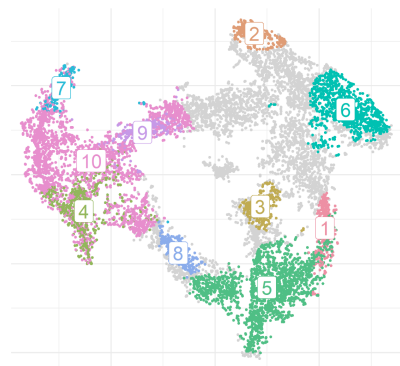


A machine-learning guided approach to explore the *cis*-regulatory code involved in neuronal differentiation

Océane Cassan^{1*}, Christophe Vroland^{1,2}, Julien Raynal^{1,2}, FANTOM consortium,
Masaki Kato³, Hazuki Takahashi³, Takeya Kasukawa³,
Piero Carninci³, Chi Wai Yip^{3*}, Laurent Bréhélin^{1*} & Charles-Henri Lecellier^{1,2*}

Poster B-330



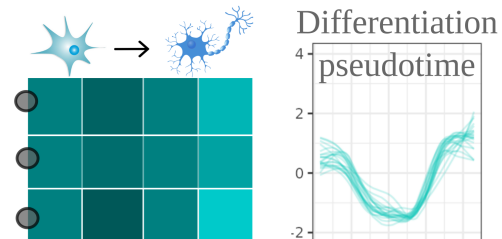
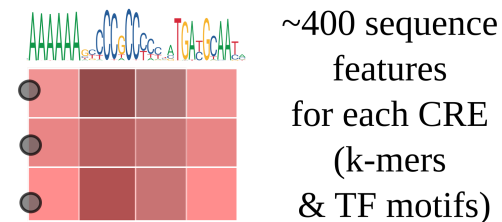
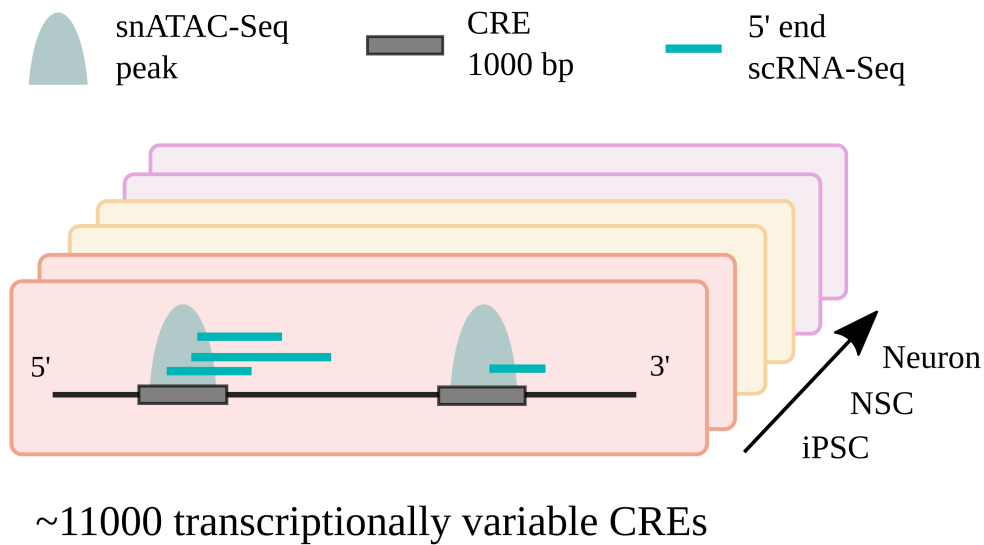
July 14, 2024
12:20 - 14:20



