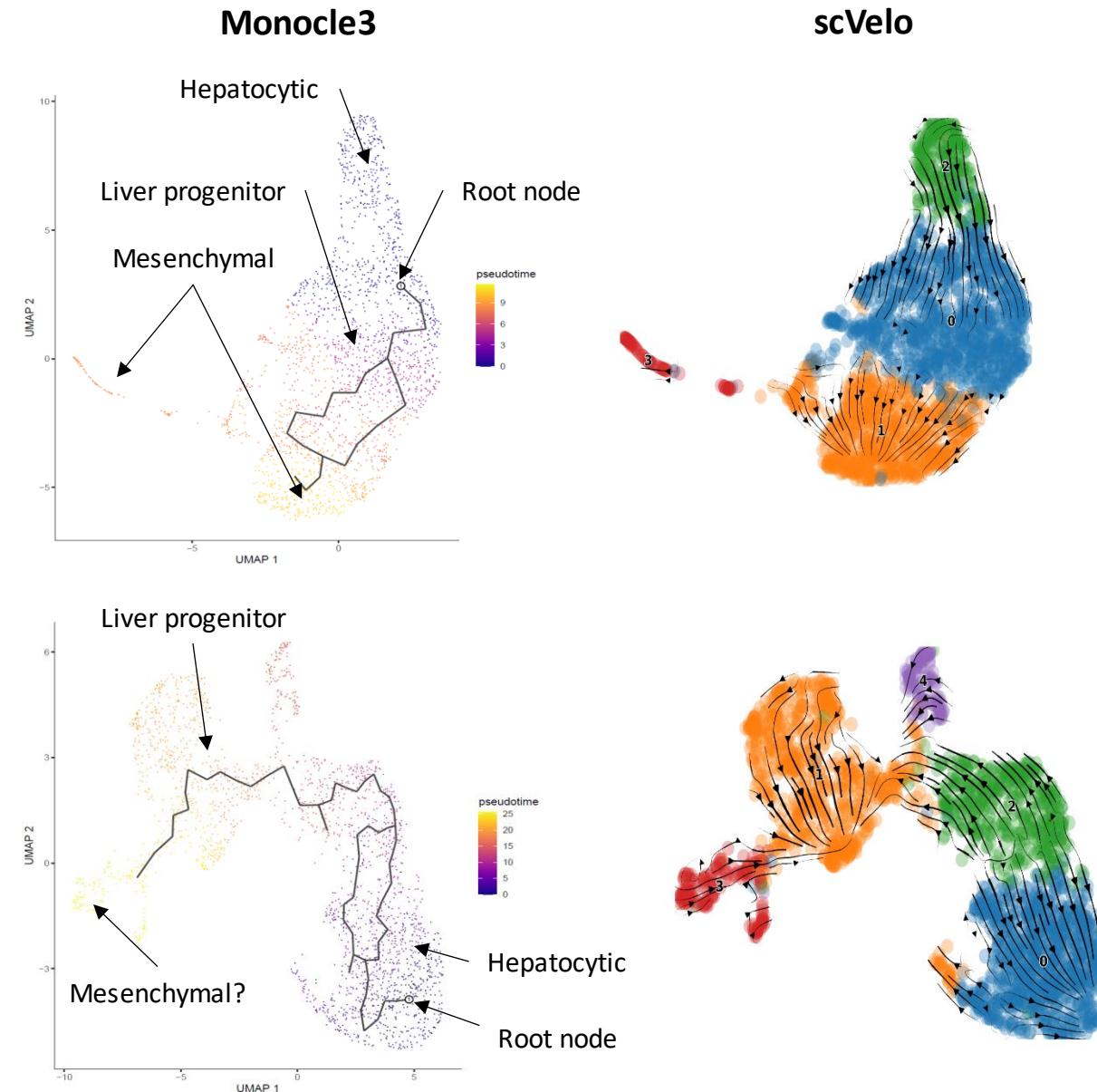
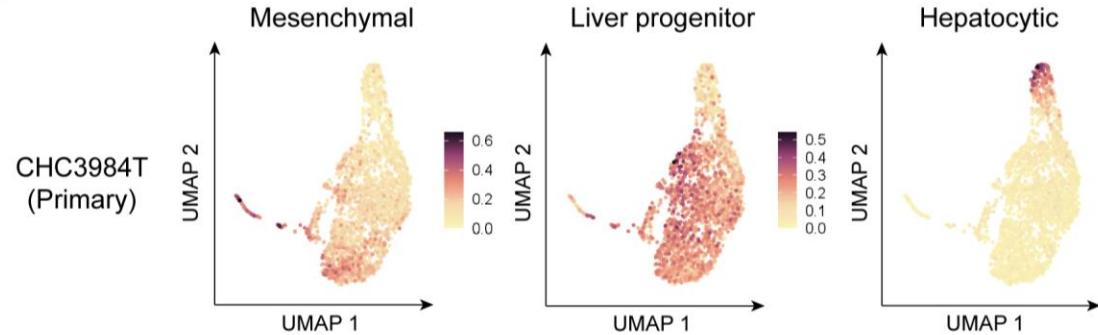


