

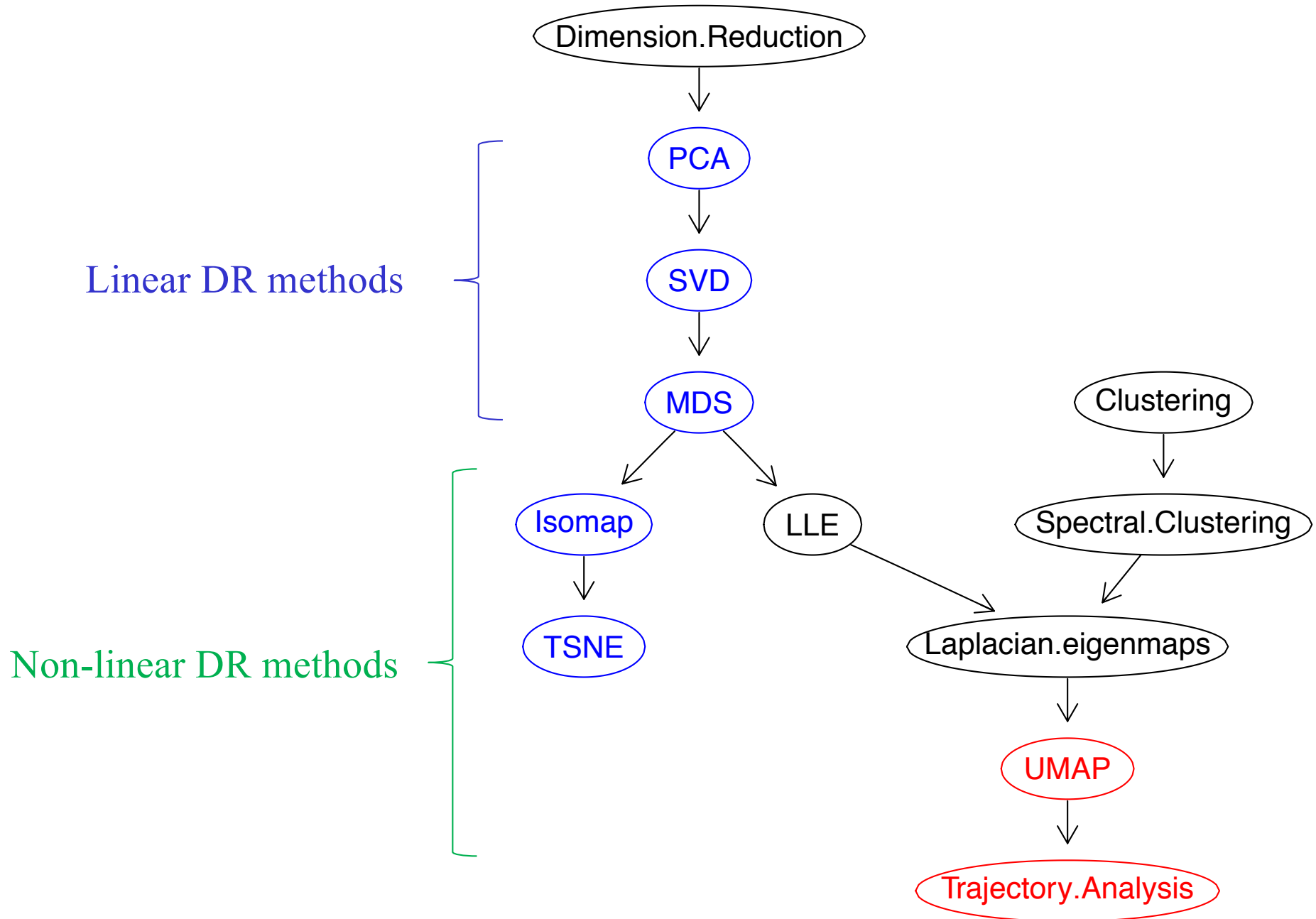
Dimension Reduction Methods: From PCA to TSNE and UMAP

Maxwell Lee

High-dimension Data Analysis Group
Laboratory of Cancer Biology and Genetics
Center for Cancer Research
National Cancer Institute

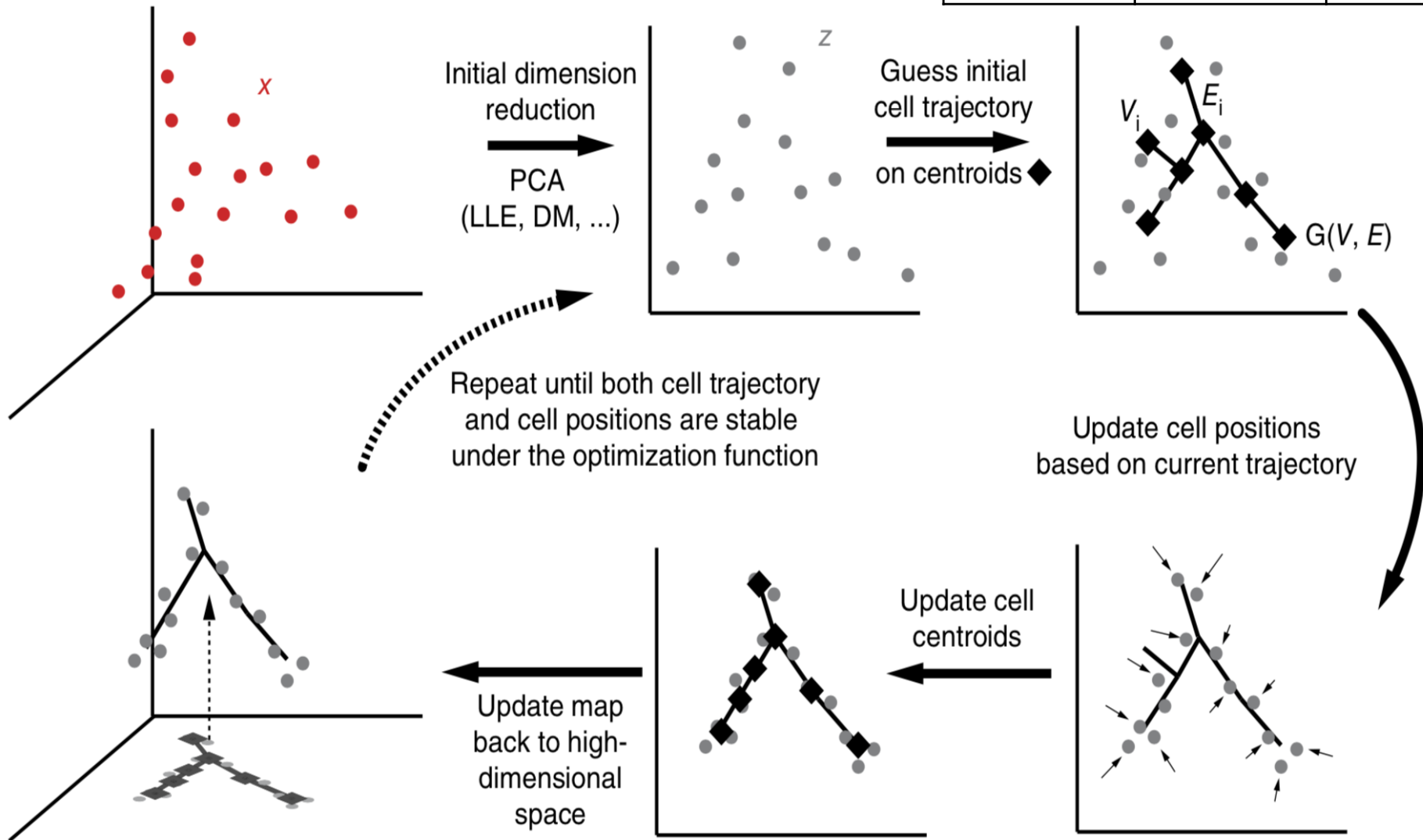
June 4, 2020

Road Map for Dimension Reduction Methods

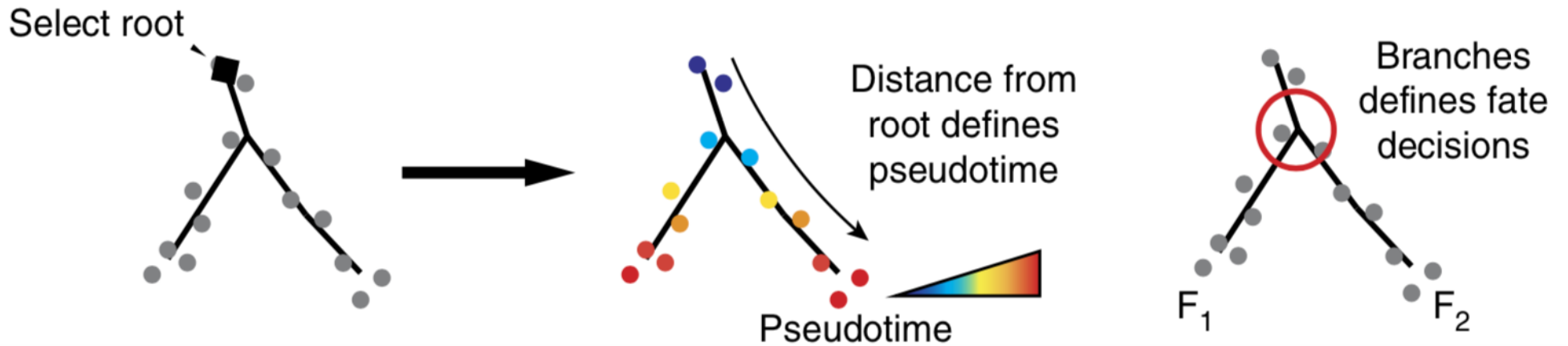


Trajectory Analysis (Monocle 2)

	monocle2	monocle3
DR	TSNE	UMAP
Graph	Tree	Graph
Partition	no	yes

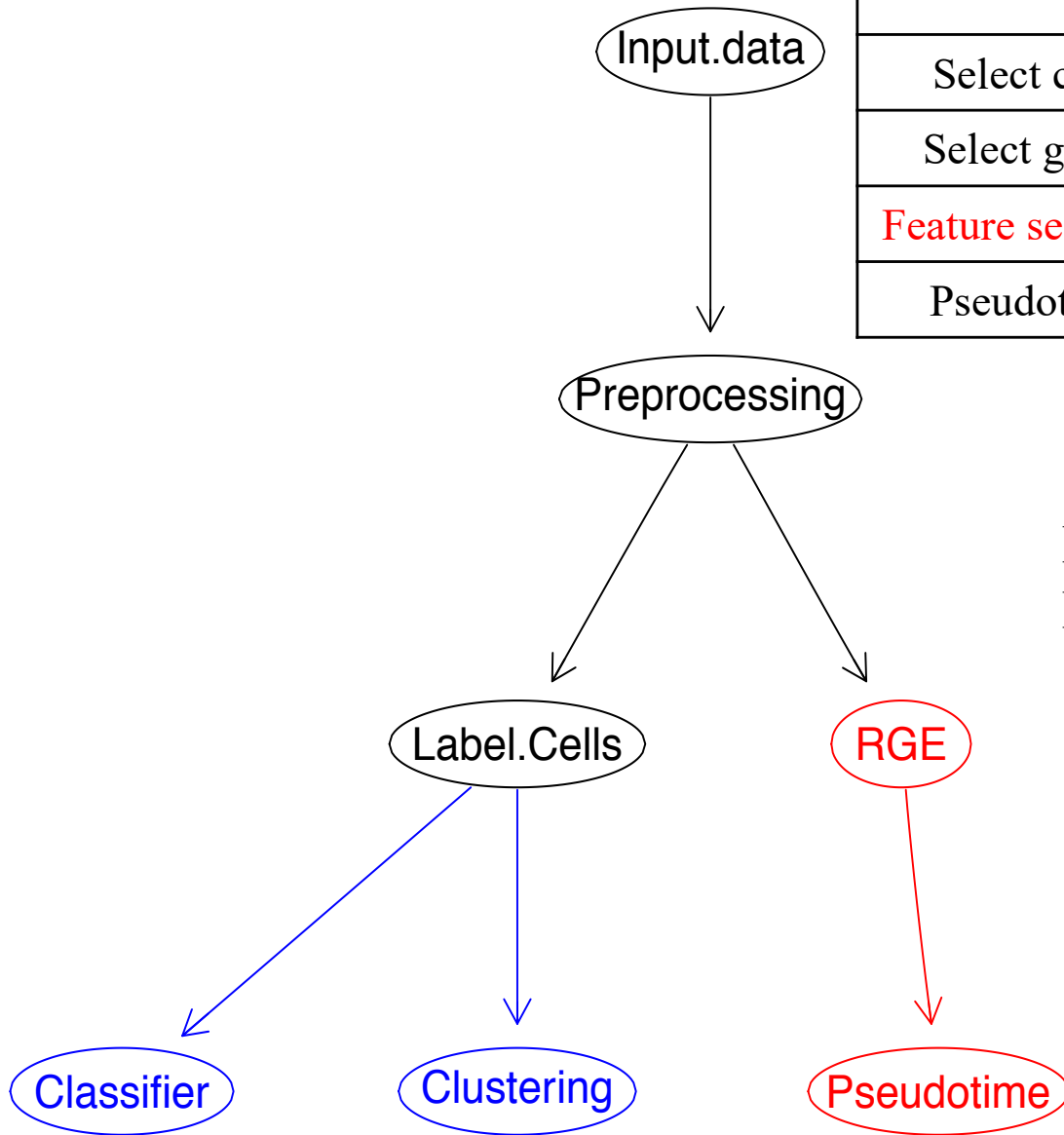


Pseudotime (Monocle 2)



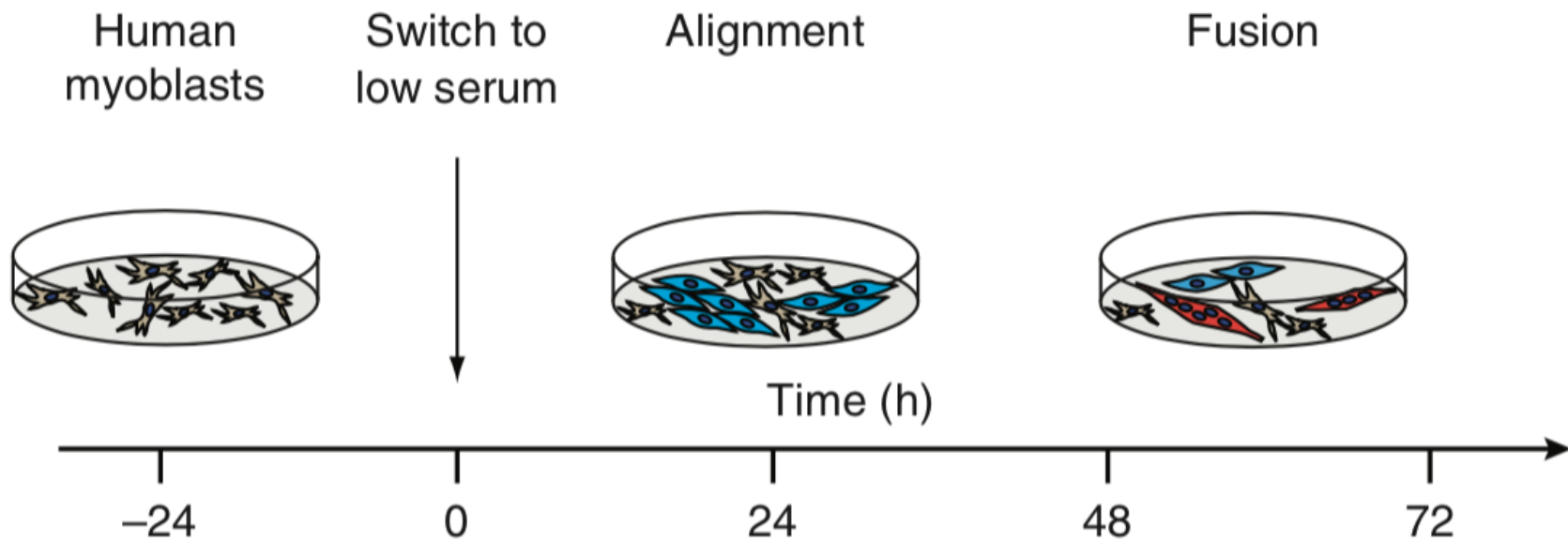
Flowchart of Trajectory Analysis with Monocle Package

	function	Description
Select cells	QC cells	
Select genes	variable genes	
Feature selection	DEG	Select subspace
Pseudotime	Select root	



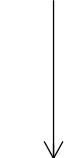
DEG: differential expression gene
RGE: reversed graph embedding

