Adaptation in genes, duplicates, families, functional modules and genomes

Contents

Contents **2**

1 Introduction **4**
1.1 Adaptive changes to evolutionary speed 5
1.2 Evolution, and the detection at molecular level 5
1.3 Grouping genes and finding evolutionary patterns 5
1.4 What is DNA?How genes rose? 5
1.5 Life in DNA, from genes to repetitive elements. 5

I Structure and dynamics of genomes **6**

2 Random-like structure of DNA **7**
2.1 Introduction 8
2.2 Results and Discussion 8
2.3 Material and methods 8

3 Life inside genomes, dynamics and predictions **9**
3.1 Genomic elements, dispersion and abundance 10
3.2 *Species* Abundance Diversity in genomes 10
3.3 Neutrality of SAD 10
3.4 Material Methods 10
3.4.1 Ecolopy 10

II Detection of selective pressures in genomes **11**

4 Searching for evolutionary patterns in funcionally linked group of genes **12**
4.1 Introduction 13
4.2 Material and Methods 13
4.2.1 Dataset 13
Five mammals 13
6 *Drosophila* 13
4.2.2 Alignments 13

4.3 open on colocalization → not random 13

5 Tools, programs, methods **15**
5.1 ETE-evol plugin 16
5.1.1 BRANCHED1 16
5.1.2 Protamines Rodents and Primates 16
5.2 Pipeline for study of adaptation at genomic scale 16
5.2.1 Selective pressure on duplicated genes in Drosophila . . . 16
5.3 Phylemon 16

6 Conclusions **17**

Bibliography **i**

List of Figures **ii**

List of Tables **iii**

Chapter 1

Introduction

1.1 Adaptive changes to evolutionary speed

1.2 Evolution, and the detection at molecular level

1.3 Grouping genes and finding evolutionary patterns

1.4 What is DNA? How genes rose?

1.5 Life in DNA, from genes to repetitive elements.

Part I

Structure and dynamics of genomes

Chapter 2

Random-like structure of DNA

2.1 Introduction

2.2 Results and Discussion

2.3 Material and methods

Chapter 3

Life inside genomes, dynamics and predictions

3.1 Genomic elements, dispersion and abundance

3.2 *Species* Abundance Diversity in genomes

3.3 Neutrality of SAD

3.4 Material Methods

3.4.1 Ecolopy

Part II

Detection of selective pressures in genomes

Chapter 4

Searching for evolutionary patterns in funcionally linked group of genes

4.1 Introduction

4.2 Material and Methods

4.2.1 Dataset

Five mammals

Complete genomes of 5 mammals species (*Homo sapiens*, *Pan troglodytes*, *Mus musculus*, *Rattus norvegicus* and *Canis familiaris*) where retrieved from *Ensembl* [Flicek *et al.*2011]. Also orthology prediction between each pair of species possibly done between human and the others was retrieved from *Ensembl Compara* [Vilella *et al.*2009] using biomart [Kinsella *et al.*2011]. Only groups of orthologs *one-to-one* with one representative of each species where kept in the final dataset. 4.1 NUMBERS

6 *Drosophila*

4.2.2 Alignments

Each of the group of orthologous sequences were aligned with Muscle [Edgar2004], and, once aligned sequences were cleaned with trimal [Capella-Gutiérrez *et al.*2009] keeping all sequences but trimming aligment columns with the euristic1 method.

4.3 open on colocalization → not random

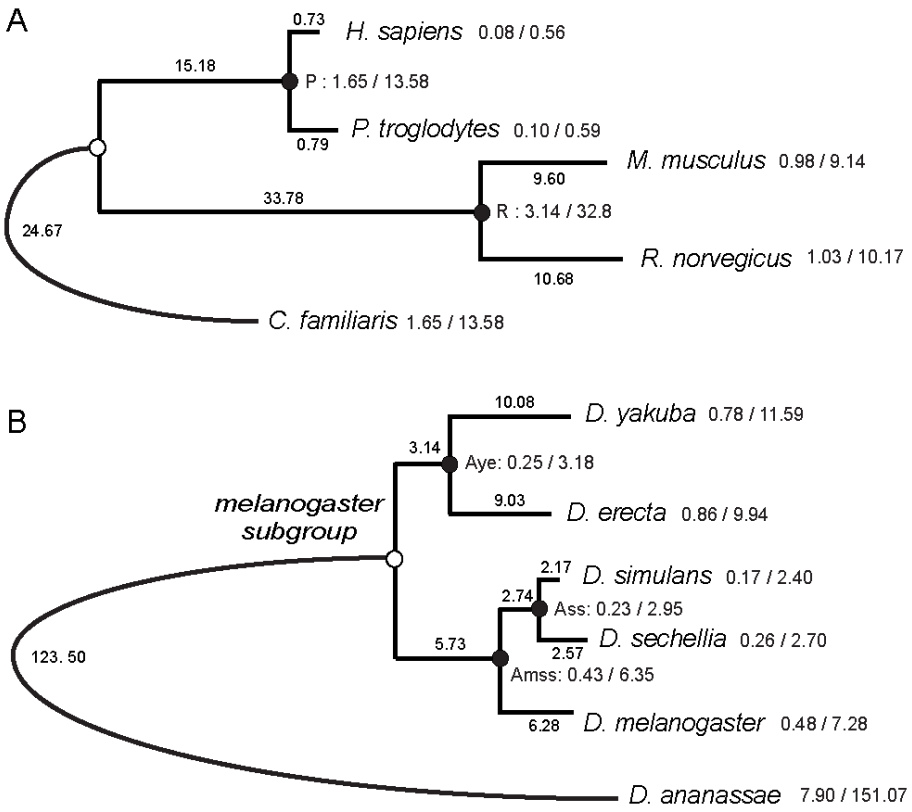

Figure 4.1: Mammals and *Drosophila* phylogeny. blabli blob lu dkfnlskjdf

Chapter 5

Tools, programs, methods

5.1 ETE-evol plugin

5.1.1 BRANCHED1

5.1.2 Protamines Rodents and Primates

5.2 Pipeline for study of adaptation at genomic scale

5.2.1 Selective pressure on duplicated genes in Drosophila

5.3 Phylemon

Chapter 6

Conclusions

Bibliography

List of Figures

4.1 Mammals and *Drosophila* phylogeny 14

List of Tables