Spatial and temporal transcriptional landscape during fly development

Cecilia Coimbra Klein, Marina Ruiz-Romero, Silvia Perez-Lluch, Alessandra Breschi, Amaya Abad, Emilio Palumbo and Roderic Guigó

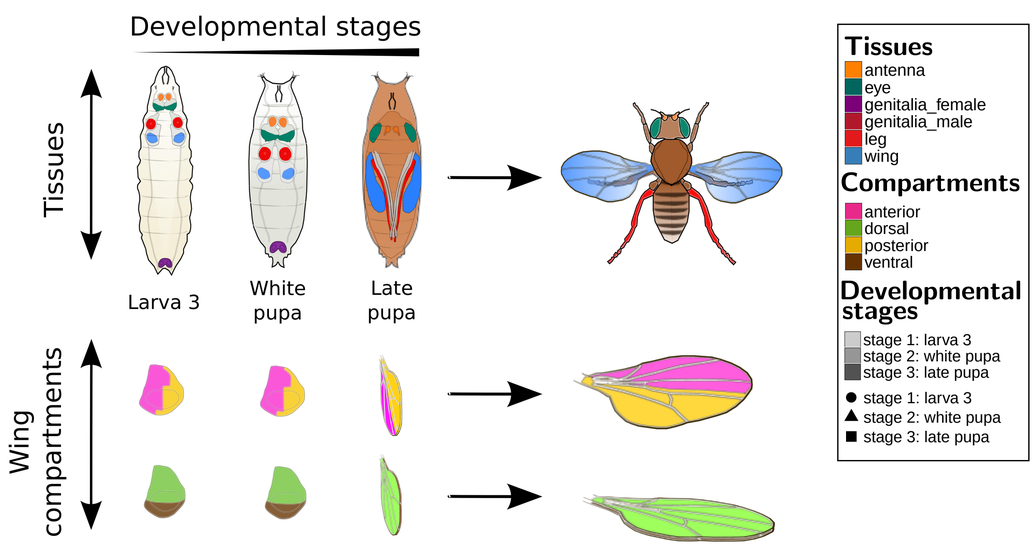
Centre for Genomic Regulation (CRG) and UPF, Barcelona, Spain
Institut Hospital del Mar d'Investigacions Mediques (IMIM), Barcelona 08003, Spain

Abstract

During development most tissues undergo fast changes in order to develop into a functional organs. In this scenario regulation of gene expression turns to be essential to determine cell and tissue specificity. Although many studies aim to decipher tissue signature through the analysis of their transcriptome profiles most lack temporal information during tissue differentiation. In this study we track down the transcriptome of pre-determined cells throughout differentiation to identify the dynamic transcriptional profile from the antenna, eye, leg, genitalia and wing of *Drosophila melanogaster* development. We identified three main sets of genes: commonly expressed genes that change across time, tissue-specific genes and time-tissue specific genes. Our analyses suggest that although differences among tissues increase over time a conserved gene regulatory network is leading the differentiation process. At the early stages of differentiation, the genes contributing to tissue morphogenesis are related to cell cycle and proliferation while later to cuticle formation and neural development. A comparison of the splicing patterns shows that there are fewer differences in splicing when compared to gene expression. Nevertheless, the differences in isoform usage are mainly associated to genes known to play a role in neural fate and are found in late stage, suggesting that splicing may play a role during differentiation. Overall we observe that the transcriptome diverge as tissues become more specified. Finally to further characterize cell sub-populations in tissue development we analyzed the four compartments of wing primordia and identified genes that are essential to maintain organ structure and compartment formation.