Human Skeletal Muscle Myoblasts (HSMM) Data

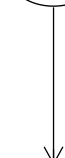


Tree Topology of HSMM Myoblast Data

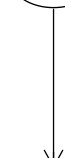
h0



h24



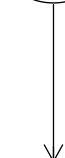
h48



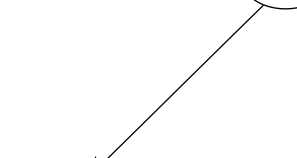
h72

linear

h0



i



h24

h48



h72

branch 1

h0



h24



i



h48

h72

branch 2

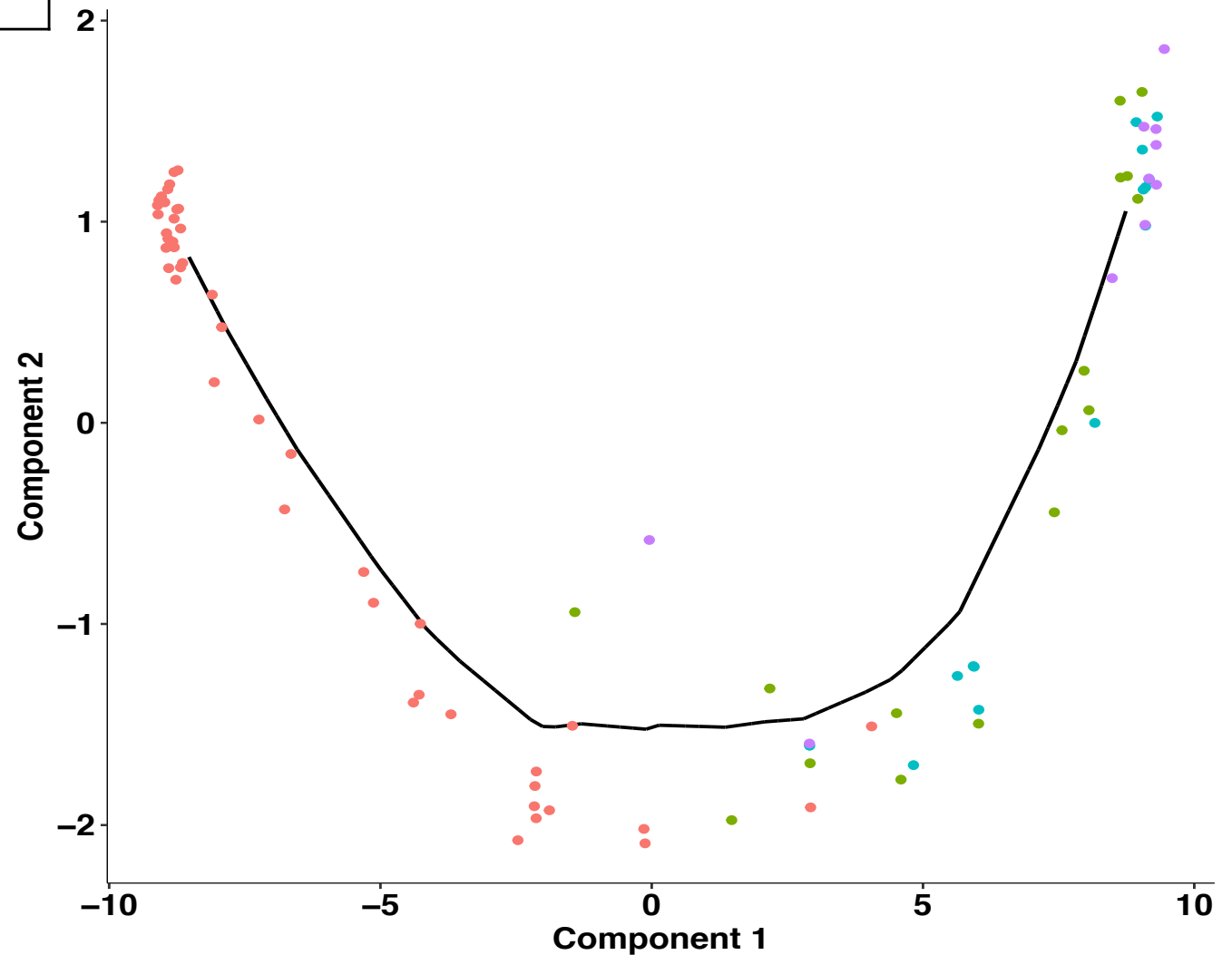
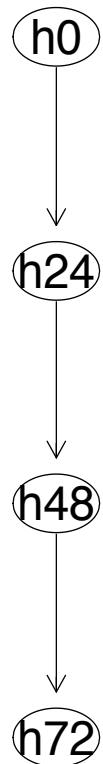
i: internal node

Trajectory Analysis of HSMM Data with Monocle 2

Media	Hours	freq
GM	0	45
DM	24	16
DM	48	14
DM	72	10

DEG: Media
nfeature: 419 genes

Hours ● 0 ● 24 ● 48 ● 72

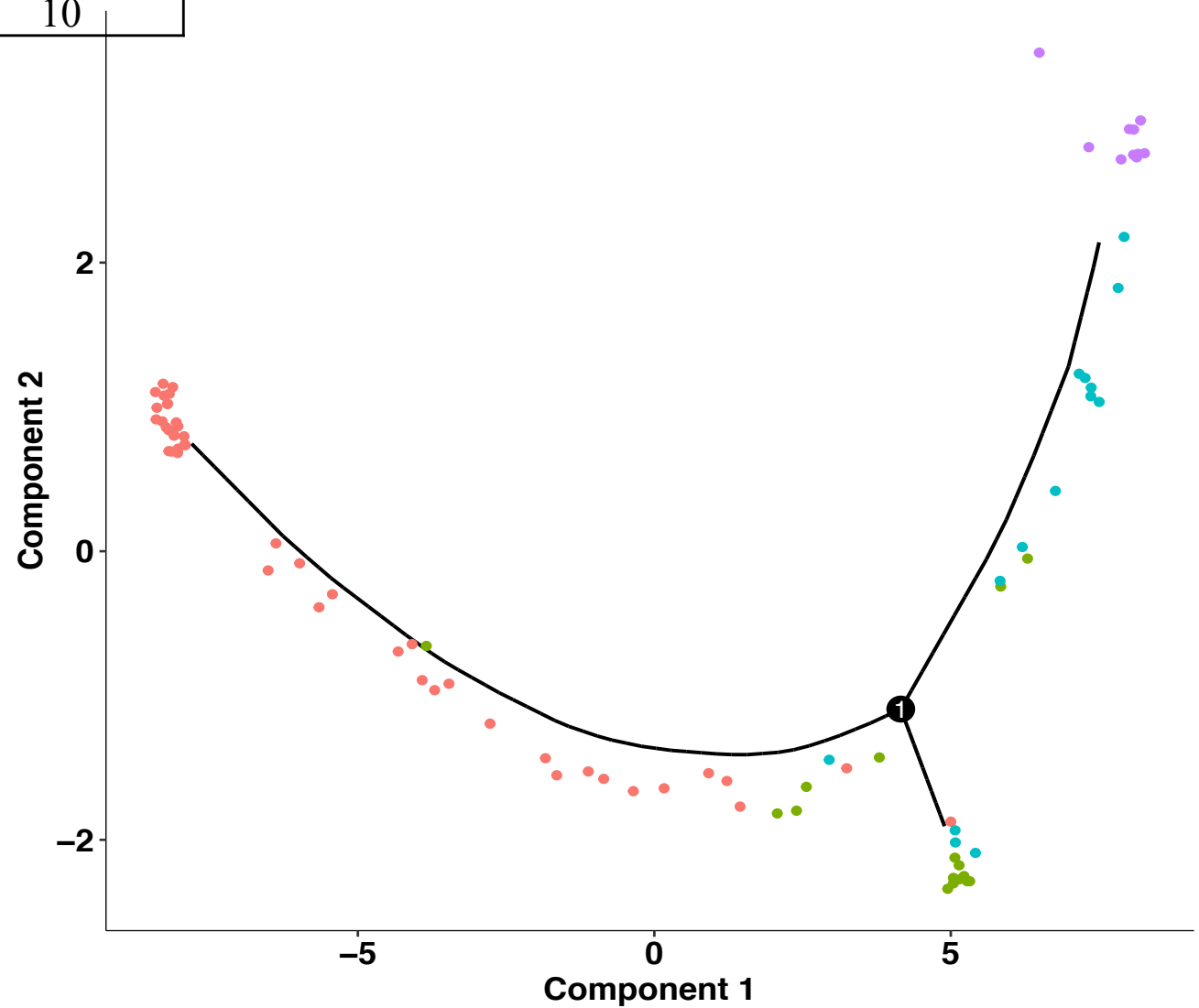
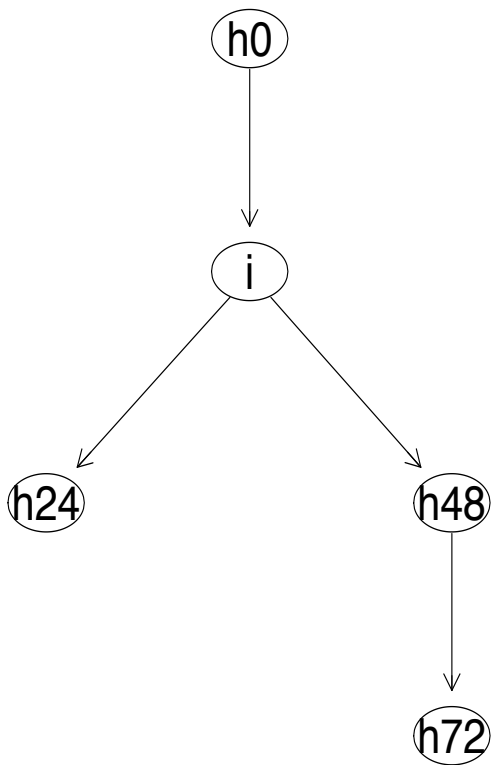


Trajectory Analysis of HSMM Data with Monocle 2

Media	Hours	freq
GM	0	45
DM	24	16
DM	48	14
DM	72	10

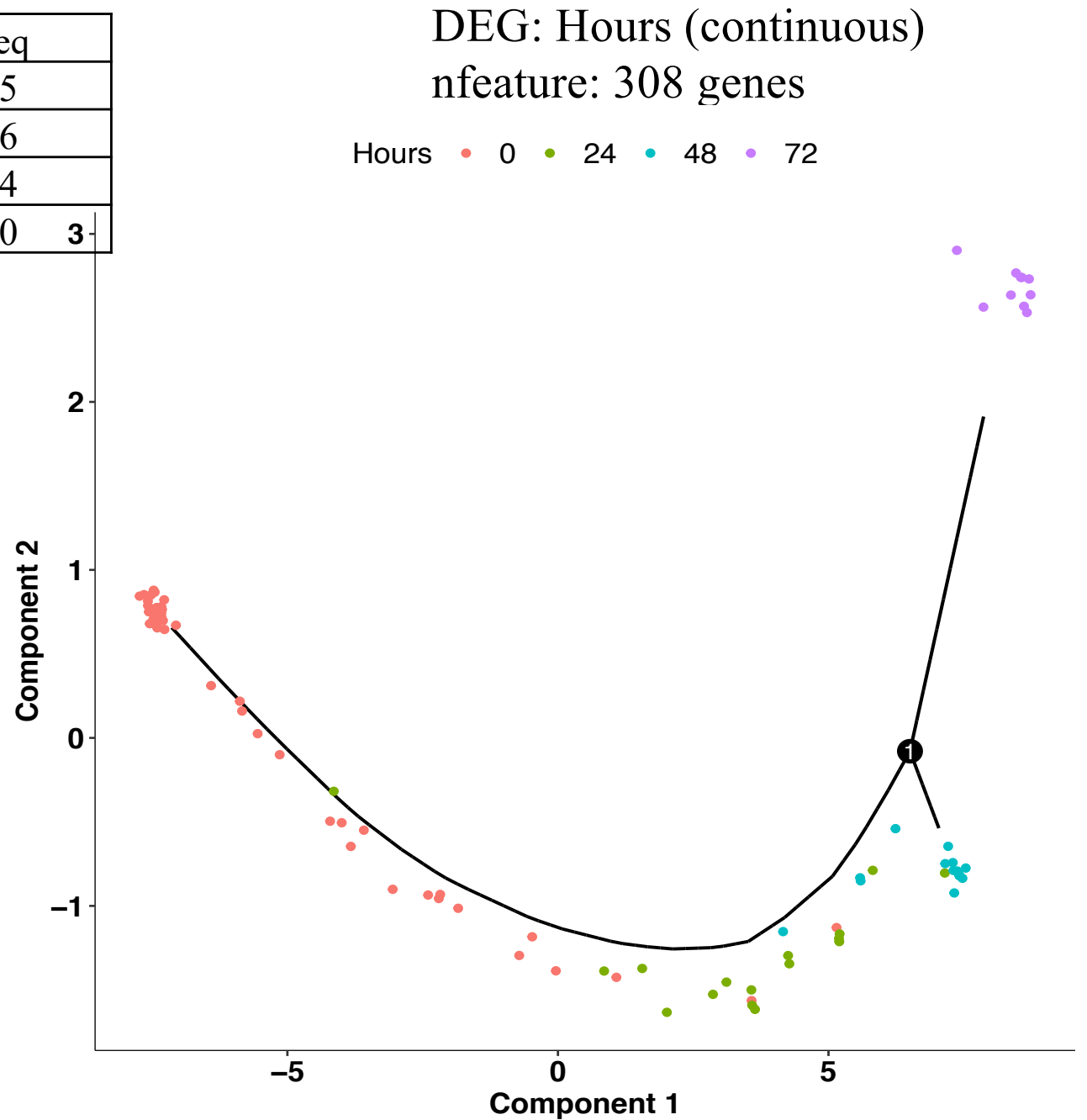
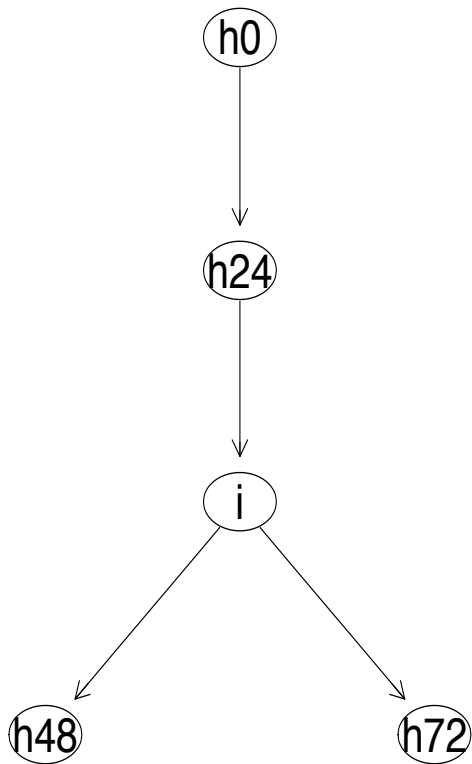
DEG: Hours (categorical)
nfeature: 385 genes

