Scientific Programming Practical 1

Introduction

Luca Bianco - Academic Year 2018-19
luca.bianco@fmach.it

Outline

- Personal introduction
- Introduction to the practical
- Hands-on practical

About me

Computer Science

Ph.D. at the University of Verona, Italy, with thesis on Simulation of Biological Systems

Research Fellow at Cranfield University - UK

Three years at Cranfield University working at proteomics projects (GAPP, MRMaid, X-Tracker...)

Module manager and lecturer in several courses of the MSc in Bioinformatics

Bioinformatician at IASMA – FEM

Currently bioinformatician in the Computational Biology Group at Istituto Agrario di San Michele all'Adige – Fondazione Edmund Mach, Trento, Italy

Collaborator uniTN - CiBio

I run the Scienitific Programming Lab A last year

Fondazione Edmund Mach

FEM – San Michele, Trento - Italy

Agricultural Institute

Research and Innovation Centre

Genomics, metabolomics wet labs on fruits (apple, grape, small fruits,…)

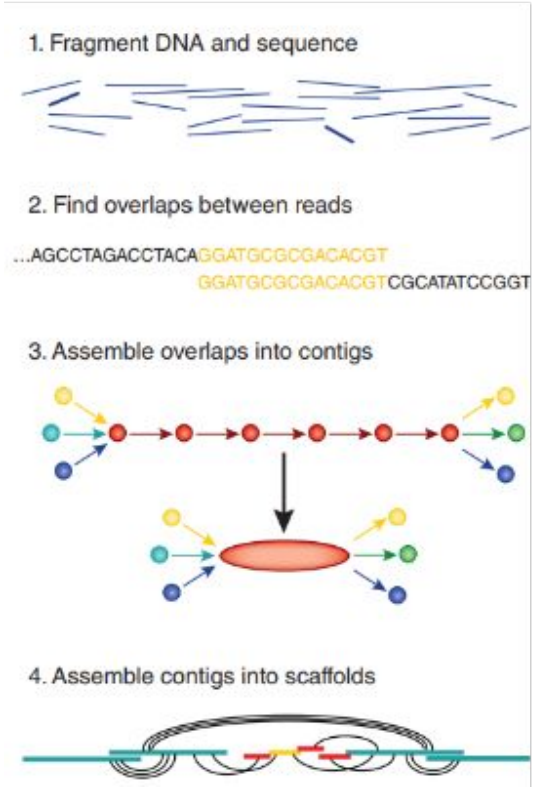
Bioinformatics and computational biology

Bioinformatics @FEM (UBC)

- Genomics
  - Assembly and annotation of complex genomes (plants, insects, etc.)
  - Development of SNP Chips for genetic screening
  - Resequencing of genomes
- Metagenomics
  - Targeted metagenomic data
  - *Feature selection* algorithms
  - Algorithms for strain-level identification from un-targeted metagenomics
- Transcriptomics
  - RNA-seq data analysis, gene and pathway enrichment
  - Data integration and compilation of expression atlases
- Metabolomics
  - Data analysis pipelines for targeted and untargeted data
  - Methods for MS imaging
- Statistical data analysis
  - Integration of –omic data and analysis of correlation networks

Bix @FEM - Examples

Genome assembly



[from M. Baker, Nature Methods, 2014]

Bix @FEM - Examples

Genome assembly of DH of Pear and Apple

Input data:

Illumina ~60x – 100x PE information + (mate pairs for Apple)
Pacific Biosciences ~ 30x + 30x (35x only for Apple)
Bionano optical maps: ~ 600x (for both)
Genetic maps: integrated genetic map from 21 mapping populations (Apple only)

Output result (for Apple, Pear is in progress):

Chromosome scale assembly

Contigs: 2150 for a total of 625Mb
N50 Contigs (hybrid dbg2olc): ~ 620Kbps
280 Scaffolds, for an N50 5,6Mb
17 chromosomes + lg0 unanchored sequences

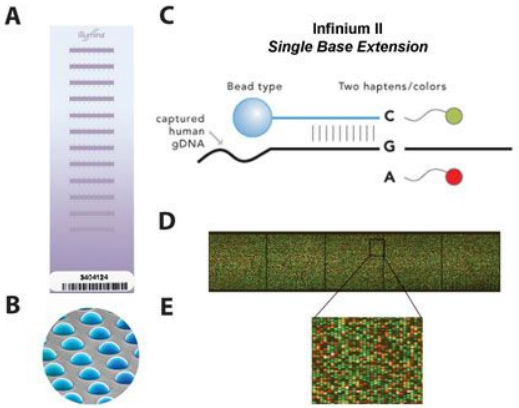
[Daccord et al, Nature Genetics, 49, 2017]