1) Model: *Drosophila* tissue differentiation

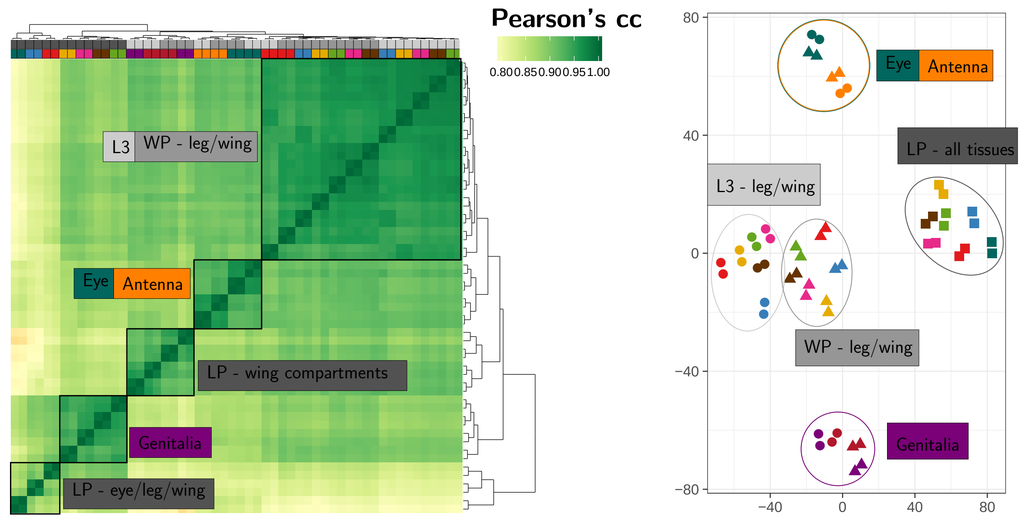
We used GAL4/UAS system for labeling with GFP tissue pre-determined cells and follow them across development. Antenna, eye, leg, wing and genitalia cells from larva 3, white pupa and late pupa stages were FACS sorted and RNA was extracted to generate stranded, polyA selected libraries for NGS.



2) Hierarchical clustering and t-SNE

Tissue samples cluster predominantly by the late stage

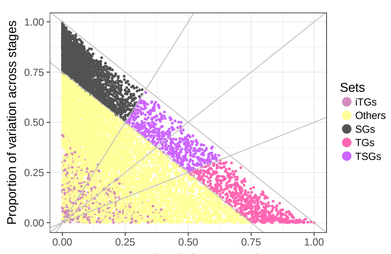
Hierarchical clustering of gene expression correlation (left) and t-distributed stochastic neighbor embedding (t-SNE[1]) (right) based on the expression of 17158 genes in 5 tissues, 4 wing compartments and 3 developmental stages. The samples show a predominant clustering by the late stage (late pupa) follow by the clusters of neural tissues and genitalia in the early stages.

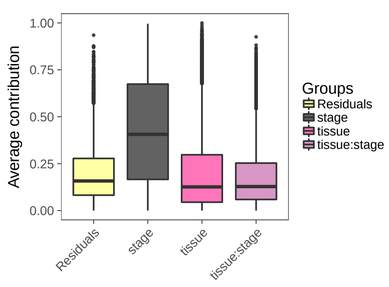


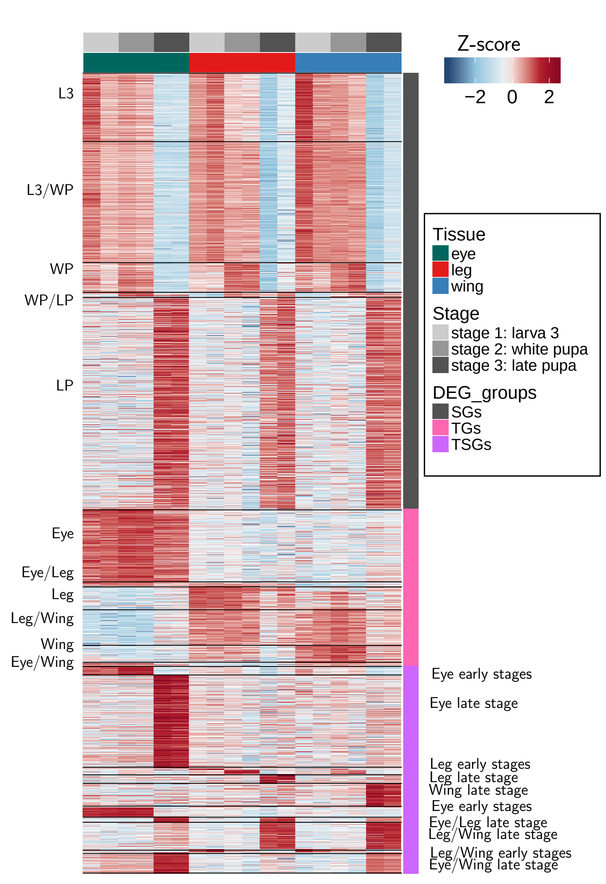
3) Differential gene expression analysis