Hours • 0 • 24 • 48 • 72



Trajectory Analysis of HSMM Data with Monocle 2

Media	Hours	freq
GM	0	45
DM	24	16
DM	48	14
DM	72	10

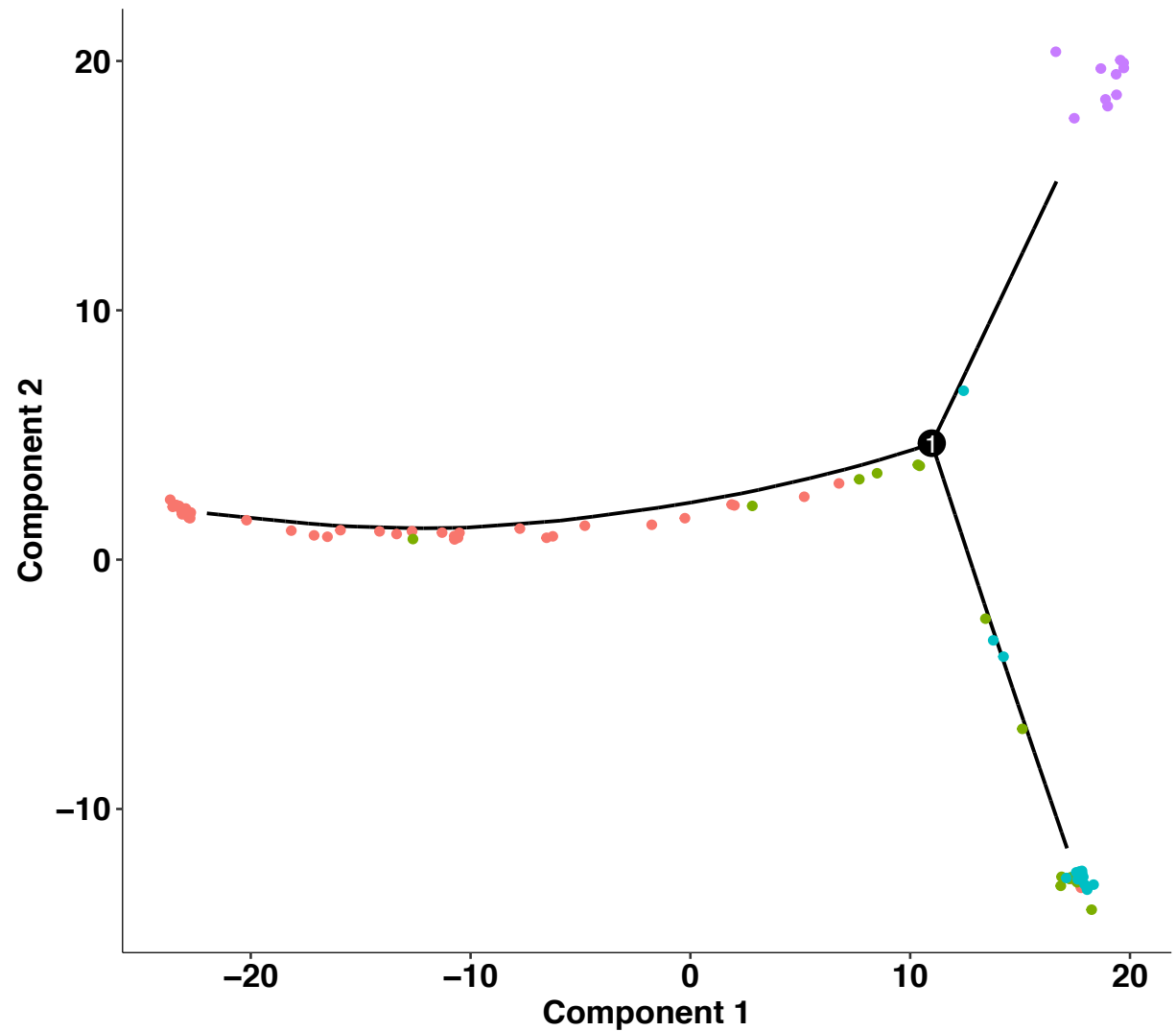
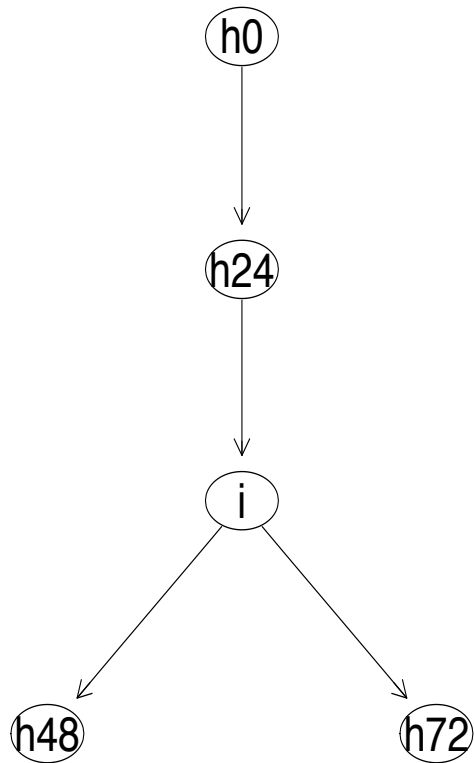


Trajectory Analysis of HSMM Data with Monocle 2

Media	Hours	freq
GM	0	45
DM	24	16
DM	48	14
DM	72	10

DEG: no filter
nfeature: 14224 genes

Hours ● 0 ● 24 ● 48 ● 72

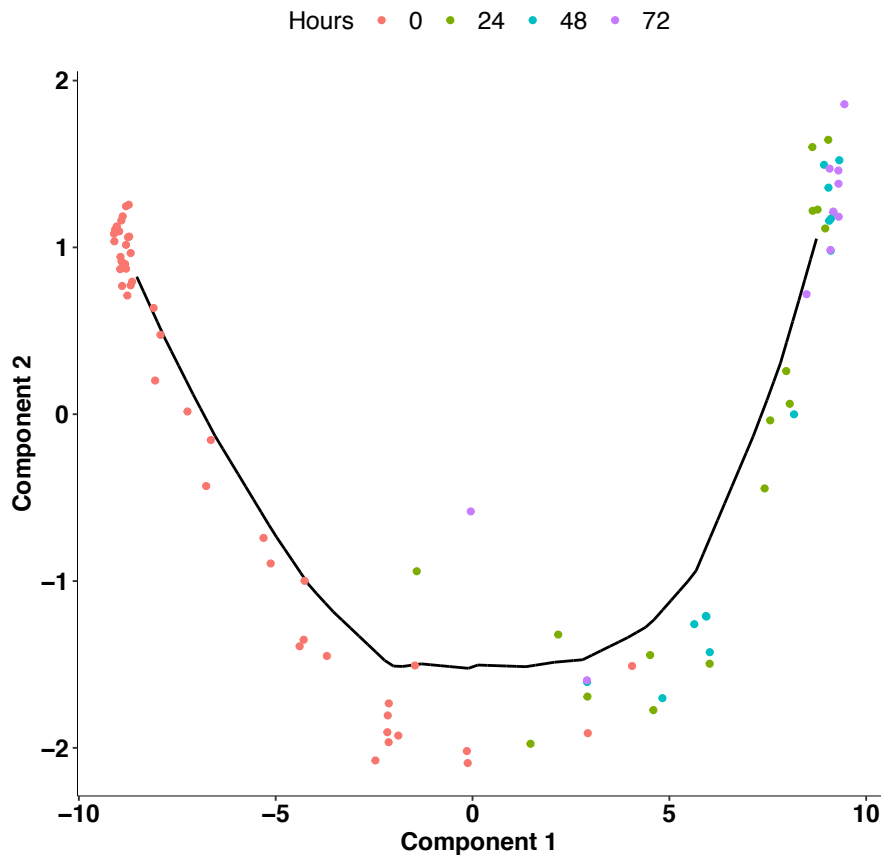


Trajectory Analysis of HSMM Data with Monocle 3

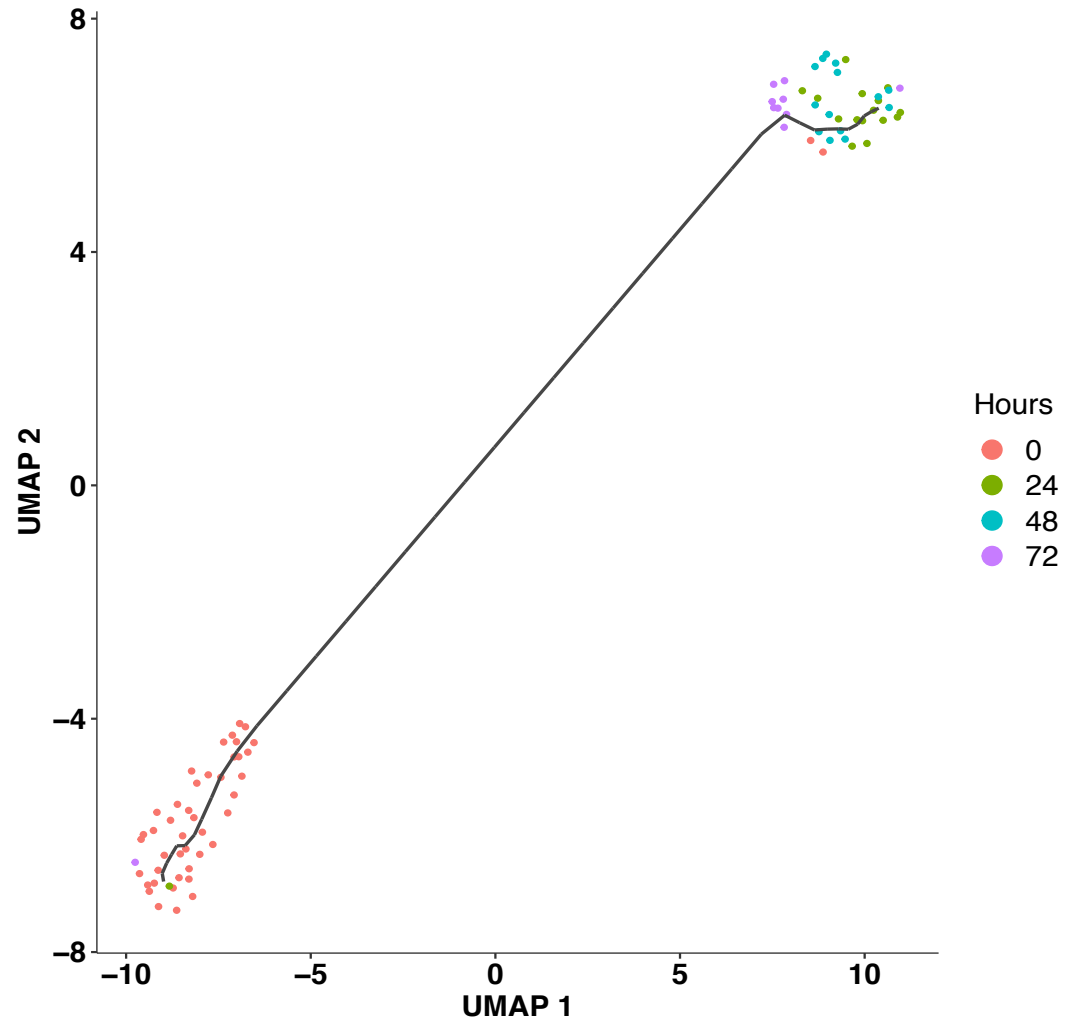
Media	Hours	freq
GM	0	45
DM	24	16
DM	48	14
DM	72	10

DEG: Media
nfeature: 419 genes

Monocle 2



Monocle 3



Trajectory Analysis of HSMM Data with Monocle 3

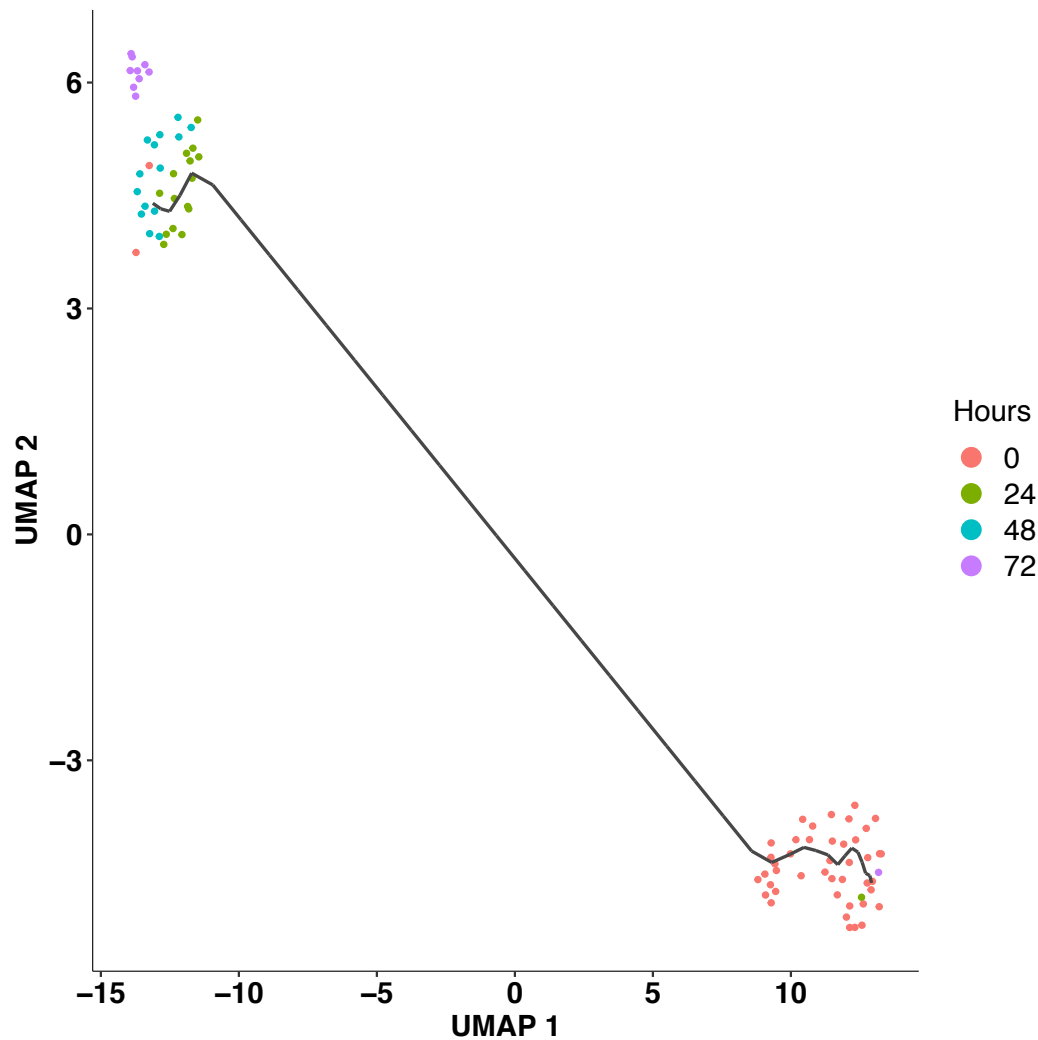
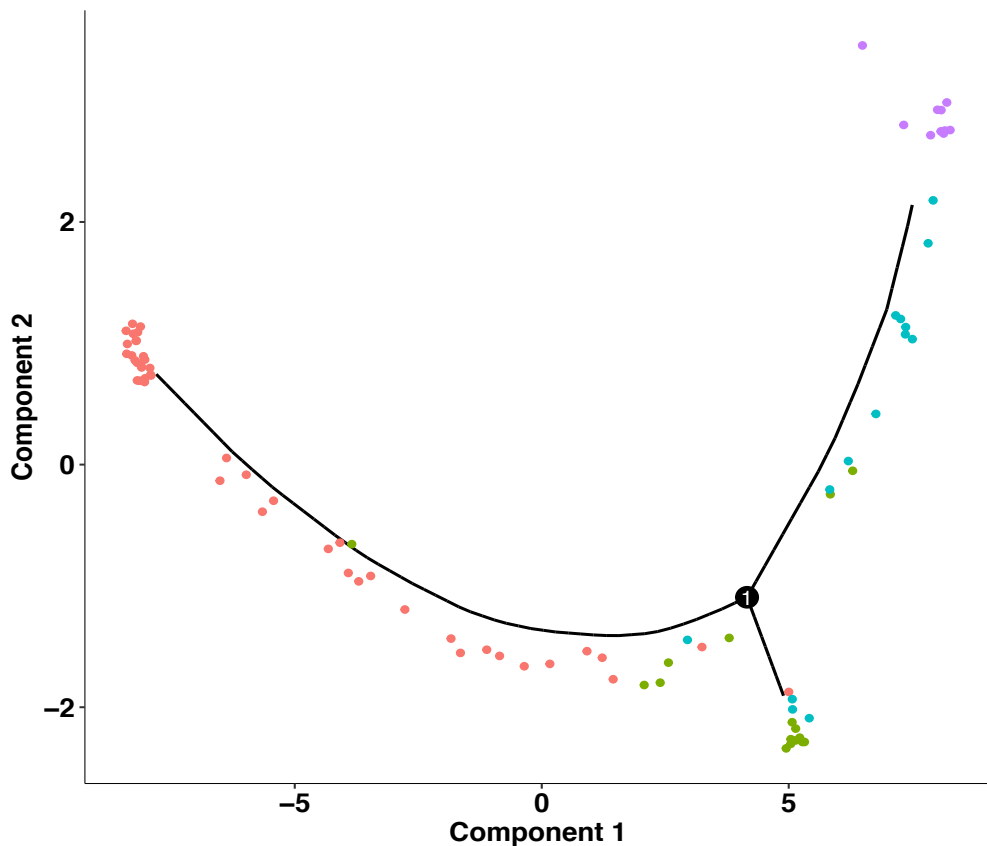
Media	Hours	freq
GM	0	45
DM	24	16
DM	48	14
DM	72	10