→ Use methods to infer cell trajectory

- **Monocle3:** trajectory, pseudotime
- **scVelo:** RNA velocity

SN-RNASEQ OF FROZEN HB: UNDERSTAND TUMOR EVOLUTION

b



SN-RNASEQ OF FROZEN HB: UNDERSTAND TUMOR EVOLUTION

Cell plasticity → how can cells undergo transitions from one subtype to another?



→ Use methods to infer cell trajectory

- Monocle3: trajectory, pseudotime
- scVelo: RNA velocity

Questions:

→ Monocle3: not appropriate for such a geometry? + need to choose the root node manually

→ RNA velocity: not adapted to single-nucleus RNA-seq? (ratio of unspliced/spliced RNA)

SN-RNASEQ OF FROZEN HB: OUR NEEDS AND QUESTIONS

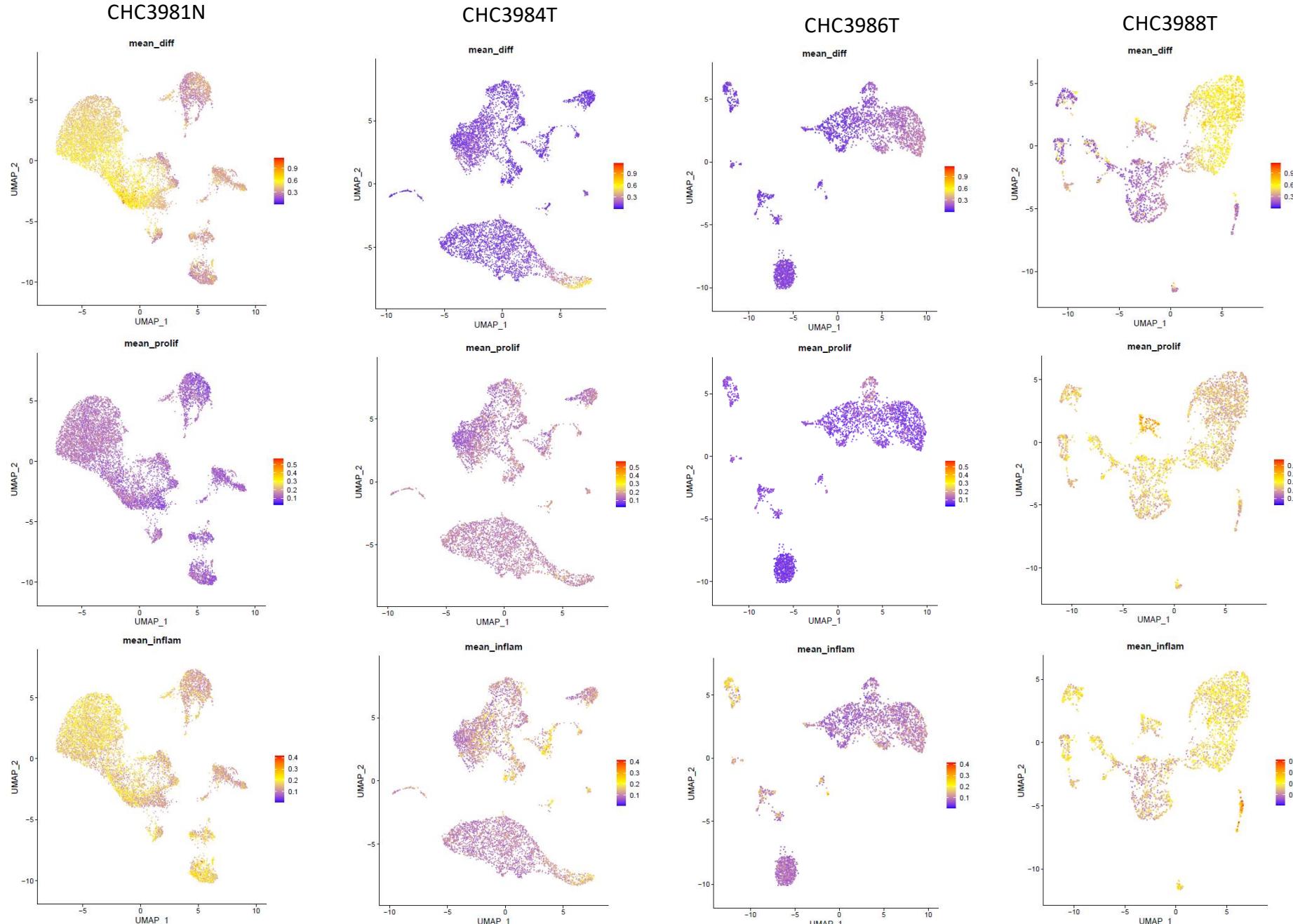
→ Need to clearly identify the **single-nuclei trajectory** to identify transitions between hepatocytic / liver progenitor / mesenchymal subtypes to characterize **HB tumor evolution and cellular plasticity**

Recap of the questions

- Methods:
 - Integration of several samples: is it possible between samples sequenced at different times? Is Harmony better than Seurat for integration?
 - How to deal with cell cycle regression?
 - What about combined snRNA-seq/ATAC-seq?
- Biological interpretation:
 - InferCNV: unexpected deletions?
 - Normal cells contaminated by tumoral RNA in the primary tumor?
 - Unexpected copy-number alteration profiles of metastasis #1, presence of normal hepatocytes?
 - How to find appropriate markers for each subtype (particularly when there are potential cell-type transitions)?
 - Which methods/strategies to identify single-cell trajectories?