Bix @FEM - Examples

SNP-Chips development for GWAS

20K SNP Illumina Infinium II Array (reseq of 16 Apple cultivars, Illumina 30x)
487K SNP Affymetrix Axiom Array (reseq of 63 Apple cultivars, Illumina 20-30x)

1. Reads alignment and filtering
2. SNP calling
3. Identification of most reliable SNPs
4. Selection of (20K) 487K target SNPs



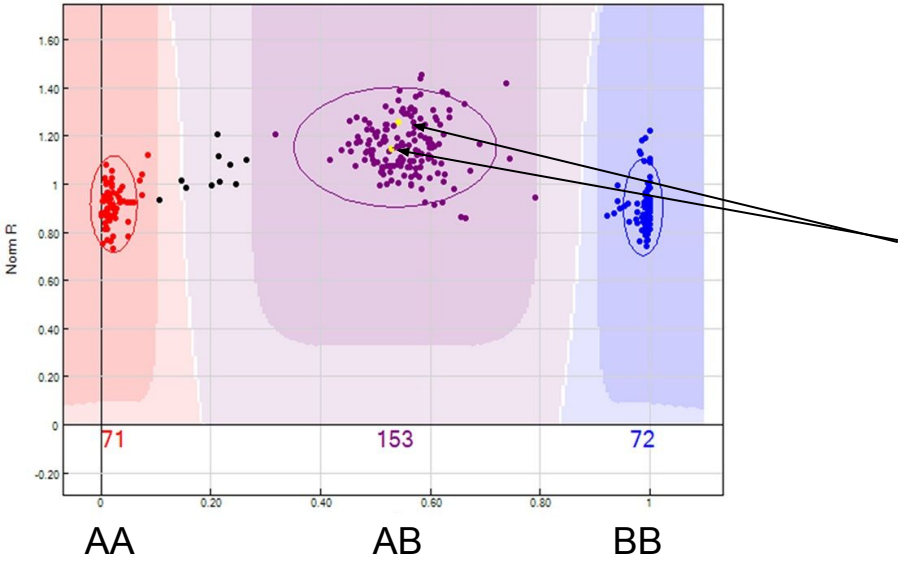
Several Terabytes of data produced!!!!

Peach, pear and walnut done too!

Bix @FEM - Examples

SNP-Chips development for GWAS

20K SNP Illumina Infinium II Array (reseq of 16 Apple cultivars, Illumina 30x)
487K SNP Affymetrix Axiom Array (reseq of 63 Apple cultivars, Illumina 20-30x)

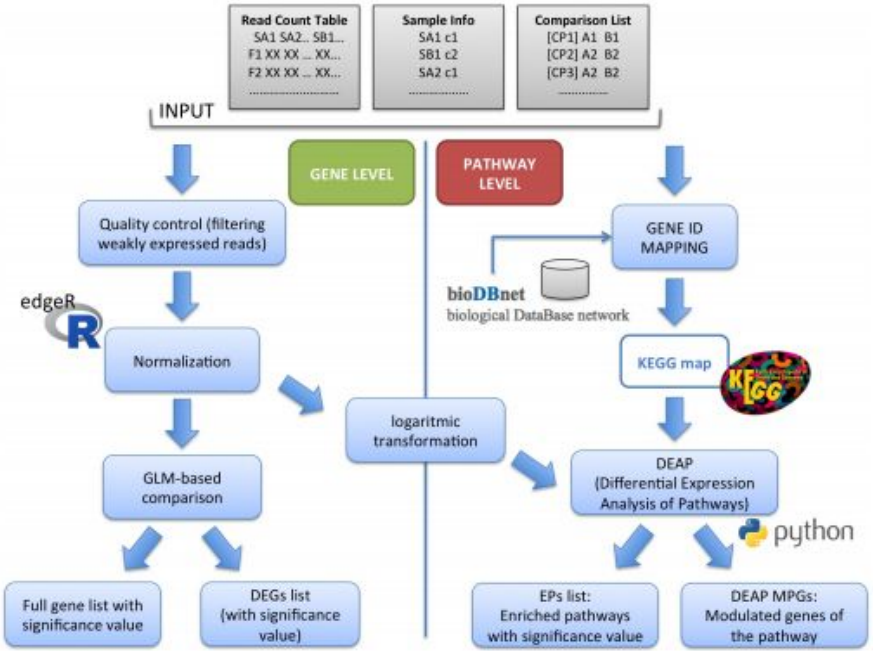


Task:
Analyze 500,000 of these…
(1 x SNP)