Gene expression varies predominantly across developmental stages than across tissues

Proportion of the gene expression variation explained by tissues (x-axis) and by developmental stages (y-axis) for 17158 genes (*dots*) using linear models[2] (left). We defined 4 sets of genes: (i) genes whose expression varies considerably across stages and little across tissues, SGs (*gray*); (ii) genes whose expression varies considerably across tissues and little across stages, TGs (*pink*); (iii) genes whose expression varies considerably across tissues as well as across stages, TSGs (*purple*); and (iv) genes with variation due to interaction tissues:stages, iTSGs (*dark pink*). Differentially expressed genes (DEG) across tissues (TGs), stages (SGs) and both tissue and stage (TSGs) were defined using edgeR[3]. Gene expression values are normalized to z-scores. There are more stage-specific genes than tissue-specific genes.



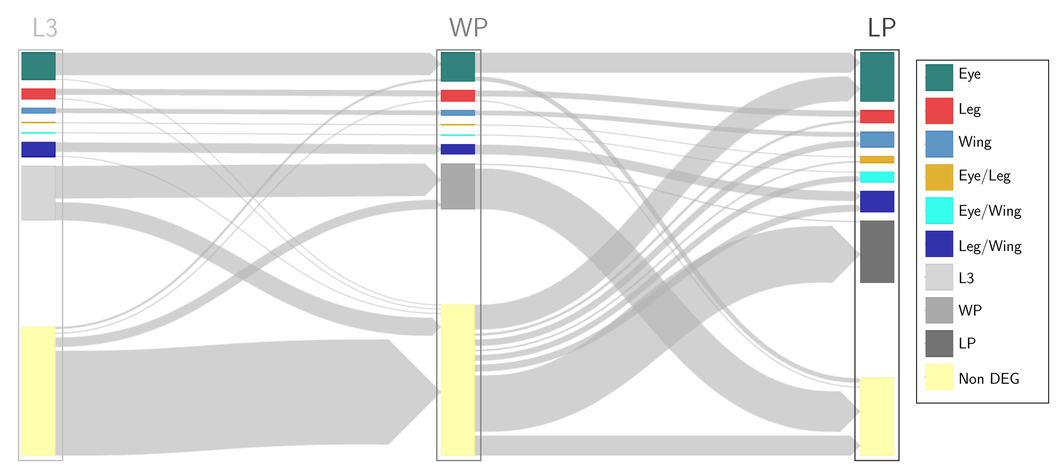




4) RNA dynamics through differentiation

Transcriptome diverges as tissues become more specialized

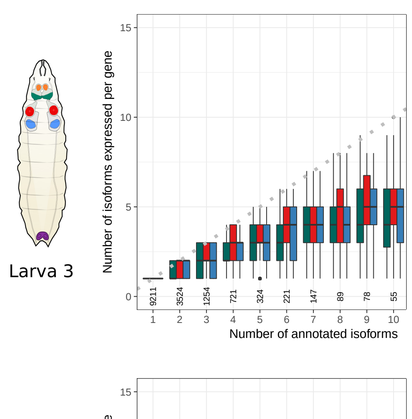
Dynamics of DEG across fly development. Notably there are more eye, leg and wing-specific genes, respectively, at the late stage. Overall, the transcriptome diverges as tissues become more specialized.