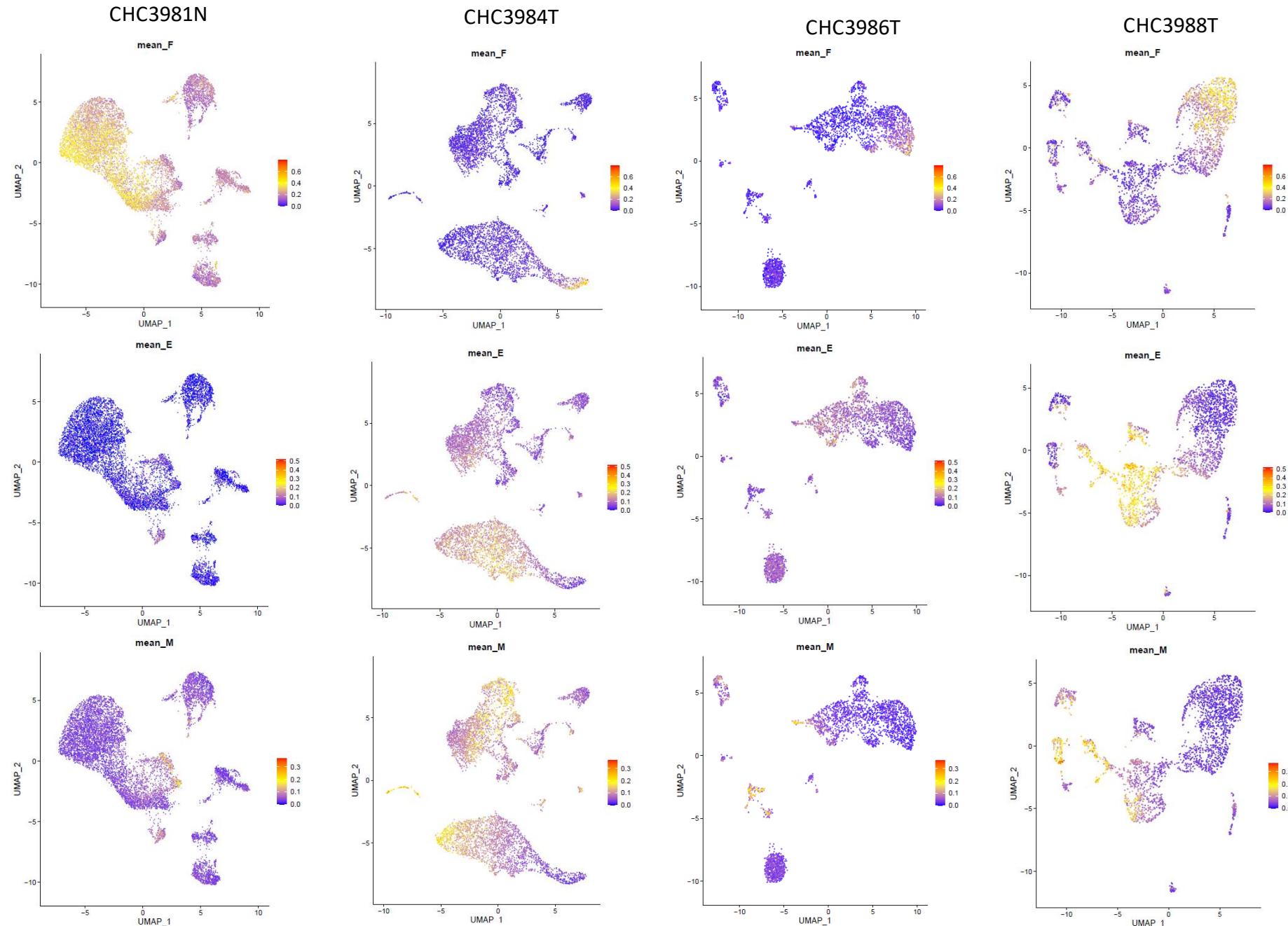
THANKS FOR YOUR ATTENTION!