parents

Bix @FEM - Examples

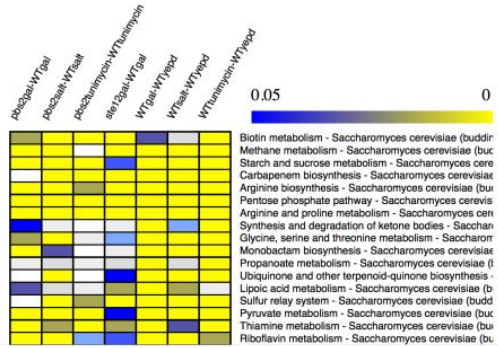
RNAseq data analysis with Pathway Inspector

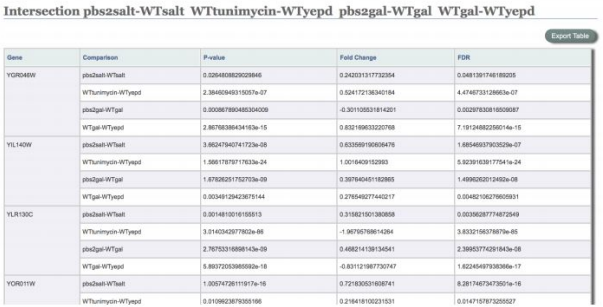


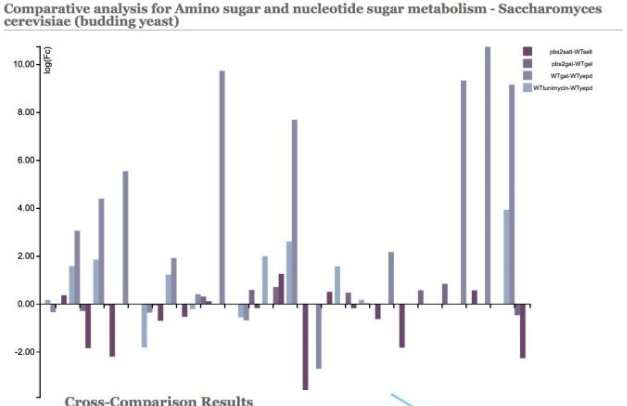
https://pathwayinspector.fmach.it

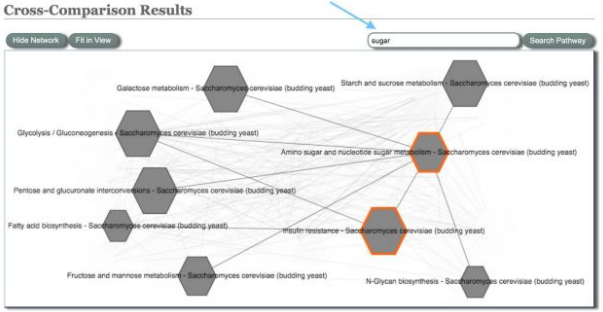
Bix @FEM - Examples

RNAseq data analysis with Pathway Inspector









https://pathwayinspector.fmach.it

Opportunities @FEM

MSc External thesis

Are you interested in a bioinformatics project in NGS data analysis, RNA Seq, data integration?

Come and talk to me or email me at:

luca.bianco@fmach.it

Scientific Programming Practical

Back to business now!