DEG: Hours (categorical)
nfeature: 385 genes

Monocle 3

Monocle 2

Hours ● 0 ● 24 ● 48 ● 72



Trajectory Analysis of HSMM Data with Monocle 3

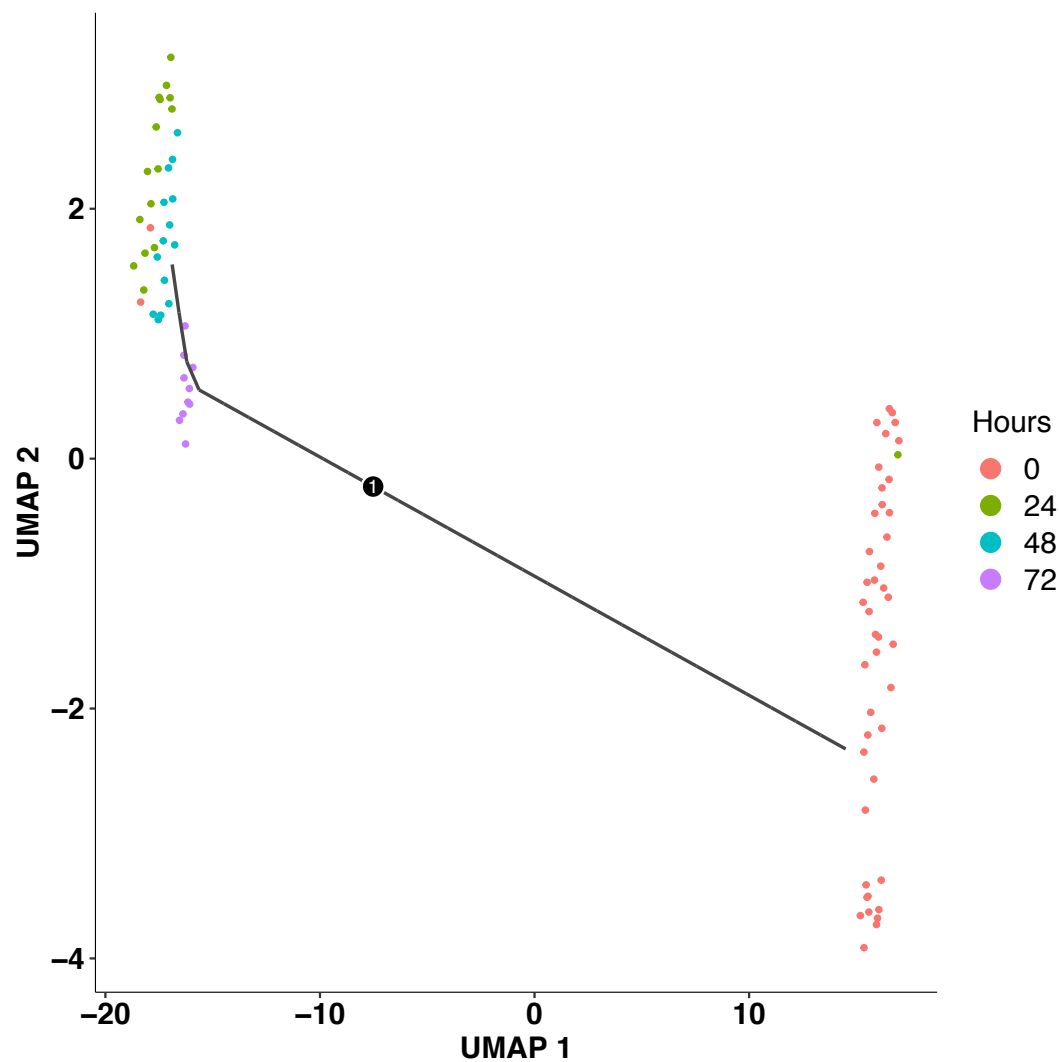
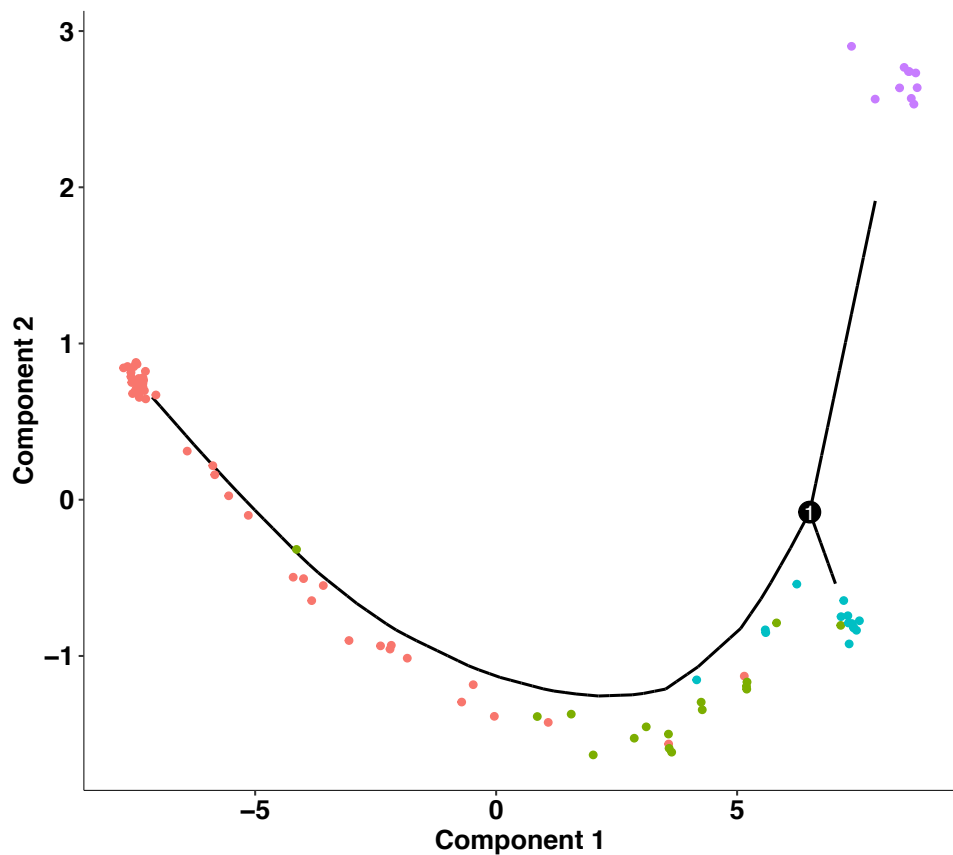
Media	Hours	freq
GM	0	45
DM	24	16
DM	48	14
DM	72	10

DEG: Hours (continuous)
nfeature: 308 genes

Monocle 3

Monocle 2

Hours ● 0 ● 24 ● 48 ● 72



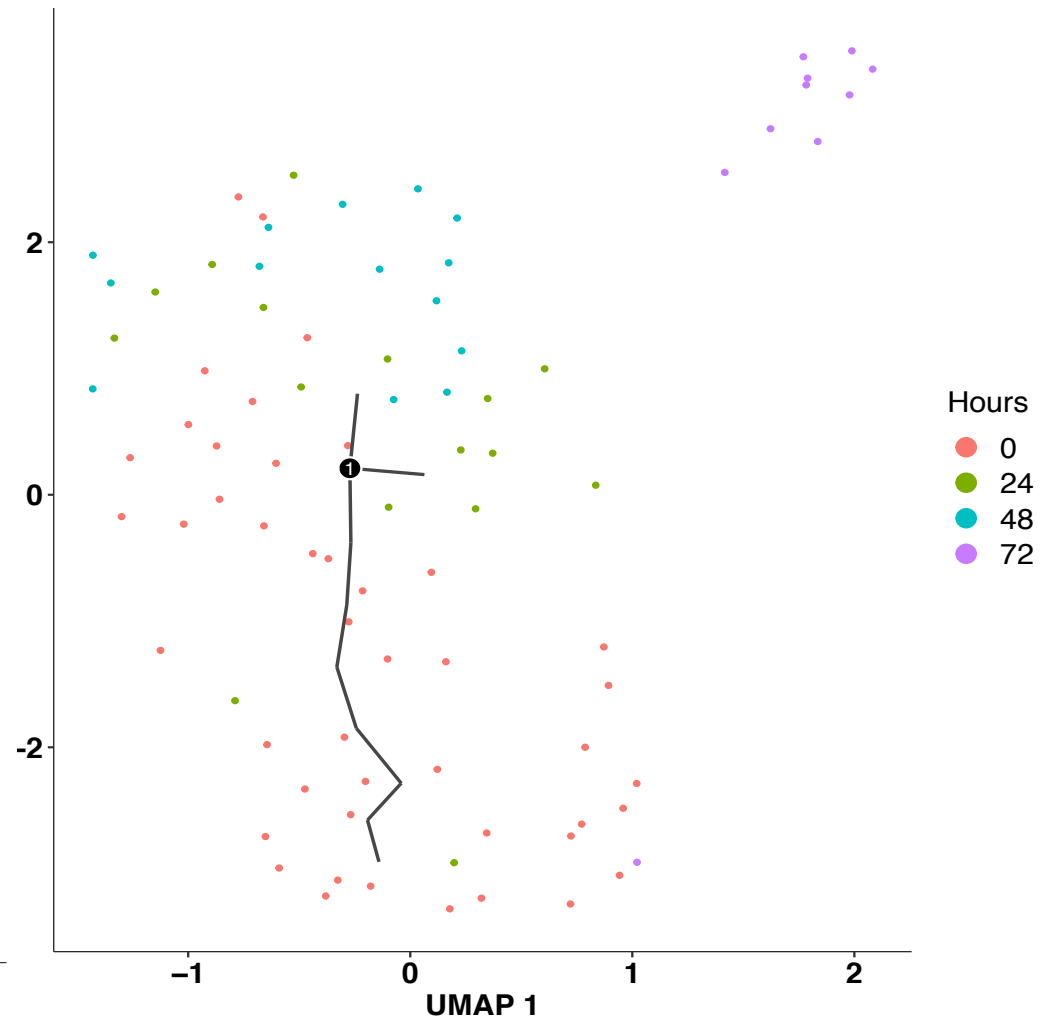
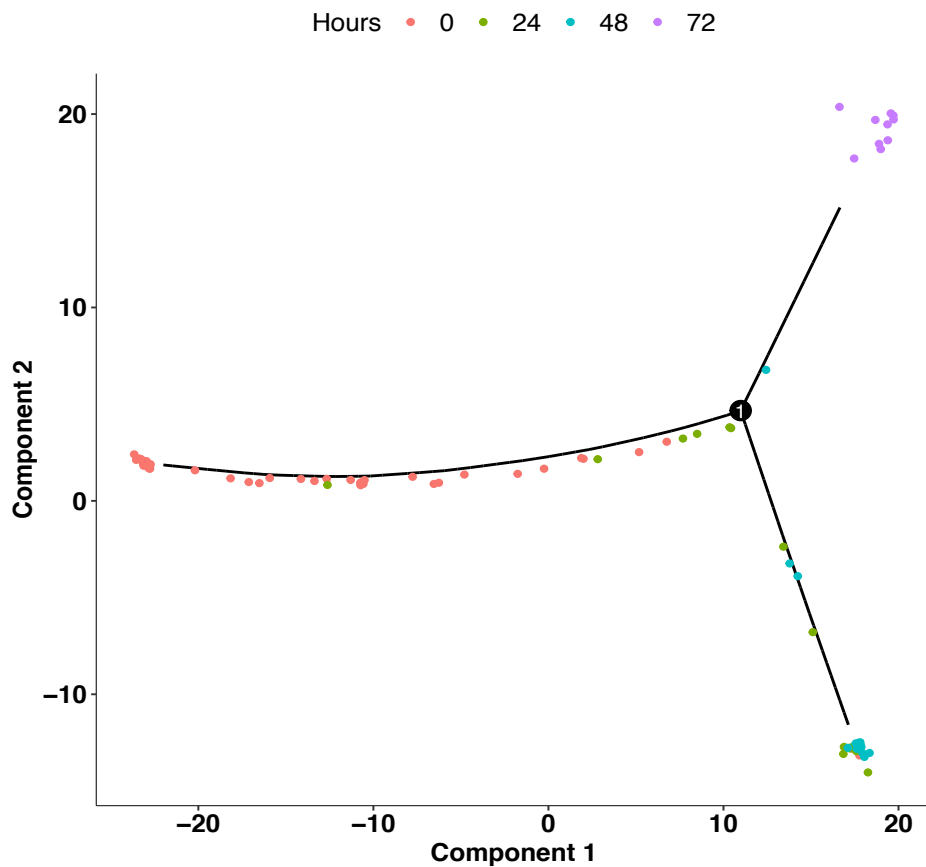
Trajectory Analysis of HSMM Data with Monocle 3