SCTransform no regression



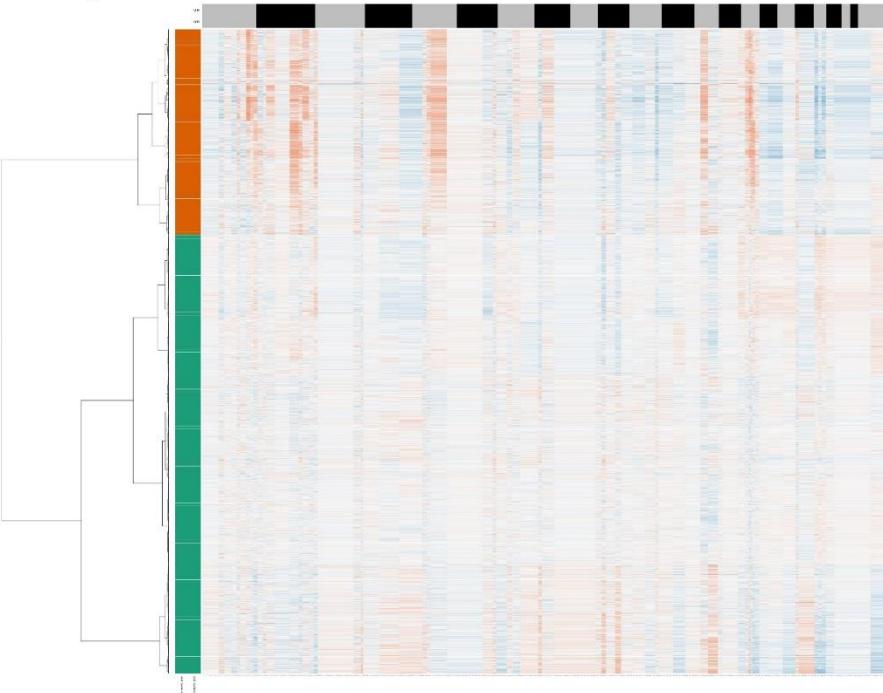
SCTransform no regression

!/\ cells sorted by max exp



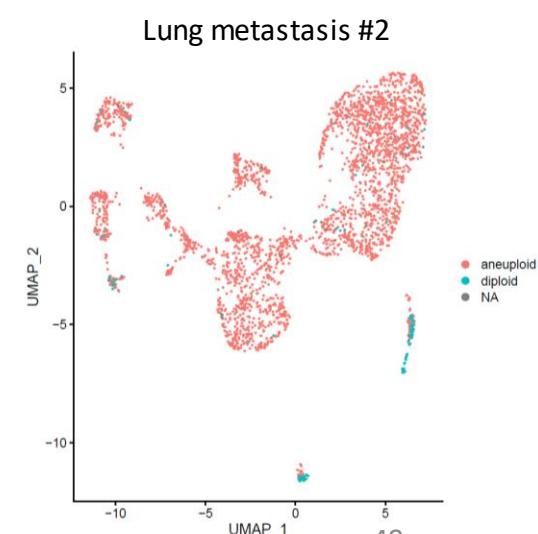