Scientific Programming Practical

In this practical you will

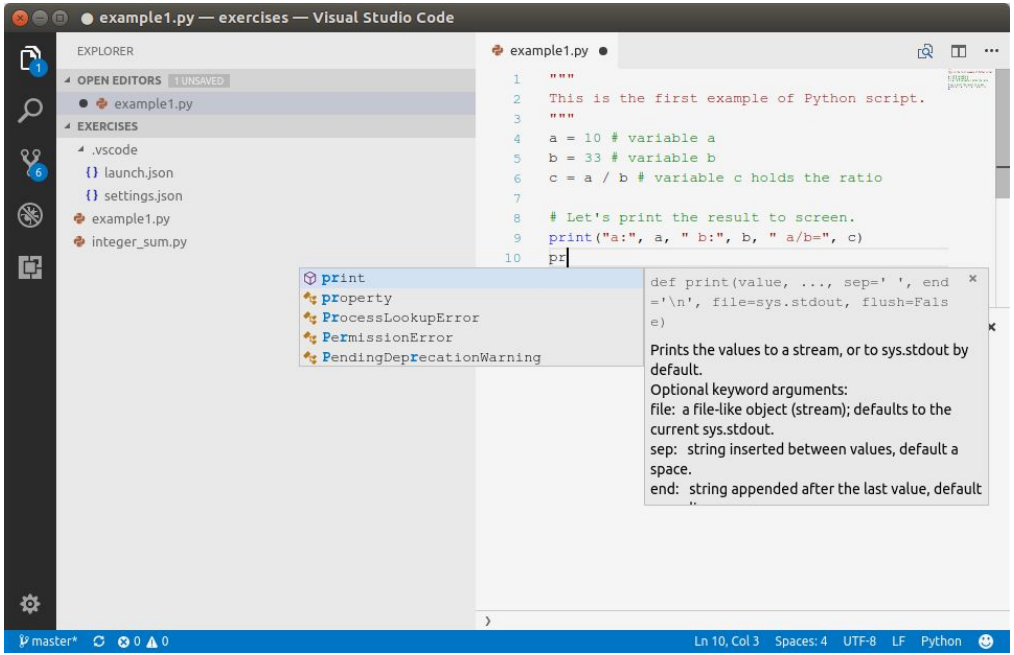
1. Install Python 3.x (and pip)
2. Install Visual Studio Code
3. Get familiar with the Python console
4. Start using Visual Studio Code and advanced features (like debugging)
5. End the session with some exercises

Scientific Programming Practical

Console VS. Integrated Development Environment (IDE)

```
biancol@bluhp:~$ python3
Python 3.5.2 (default, Aug 18 2017, 17:48:00)
[GCC 5.4.0 20160609] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>>
```

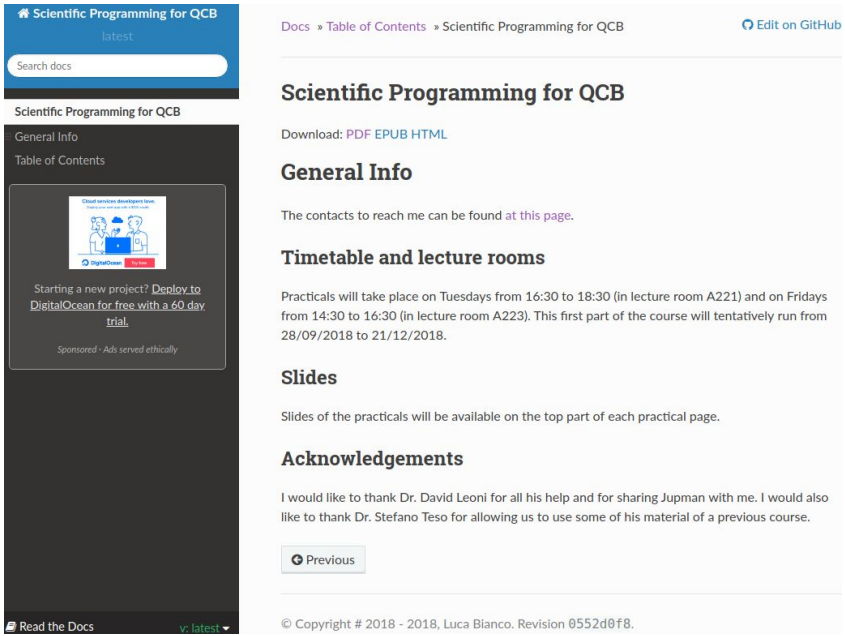
```
>>> print("Hi there")
Hi there
>>> print("{} + {} = {}".format(10,20, 10+20))
10 + 20 = 30
>>>
```