Context

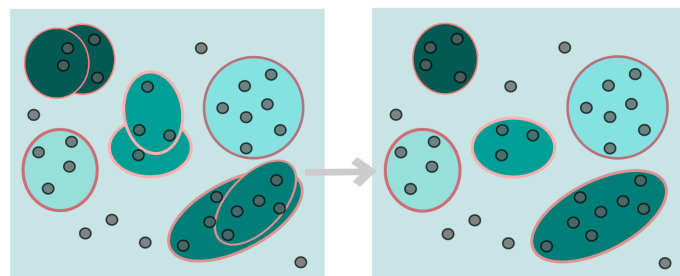
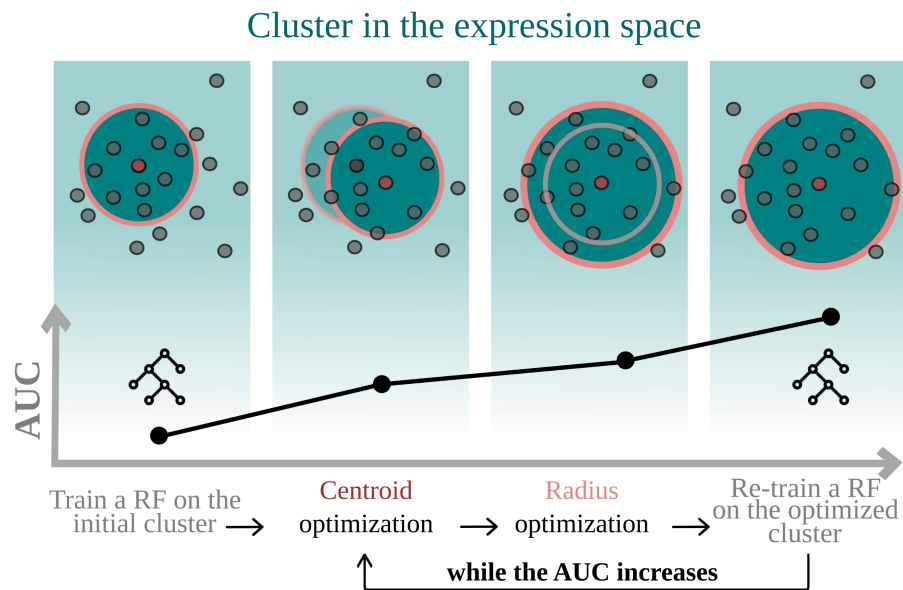
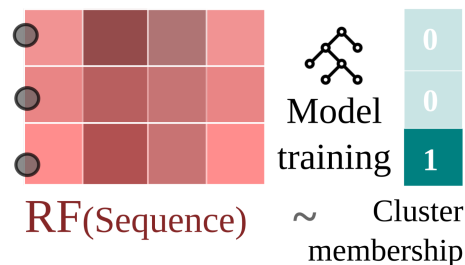
What are the sequence features underlying the coordinated changes in transcriptional activity of cis-regulatory elements (CREs) during differentiation processes?

Single cell dataset of human neuronal differentiation

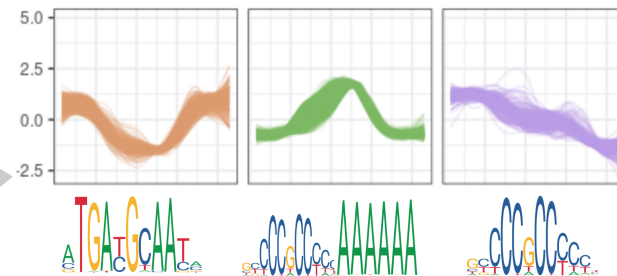


Statistical learning TO Inform Clustering

Stoic guides **co-expression clustering** with sequence information in a **supervised** framework