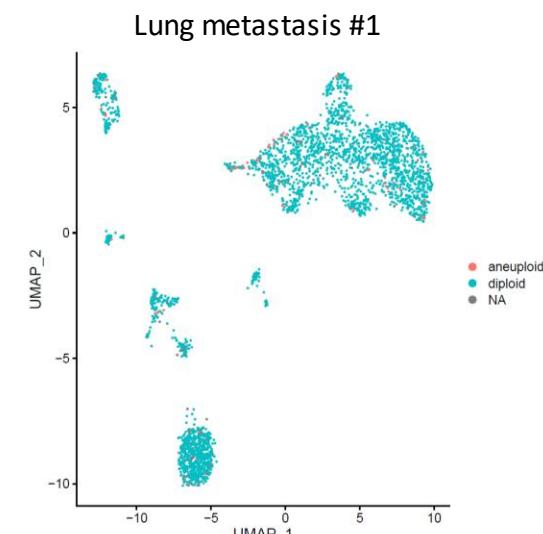
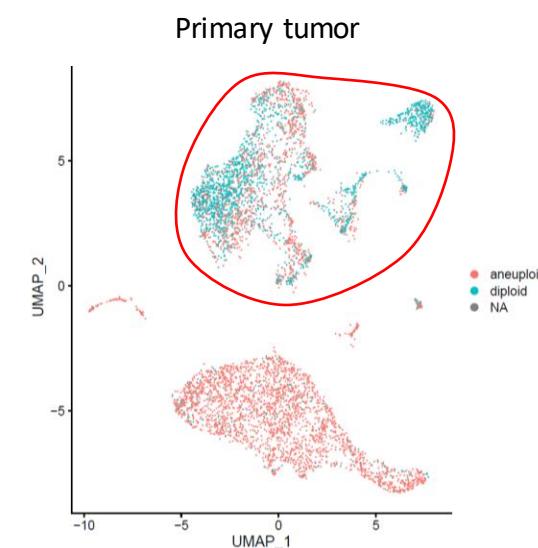
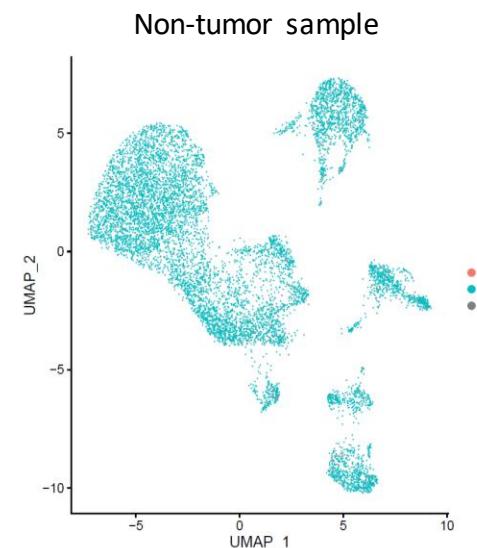
SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