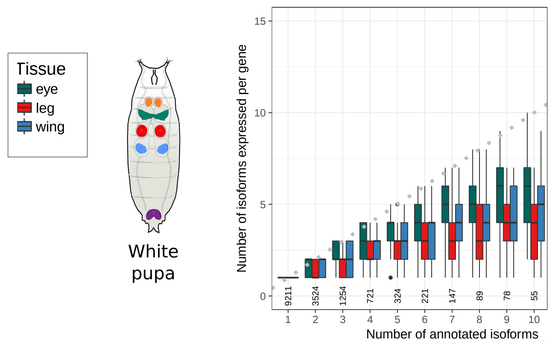


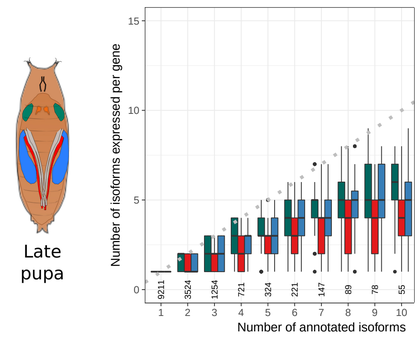
5) Isoform usage analysis

More diversity in isoform usage in differentiated tissues

We quantify transcript expression in eye, leg and wing tissues along the three developmental stages using RSEM[4]. Box plots represent the number of isoforms expressed per gene considering the number of transcripts annotated per gene. Genes with more than ten annotated isoforms were excluded. Our analysis shows more diversity in isoform usage in the late pupa stage where tissues are almost formed, meaning that the number of splicing events occuring during the process is higher in differentiated tissues. Overall tissues, eye tend to have more variability in isoform usage.



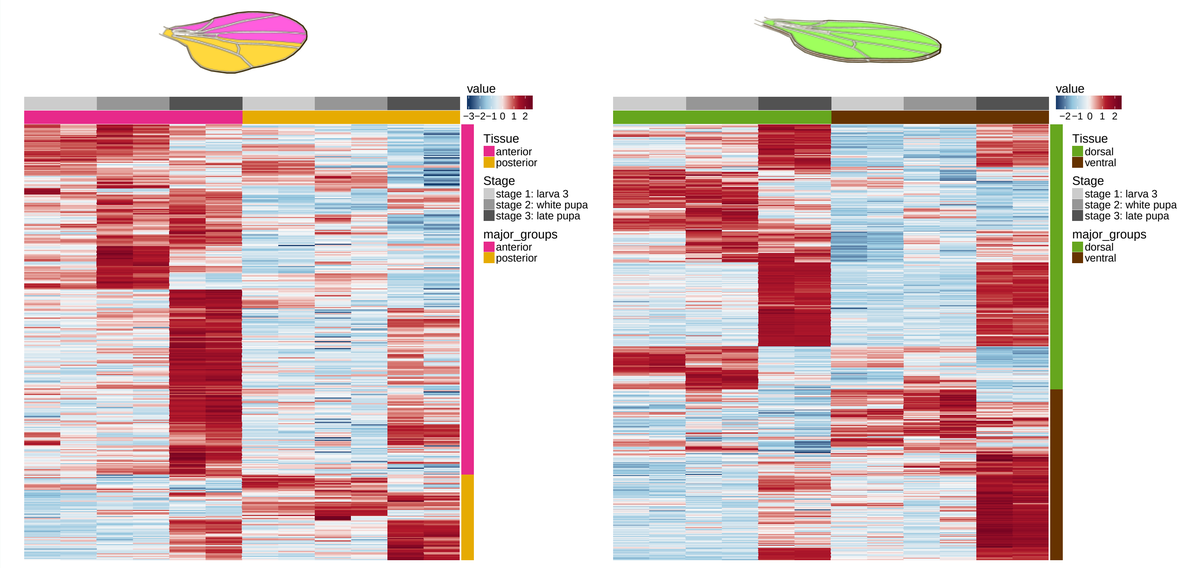




6) Compartment characterization in wing development

Compartments' signature is changing in the course of development

To characterize compartment formation during wing development, we separated anterior from posterior and ventral from dorsal cells and compared their transcriptional landscape. We used EdgeR[3] to describe differentially expressed genes (DEG) between sub-populations. Heatmaps of gene expression show that compartments' signature is changing in the course of development as few genes are DE during the whole process and many DEG appear to be stage specific, specially in the late stage.



References

[1] Laurens van der Maaten and Geoffrey E. Hinton. Visualizing high-dimensional data using t-sne. *Journal of Machine Learning Research*, 9:2579–2605, 2008.

[2] Alessandra Breschi, Sarah Djebali, Jesse Gillis, Dmitri D. Pervouchine, Alex Dobin, Carrie A. Davis, Thomas R. Gingeras, and Roderic Guigó. Gene-specific patterns of expression variation across organs and species. *Genome biology*, 17(1), 2016.

[3] Mark D. Robinson, Davis J. McCarthy, and Gordon K. Smyth. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*, 26(1):139–140, 2010.

[4] Bo Li and Colin N. Dewey. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC bioinformatics*, 12(1):323+, 2011.

cecilia.klein@crg.es