Media	Hours	freq
GM	0	45
DM	24	16
DM	48	14
DM	72	10

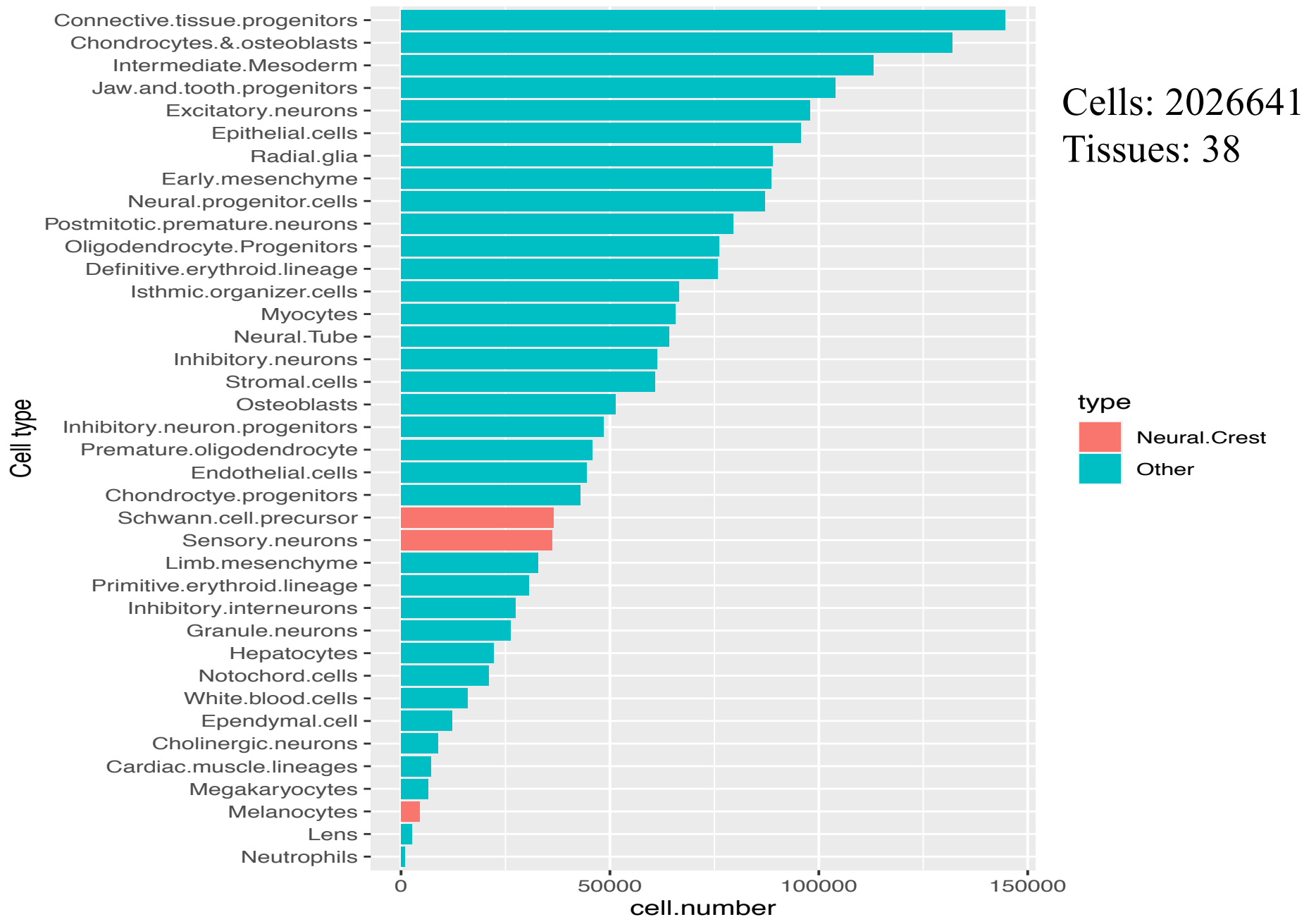
DEG: no filter
nfeature: 14224 genes

Monocle 3

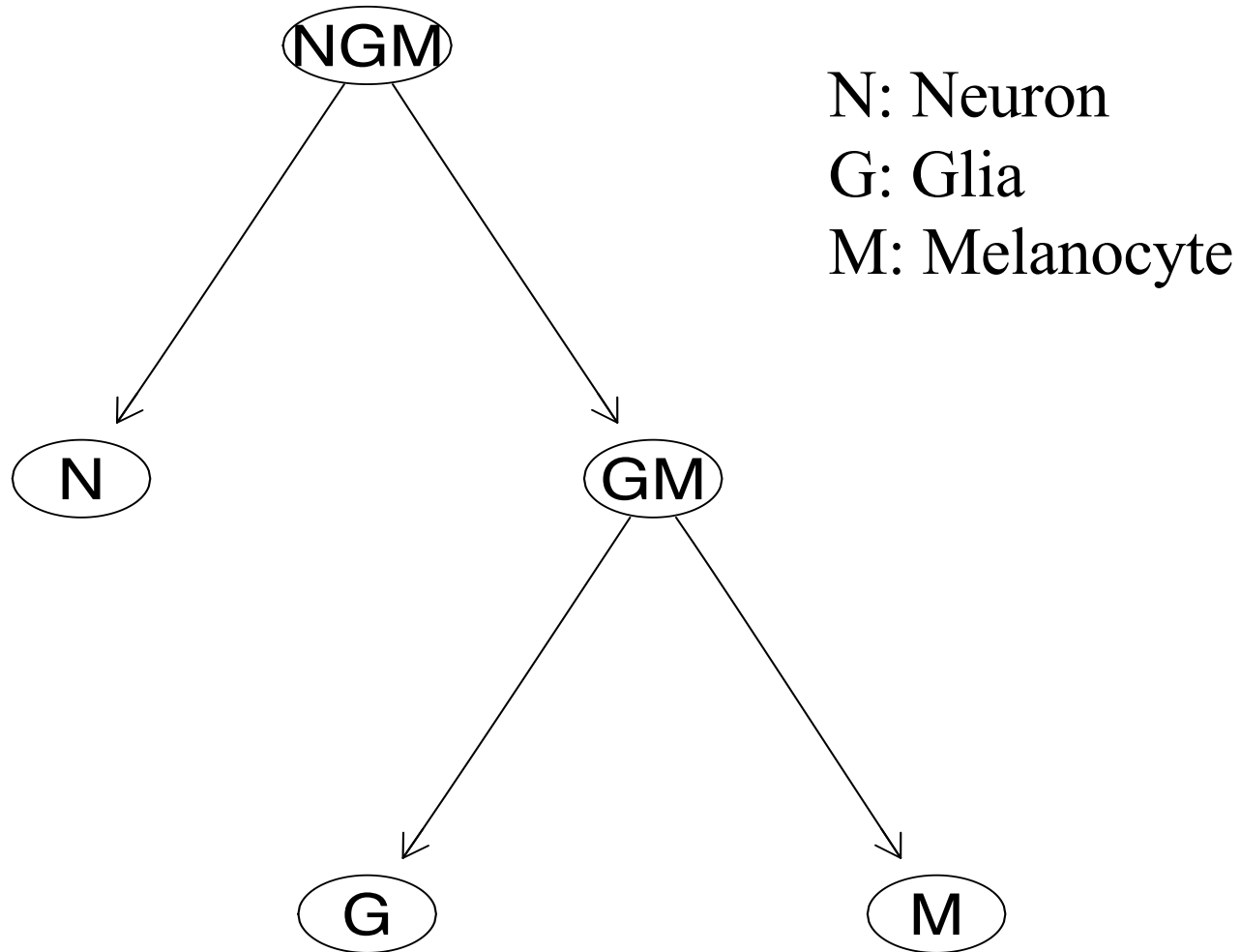
Monocle 2



Mouse Organogenesis Cell Atlas (MOCA) Data



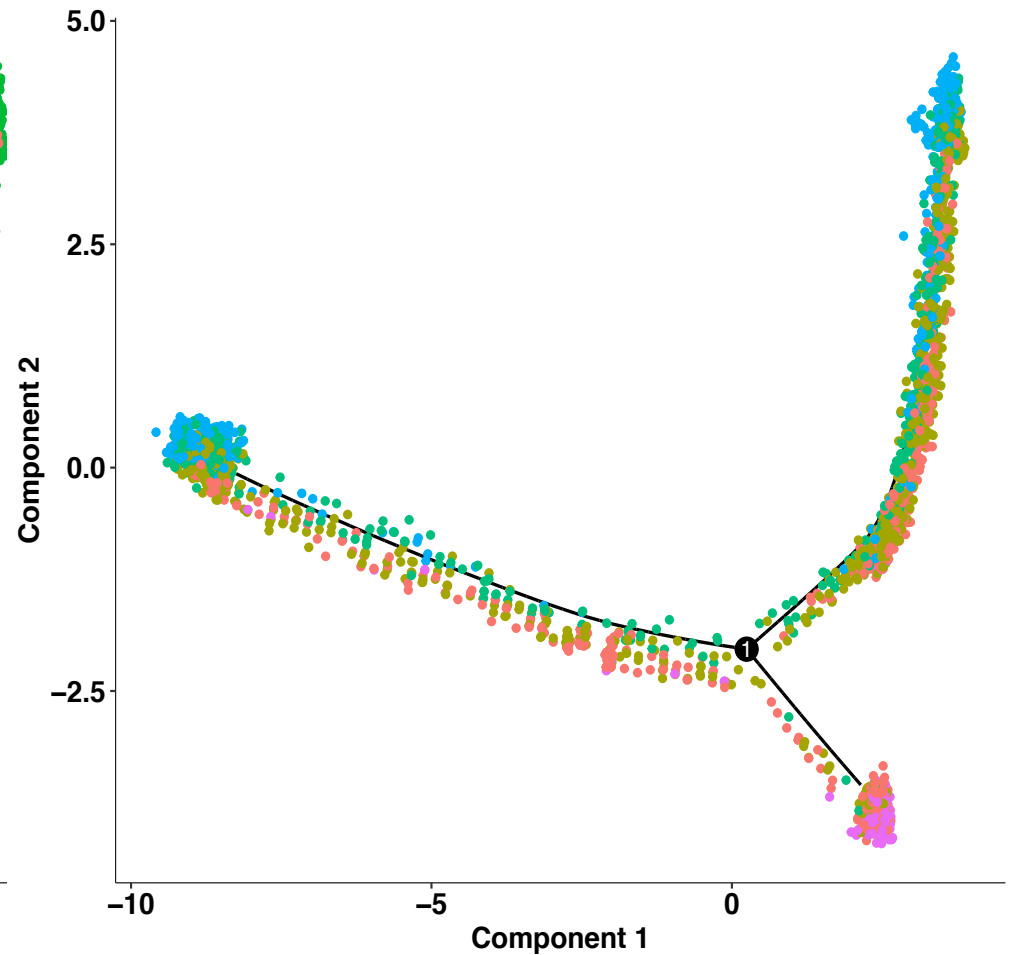
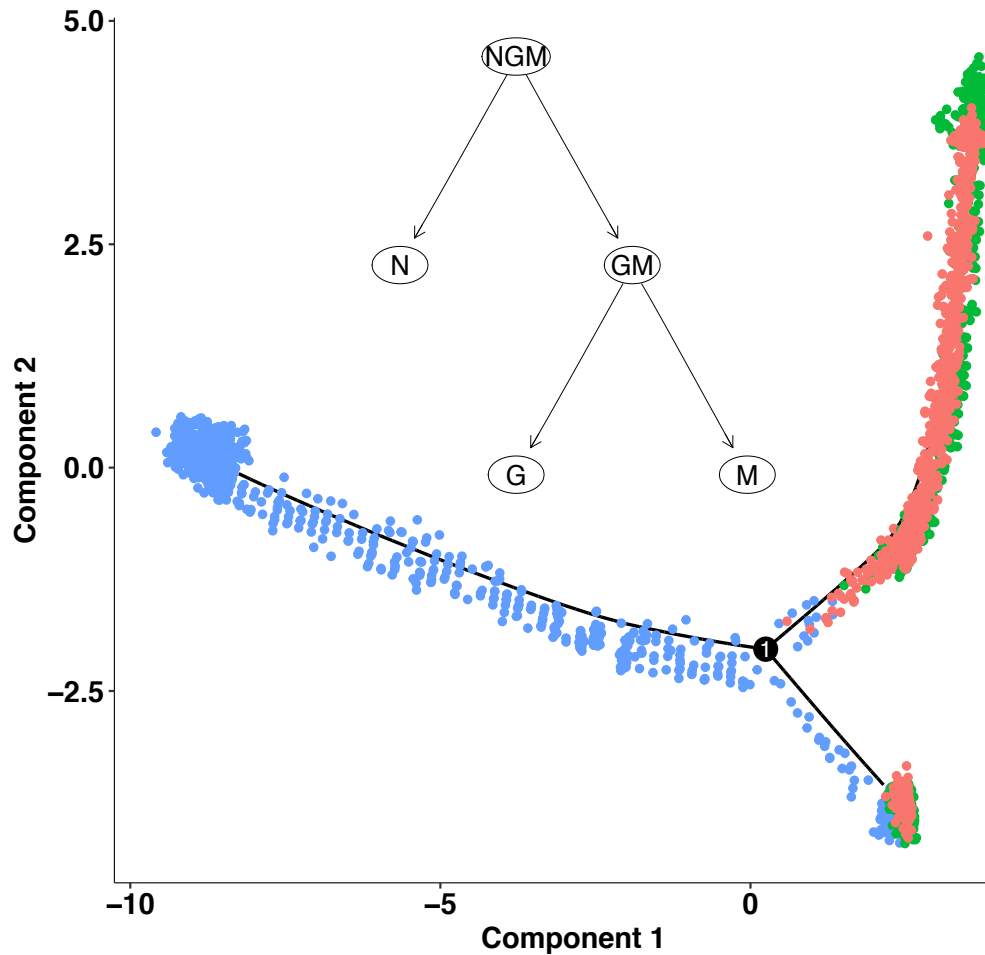
Neural Crest Lineage



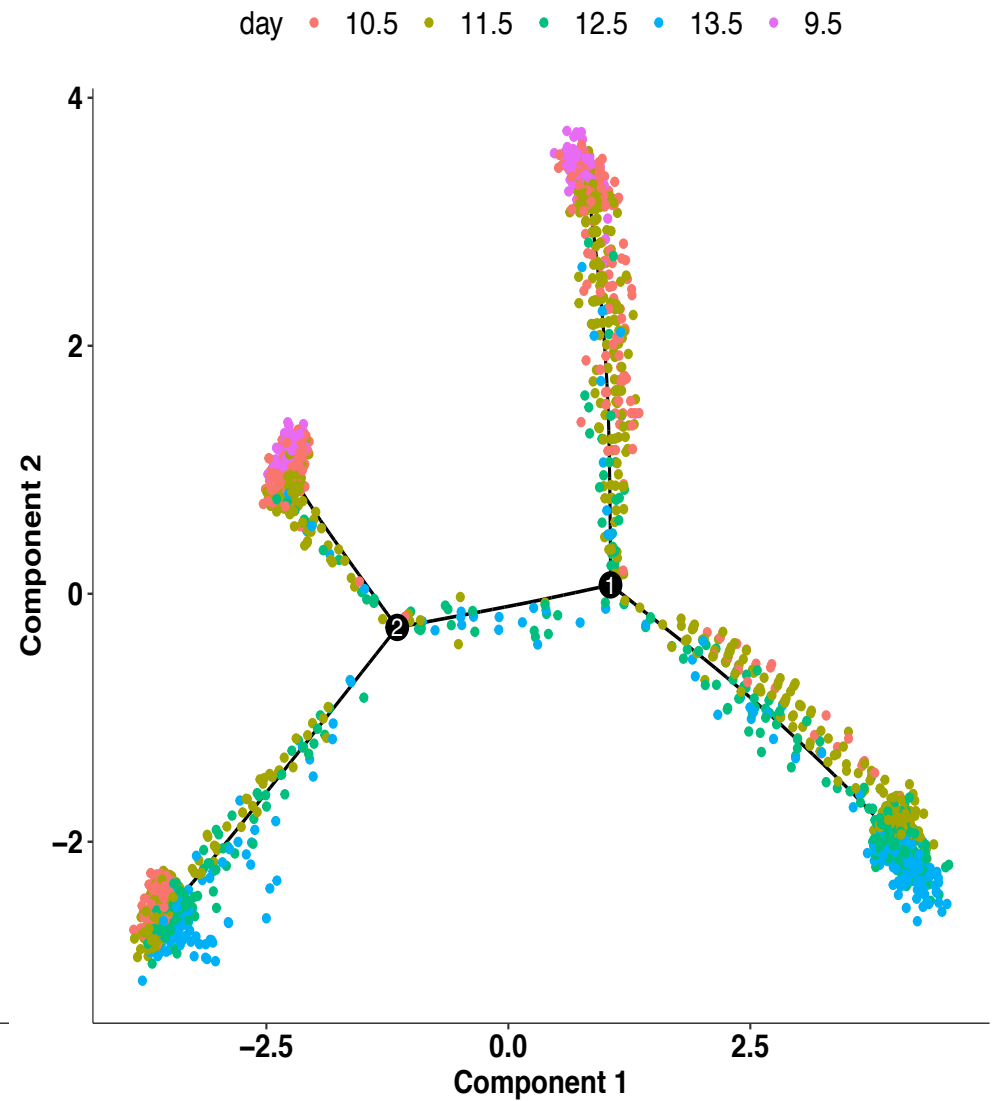
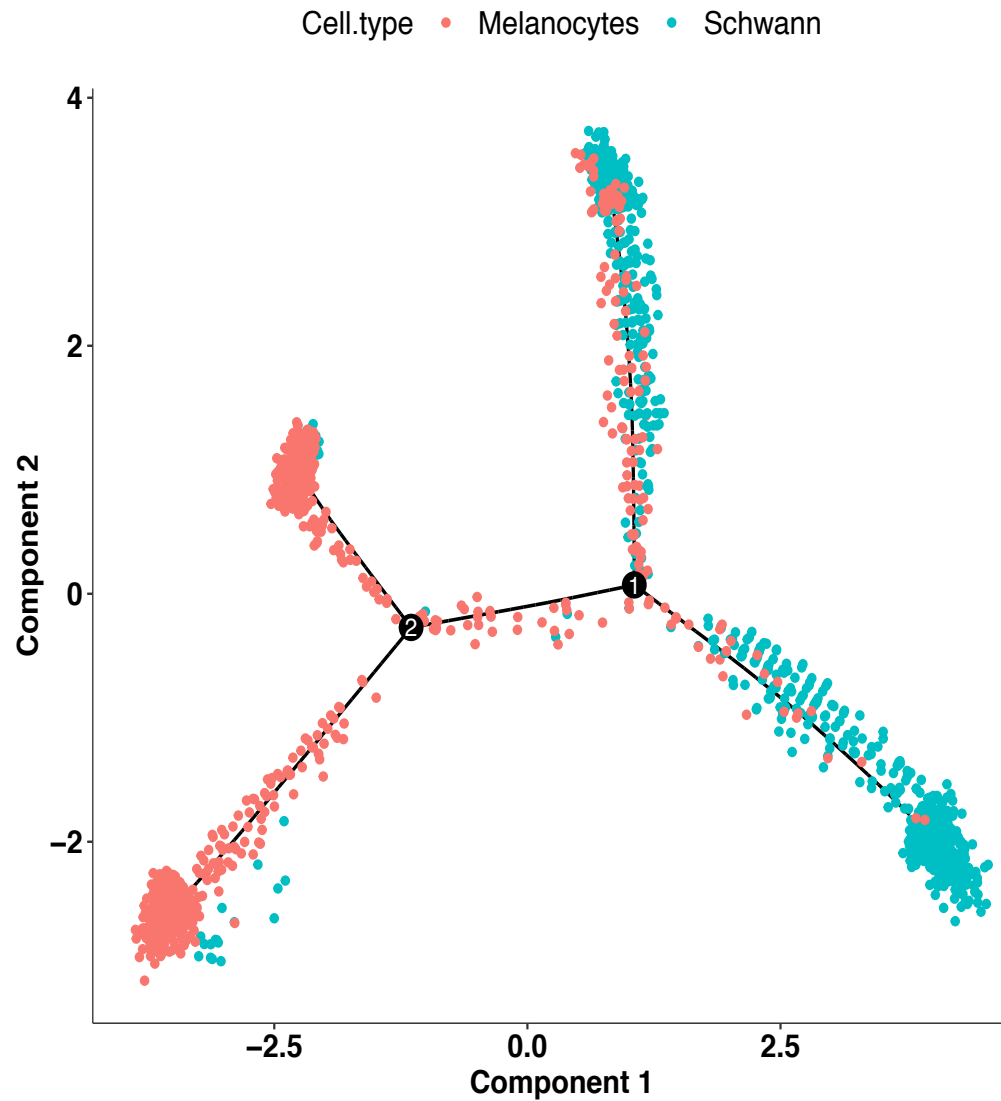
Neural Crest Lineage Analysis with Monocle 2