CopyKAT (default parameters)



Aneuploid cells (tumoral)

Diploid cells (normal)

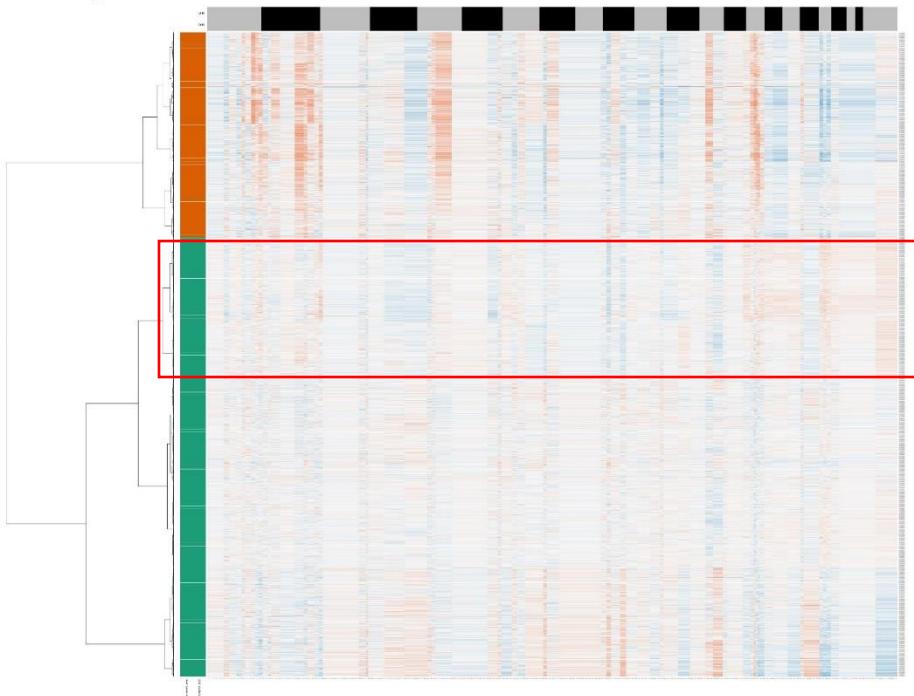


Questions:

- Assumed normal cells in the primary tumor = mix of aneuploid/diploid cells?