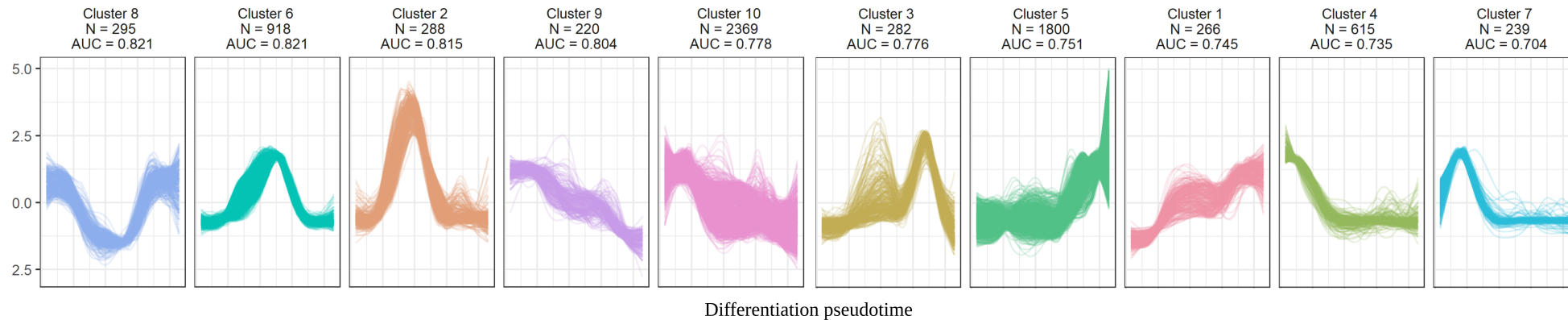
Redundant clusters pruning



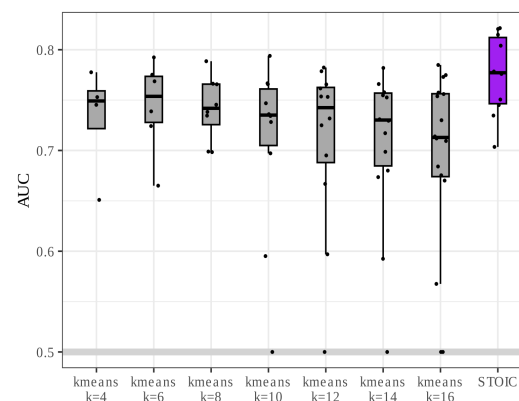
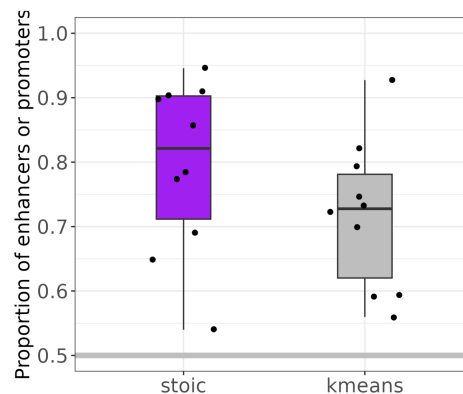
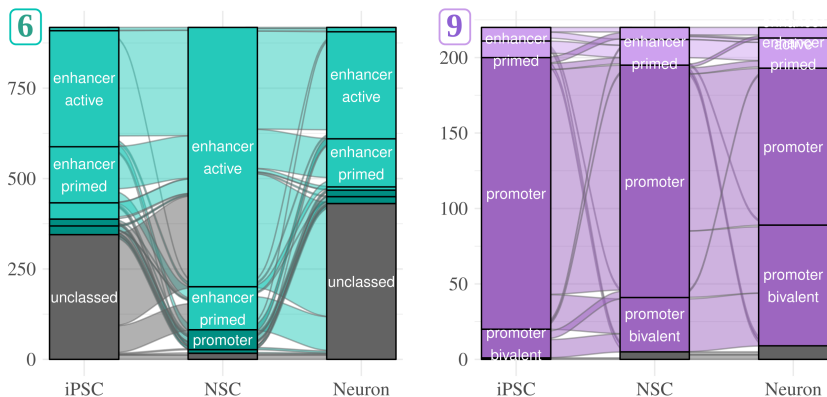
Expression clusters associated with important sequence features



Stoic clusters outline diverse expression profiles and epigenetic signatures



Most inferred clusters are either **enhancer-rich** or **promoter-rich**



Stoic provides **higher AUC** values for discriminating cluster membership using sequence features



Stoic R package

The machine-learning guided approach developed in Stoic is **applicable to any problem where the clustering of some measurements can be guided by a second matched dataset.**

Let's discuss around Poster B-330

oceane.cassan@lirmm.fr

July 14, 2024 12:20 - 14:20