Cell.type • Melanocytes • Schwann • Sensory

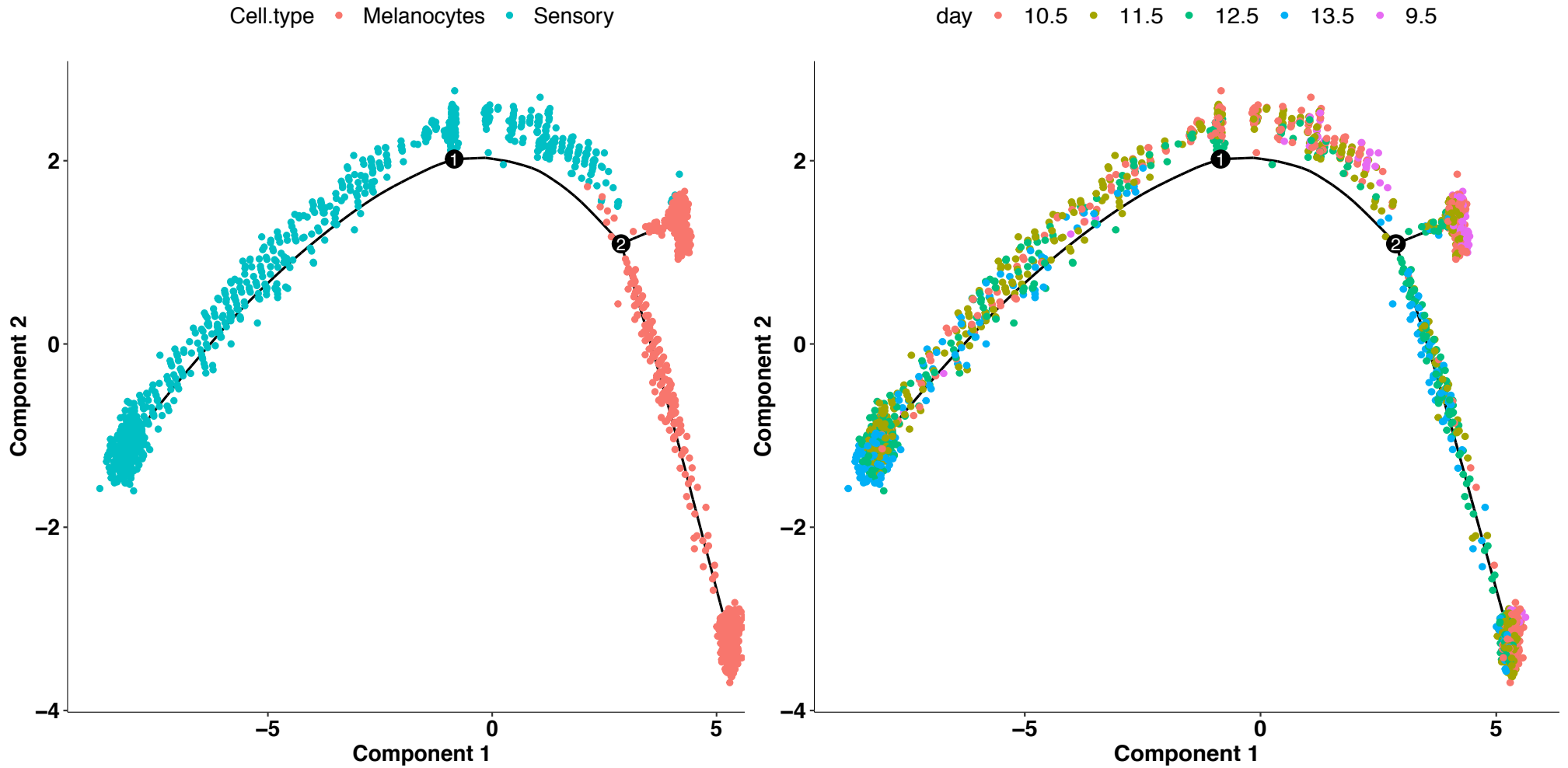
day • 10.5 • 11.5 • 12.5 • 13.5 • 9.5



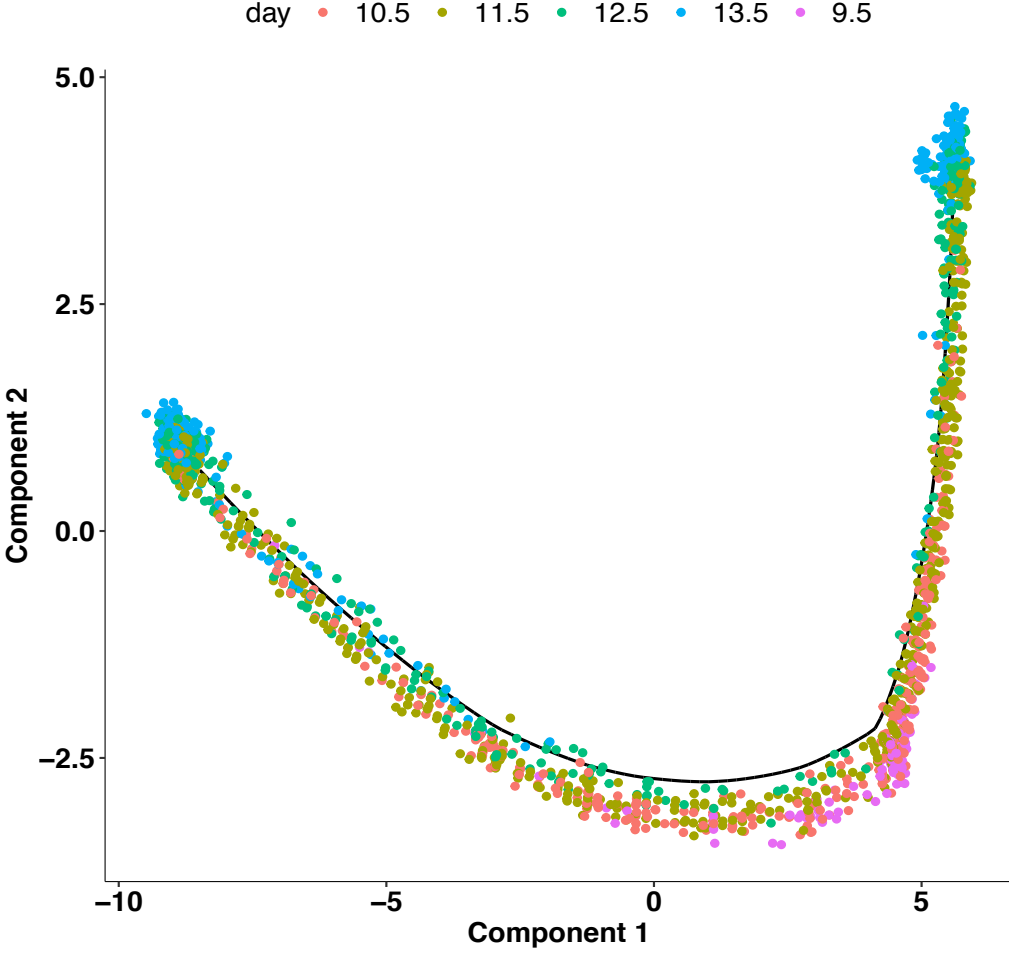
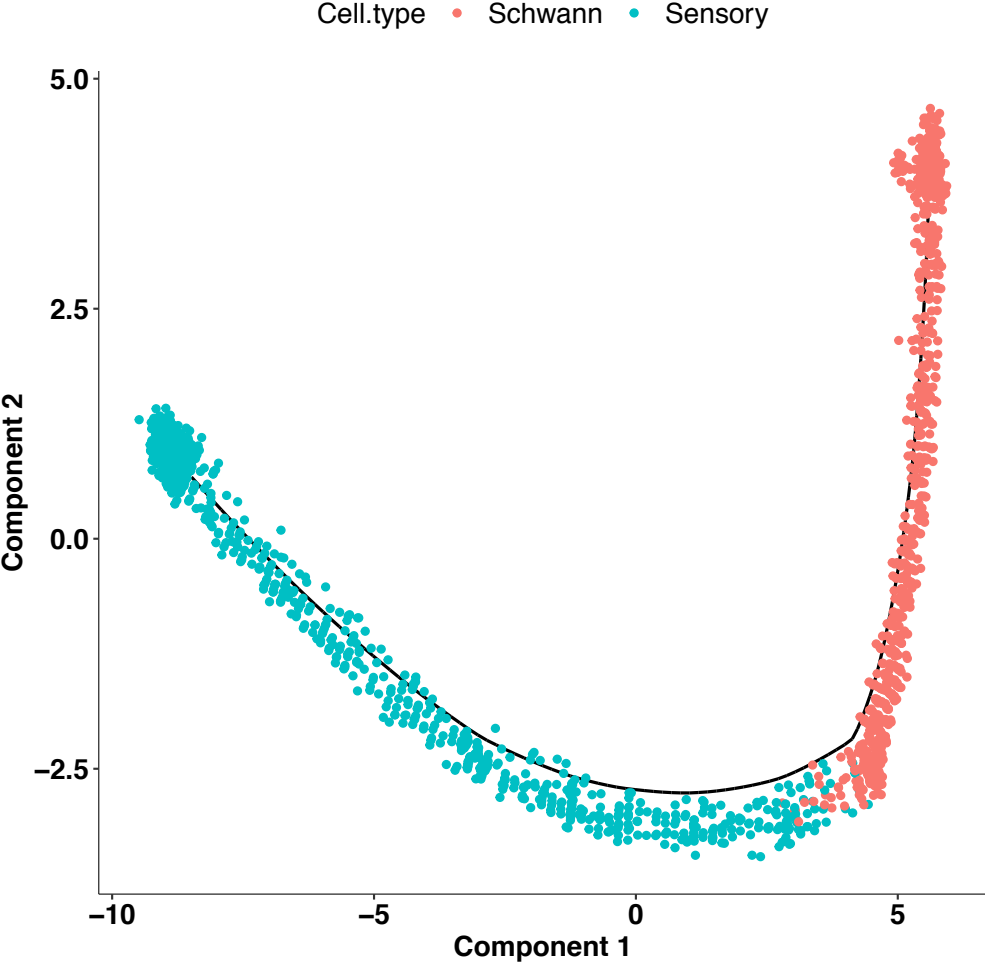
Melanocytes vs Schwann Cells with Monocle 2



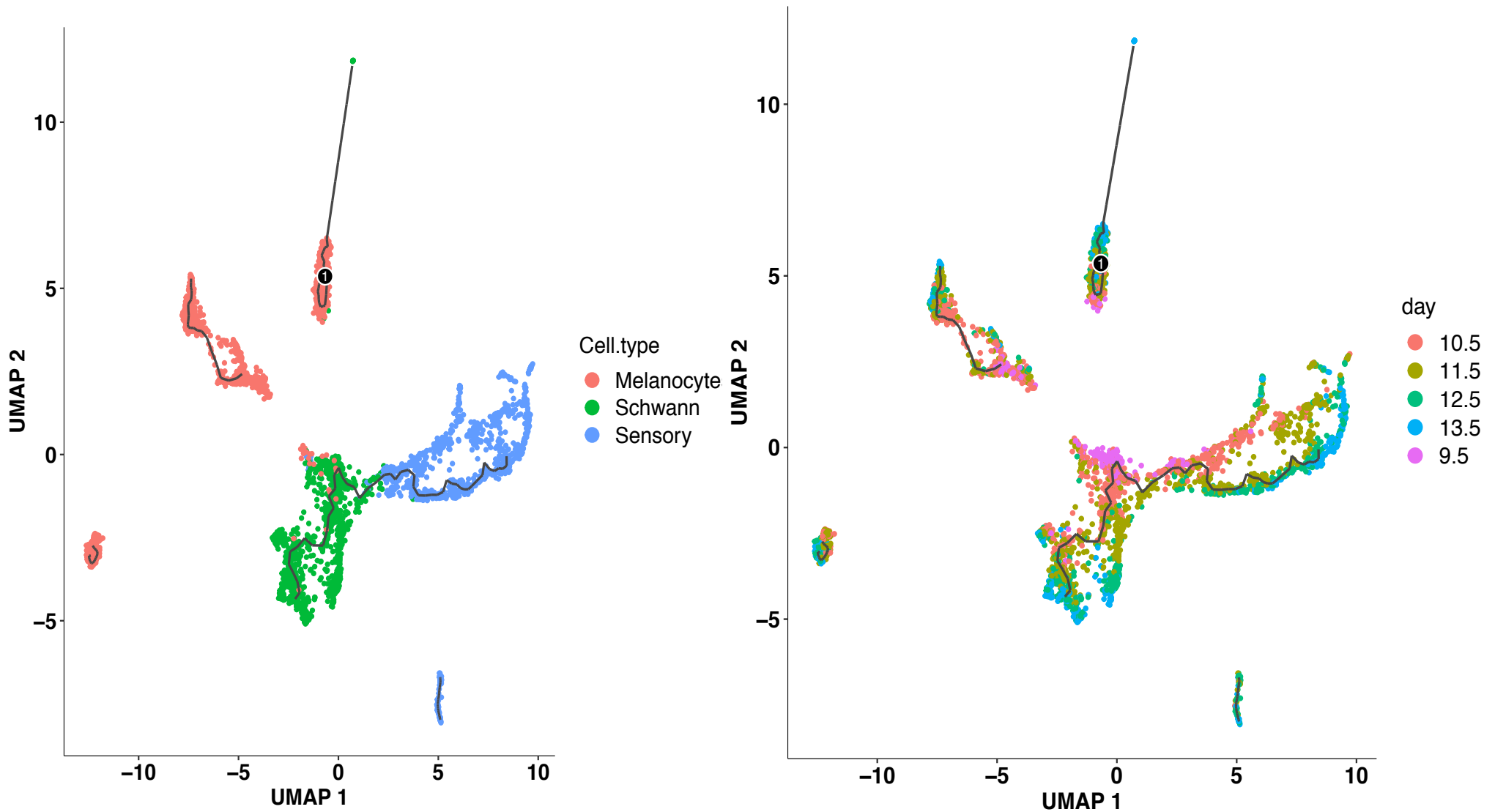
Melanocytes vs Sensory Neurons with Monocle 2



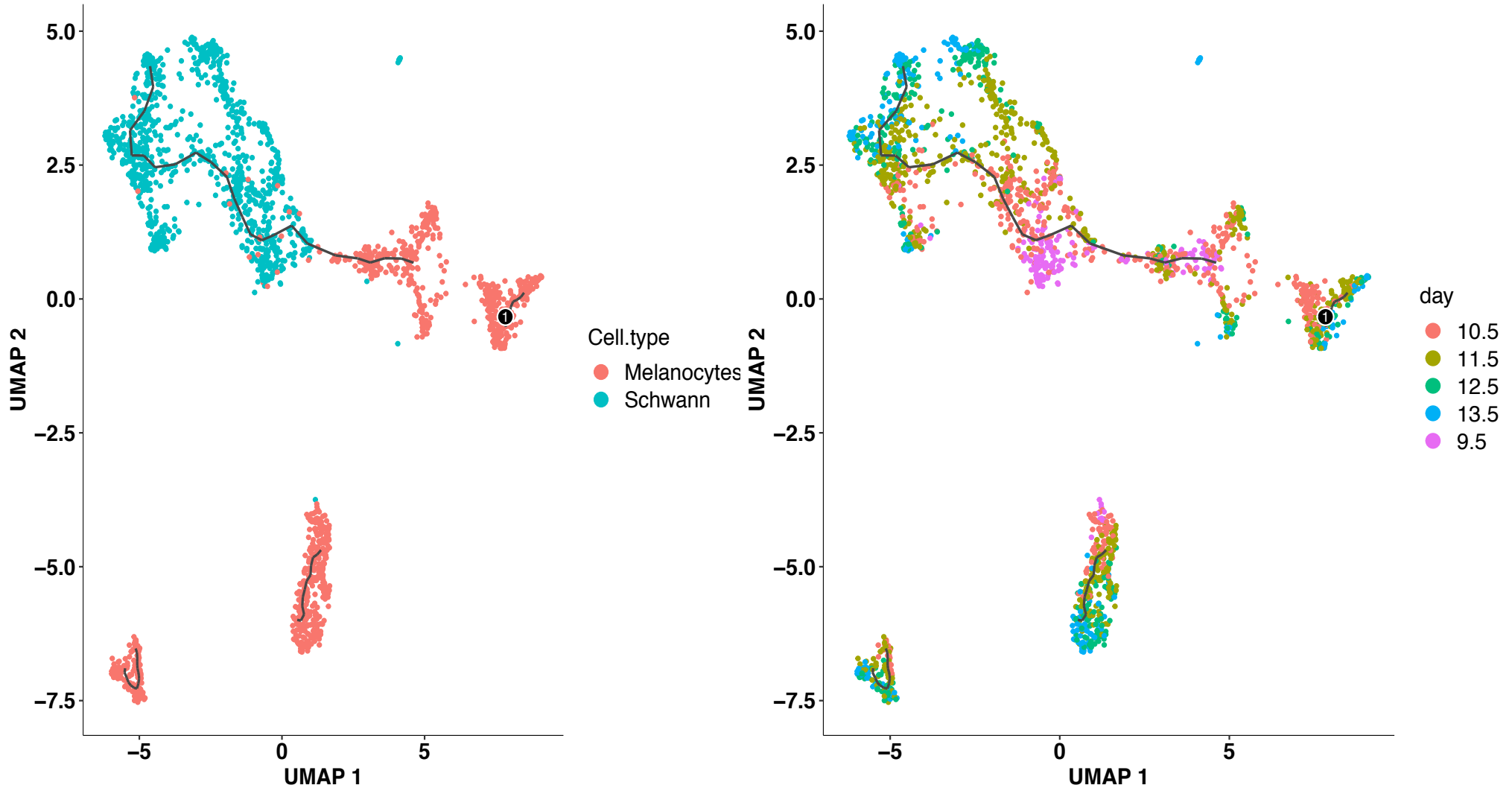
Schwann Cells vs Sensory Neurons with Monocle 2