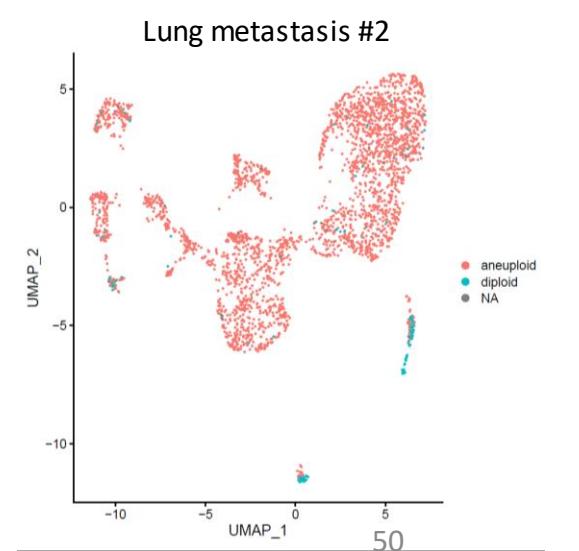
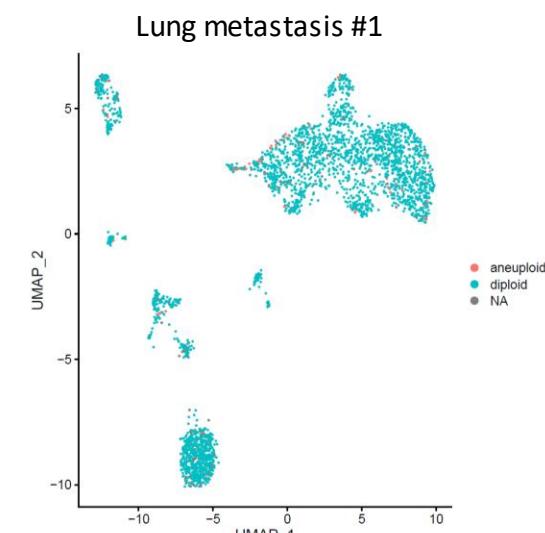
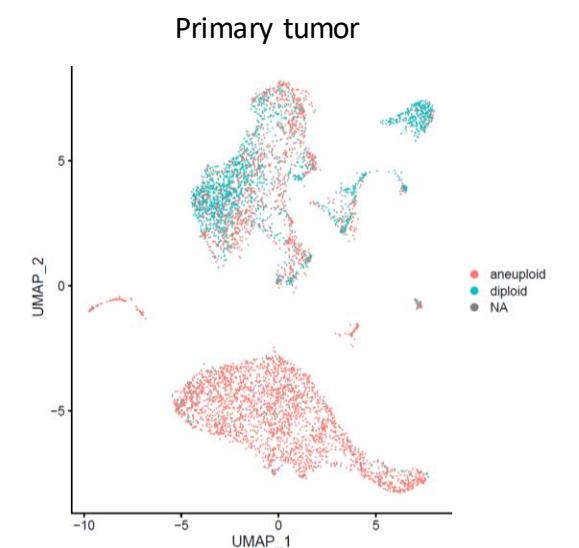
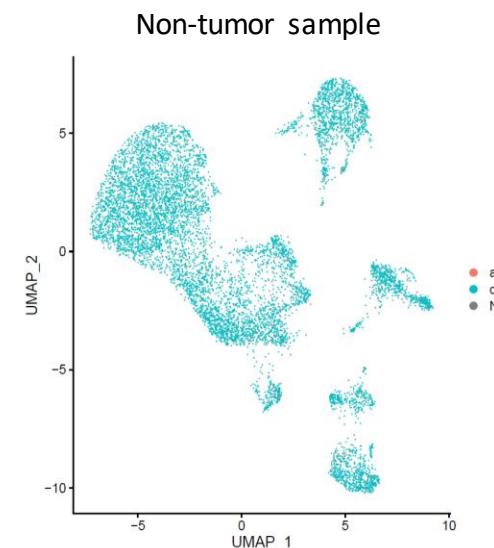
SN-RNASEQ OF FROZEN HB: IDENTIFICATION OF TUMOR CELLS

CopyKAT (default parameters)



Aneuploid cells (tumoral)

Diploid cells (normal)



Questions:

- Assumed normal cells in the primary tumor = mix of aneuploid/diploid cells?
- Lung metastasis #1: copy number profile too diluted?