

SCIENCE AT THE EDGE

2016 SEMINAR SERIES

Quantitative Biology Graduate Program | Gene Expression in Development and Disease

Erwin Frise

LAWRENCE BERKELEY NATIONAL LABORATORIES
DEPARTMENT OF ENVIRONMENTAL GENOMICS AND
SYSTEMS BIOLOGY

“Mastery of Spatial Gene Expression and Other Convolved Datasets”

Extracting information from large and complex datasets is a common problem in quantitative biology. With the goal of analyzing spatial gene expression images systematically collected during *Drosophila* embryogenesis, we have been developing novel representations and analysis strategies. Our most recent advance is an algorithm that is able to separate the composite data into basic elements, which we call principal patterns. Like other dimensionality reduction methods (e.g. PCA), principal patterns allow for a compact and interpretable representation but, more importantly, also represent the underlying biology. In case of the *Drosophila* embryo, principal patterns correspond to the regions for organ development. Using principal patterns, we developed strategies to associate genes to organs, follow their roles and investigate spatially defined regulatory interactions. More generally, principal patterns are a modern PCA, able to provide an informative insight into large complex biological datasets.

References:

Frise E, Hammonds AS, Celniker SE (2010) Systematic image-driven analysis of the spatial *Drosophila* embryonic expression landscape. *Mol Syst Biol* 6:345.

Hammonds AS, et al. (2013) Spatial expression of transcription factors in *Drosophila* embryonic organ development. *Genome Biol* 14(12):R140.

Wu S, et al (2016) Stability-driven nonnegative matrix factorization to interpret spatial gene expression and build local gene networks. *PNAS* 113:4290

FRIDAY, OCTOBER 21, 2016

11:30AM, ROOM 1400 BPS

Refreshments at 11:15