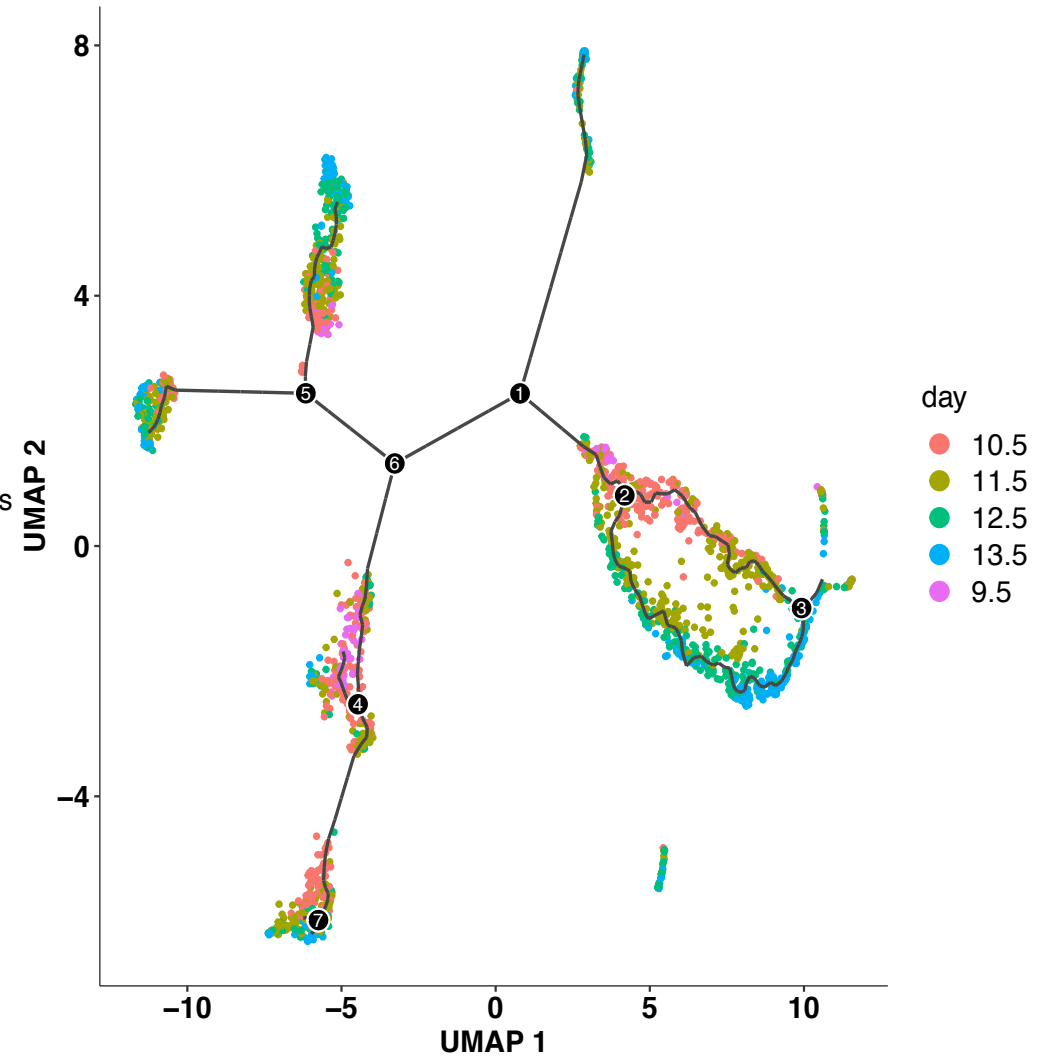
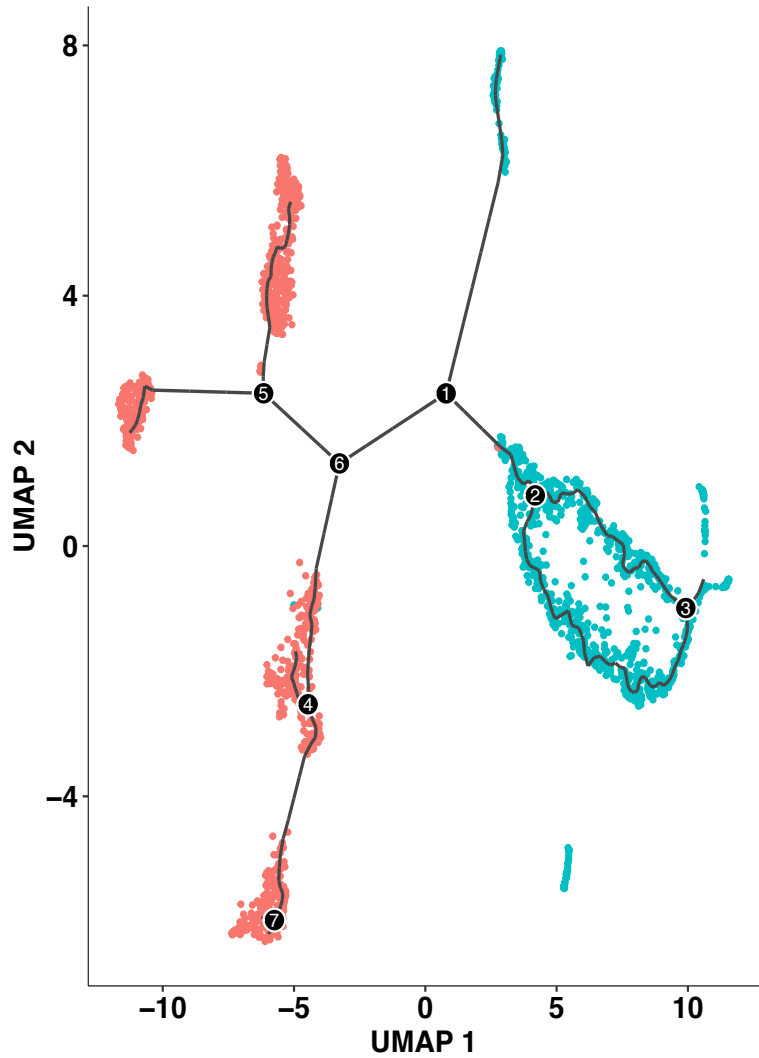
Neural Crest Lineage Analysis with Monocle 3



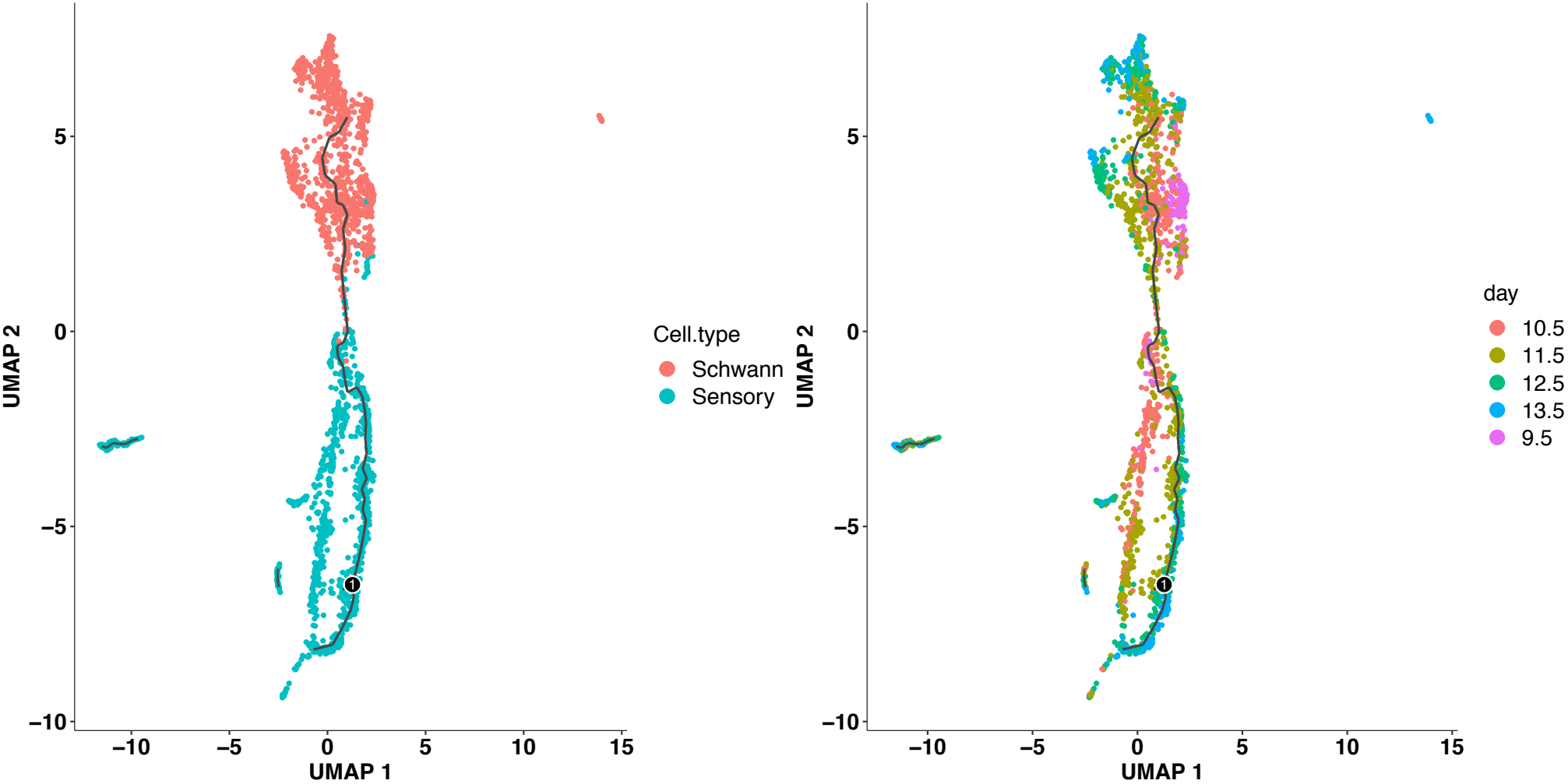
Melanocytes vs Schwann Cells with Monocle 3



Melanocytes vs Sensory Neurons with Monocle 3

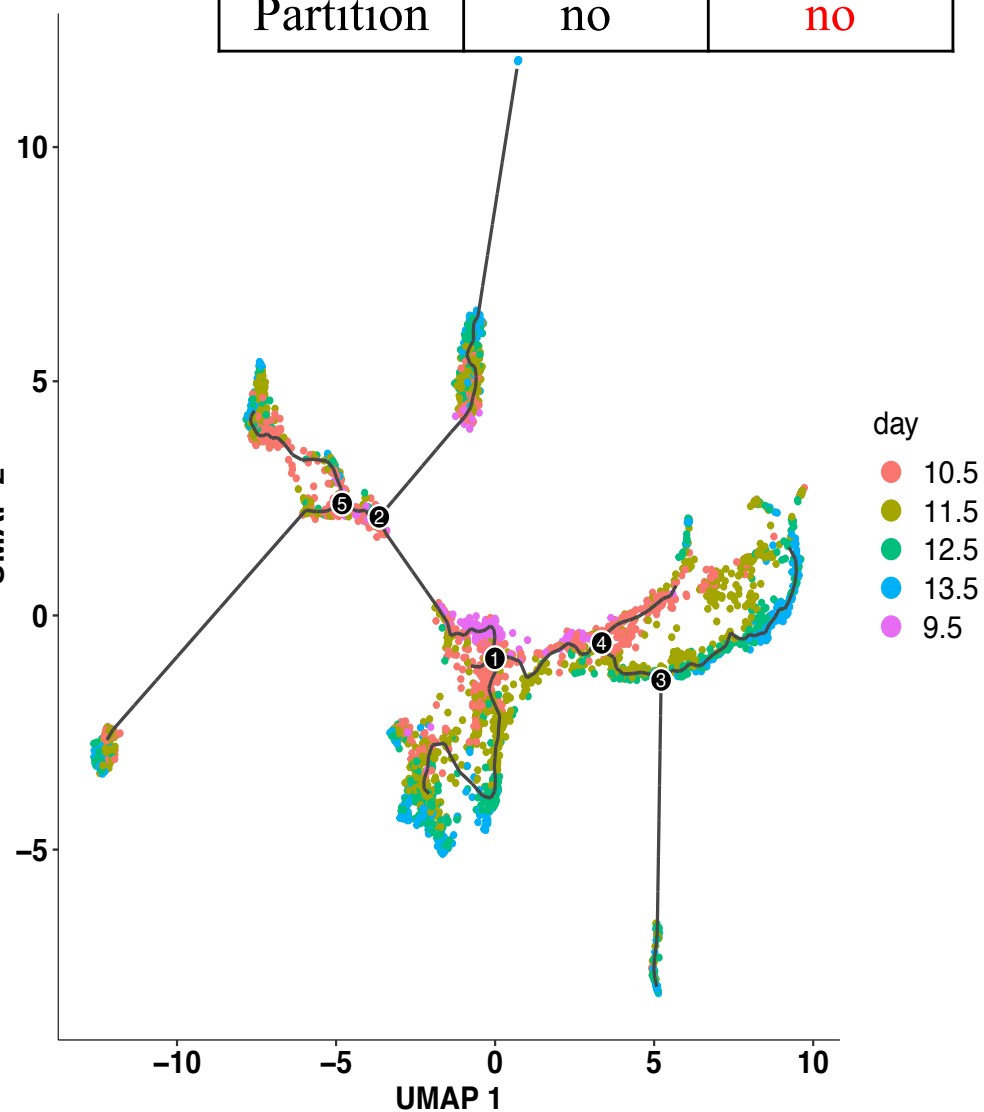
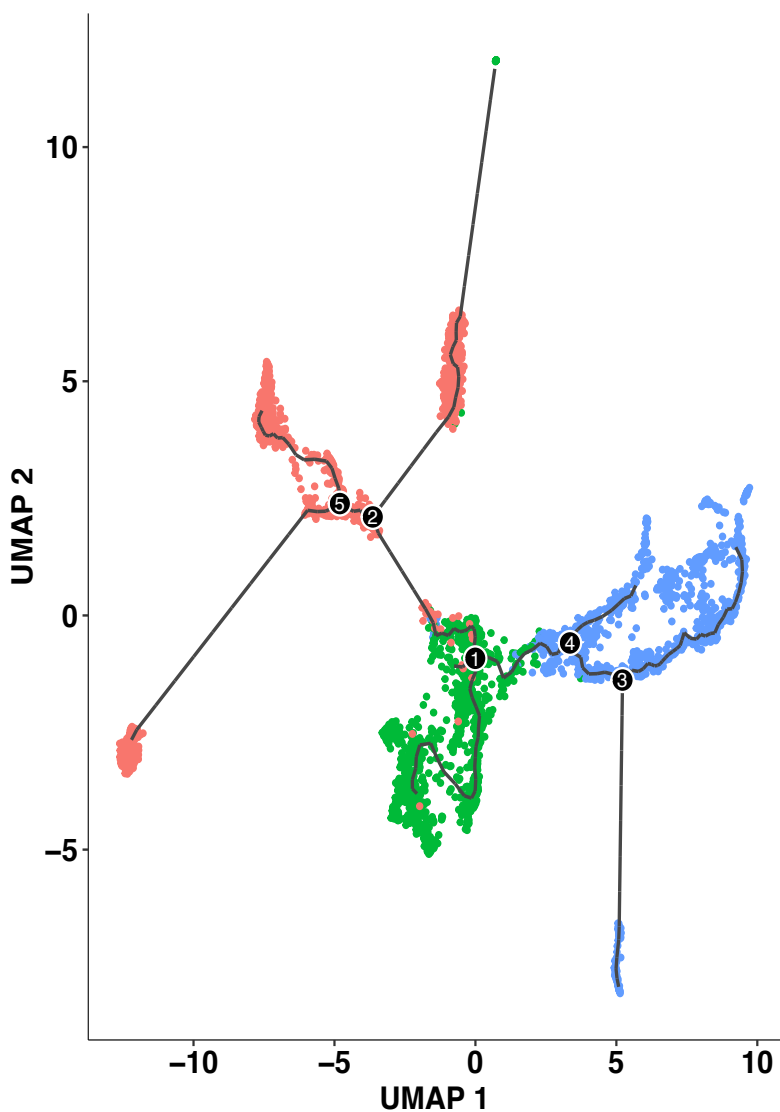


Schwann Cells vs Sensory Neurons with Monocle 3



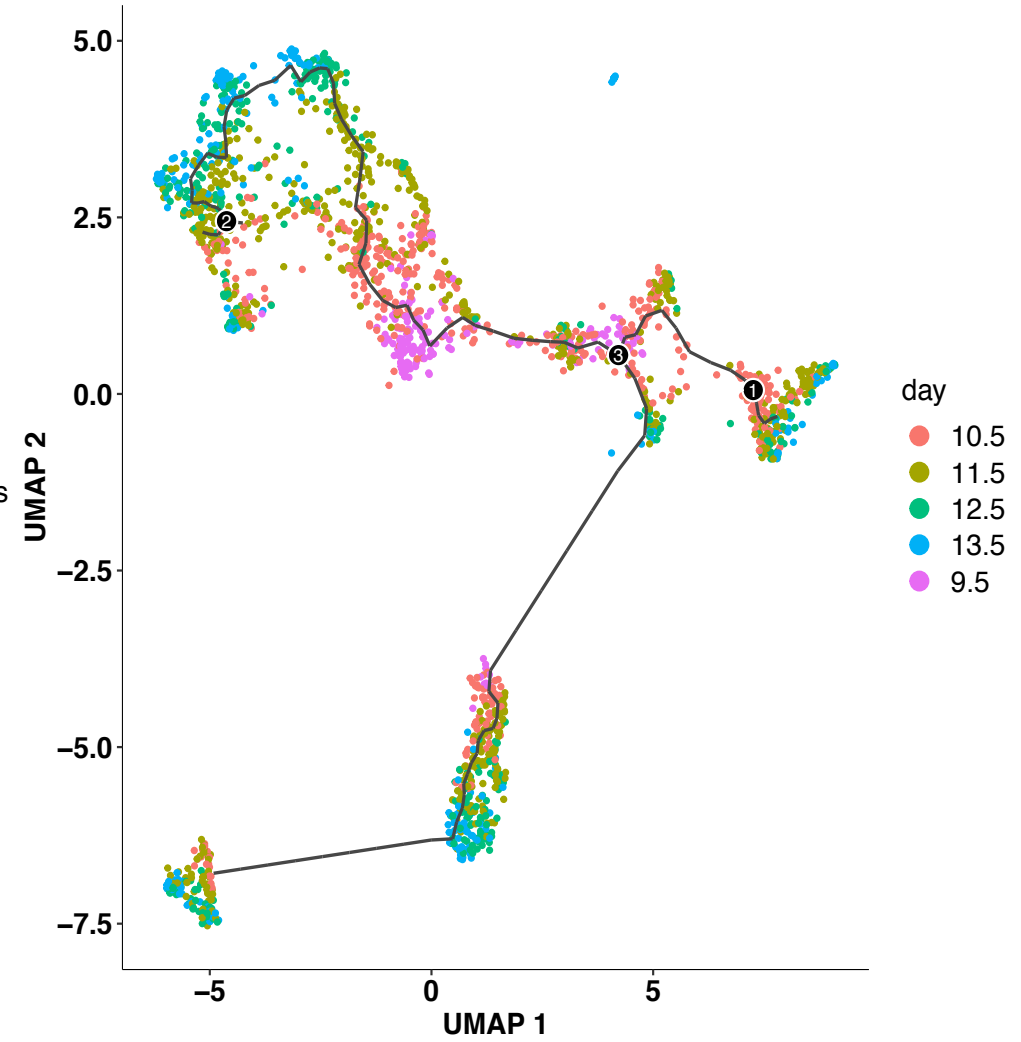
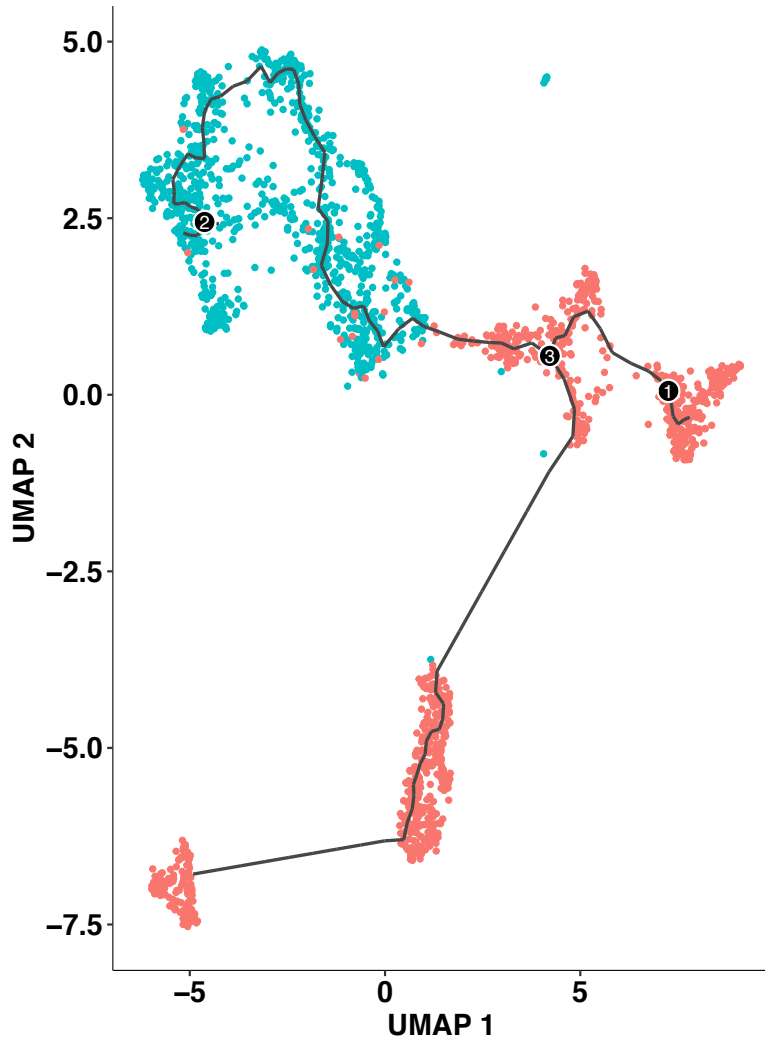
Neural Crest Lineage Analysis with Monocle 3

	monocle2	monocle3
DR	TSNE	UMAP
Graph	Tree	Tree
Partition	no	no



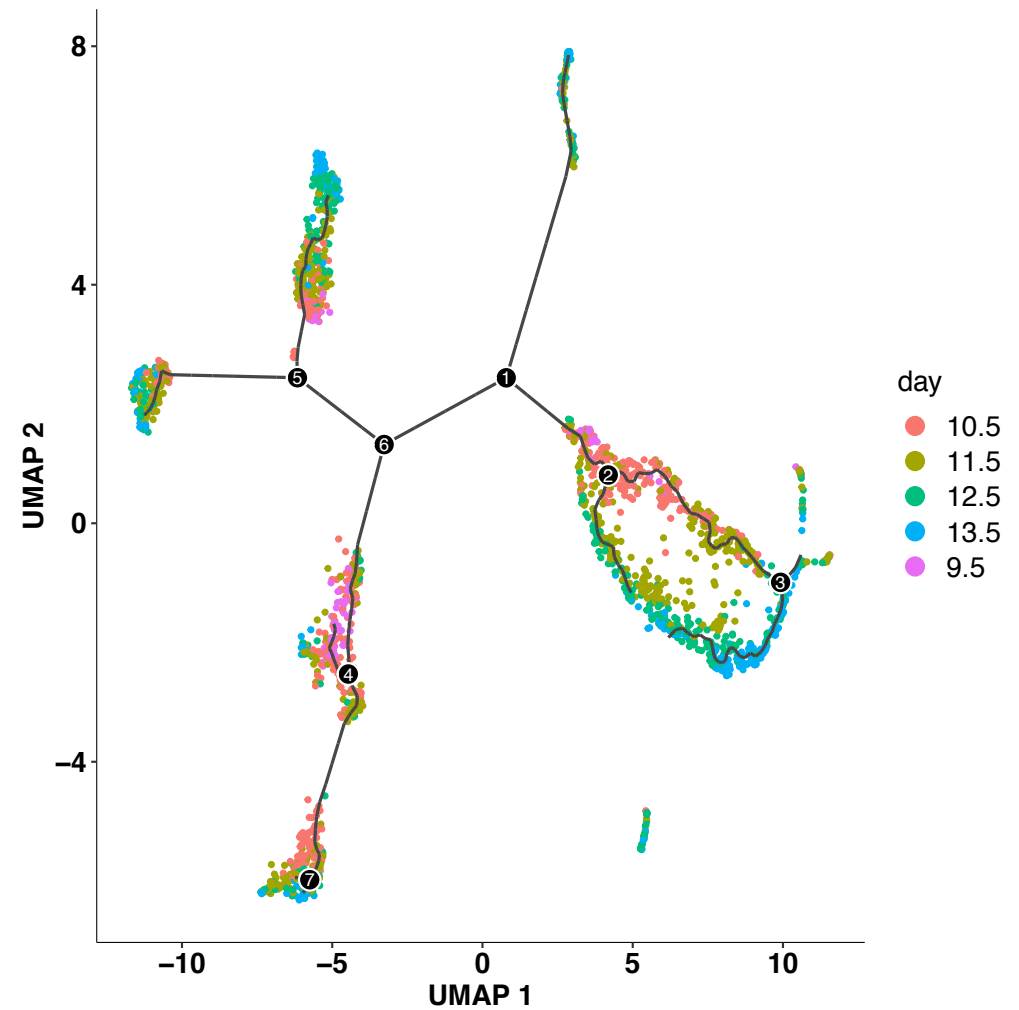
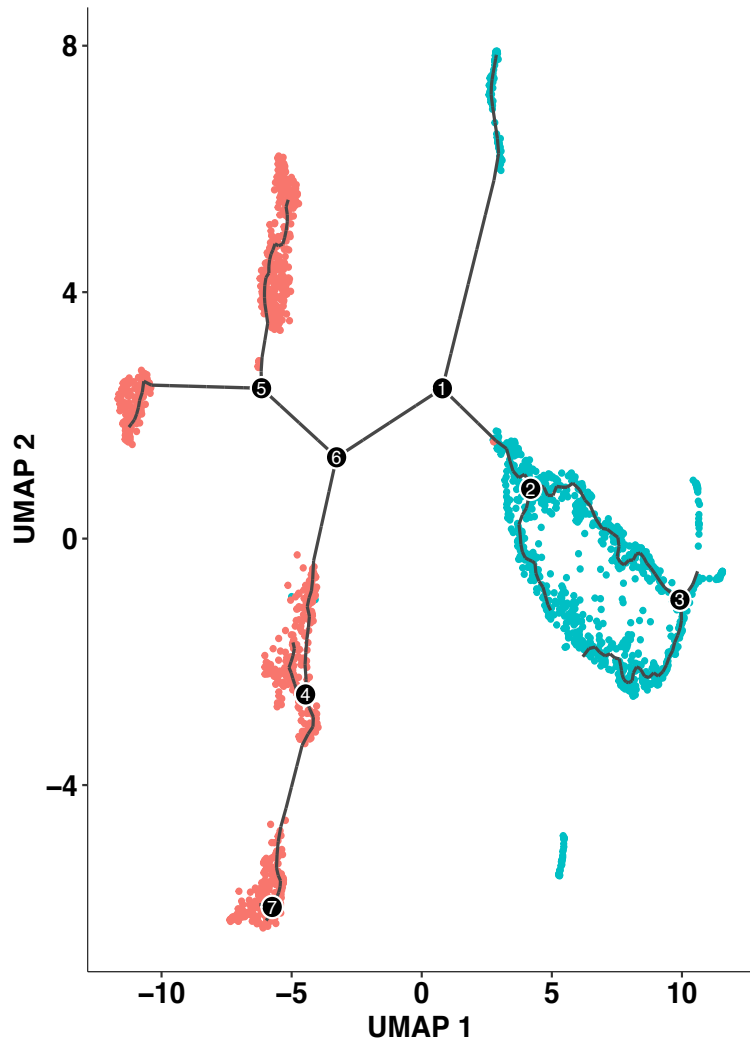
Melanocytes vs Schwann Cells with Monocle 3

	monocle2	monocle3
DR	TSNE	UMAP
Graph	Tree	Tree
Partition	no	no



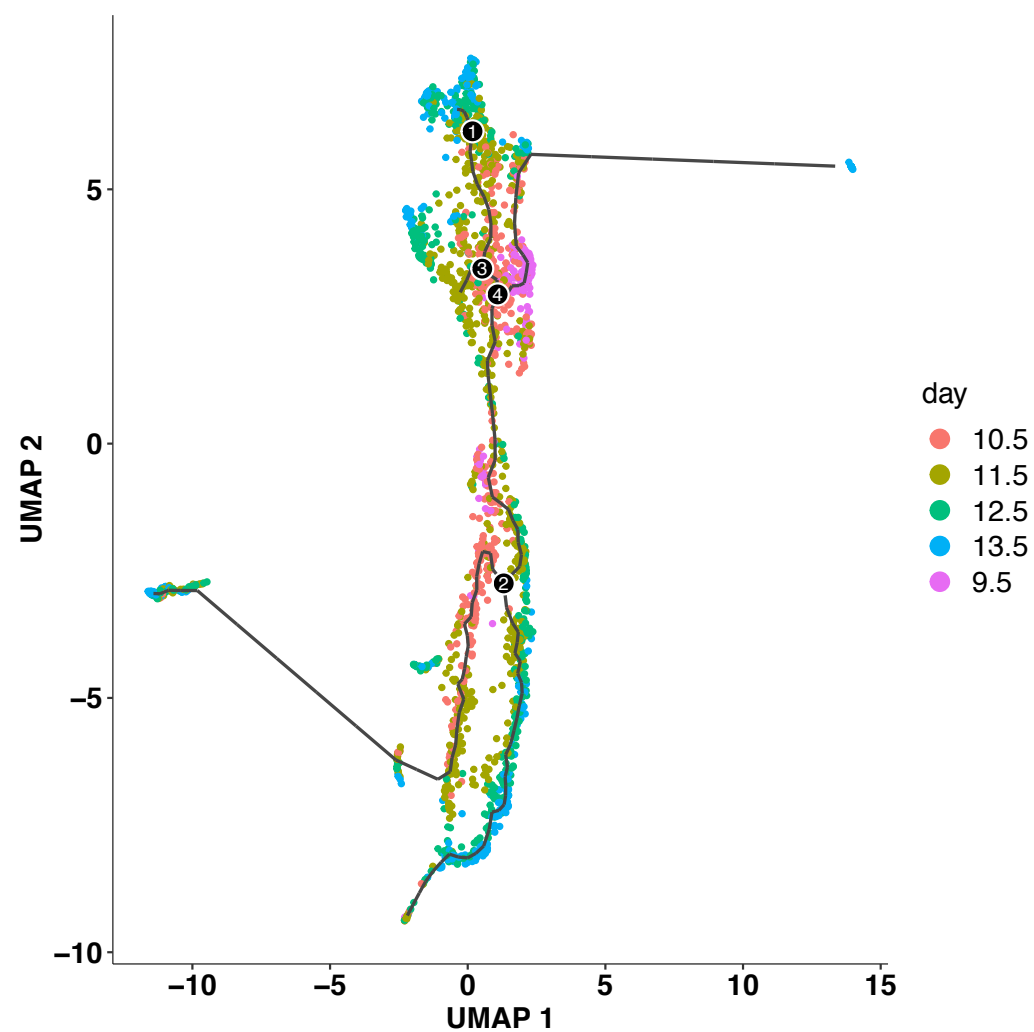
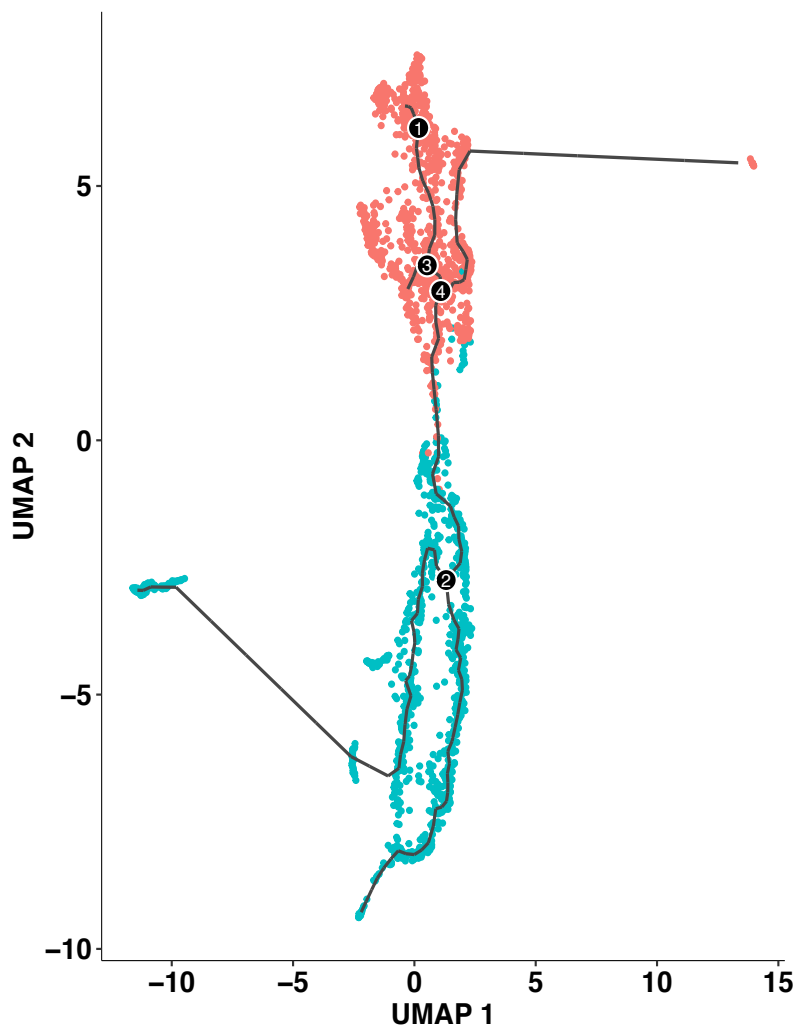
Melanocytes vs Sensory Neurons with Monocle 3

	monocle2	monocle3
DR	TSNE	UMAP
Graph	Tree	Tree
Partition	no	no



Schwann Cells vs Sensory Neurons with Monocle 3

	monocle2	monocle3
DR	TSNE	UMAP
Graph	Tree	Tree
Partition	no	no



Road Map for Dimension Reduction Methods