Resources

All material regarding practicals will be found here:

http://qcbsciprolab.readthedocs.io



@
luca.bianco@fmach.it

Timetable

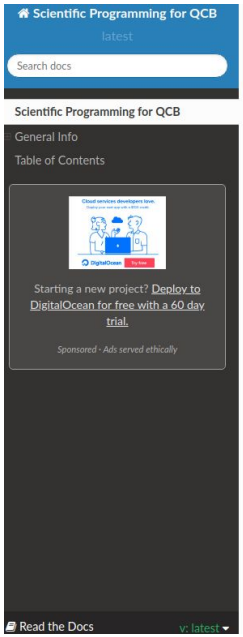
Tuesdays:

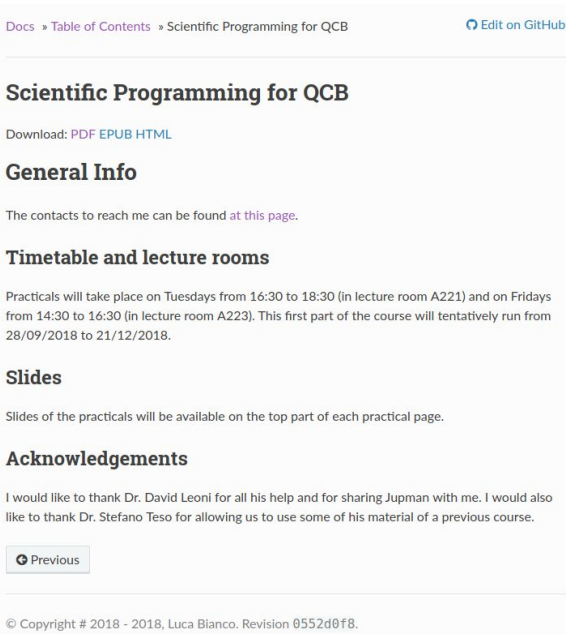
A221: 16,30 - 18,30

Fridays:

A223: 14,30 - 16,30

Most probably we will skip October 30th practical.



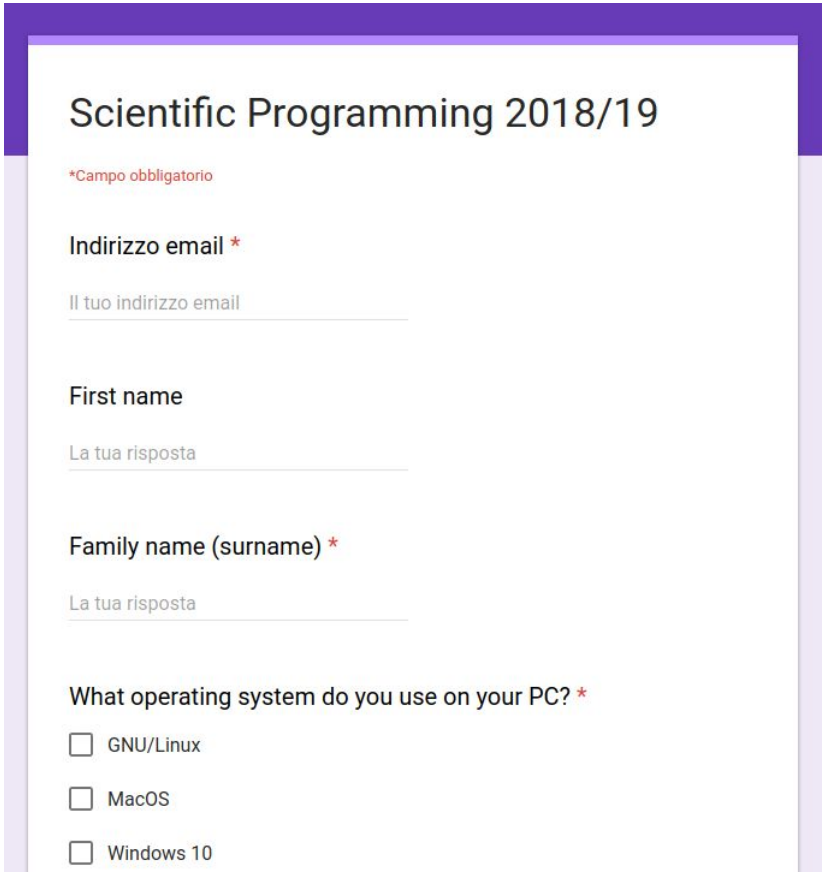


@
luca.bianco@fmach.it

Please, fill the form at

https://tinyurl.com/spquest01

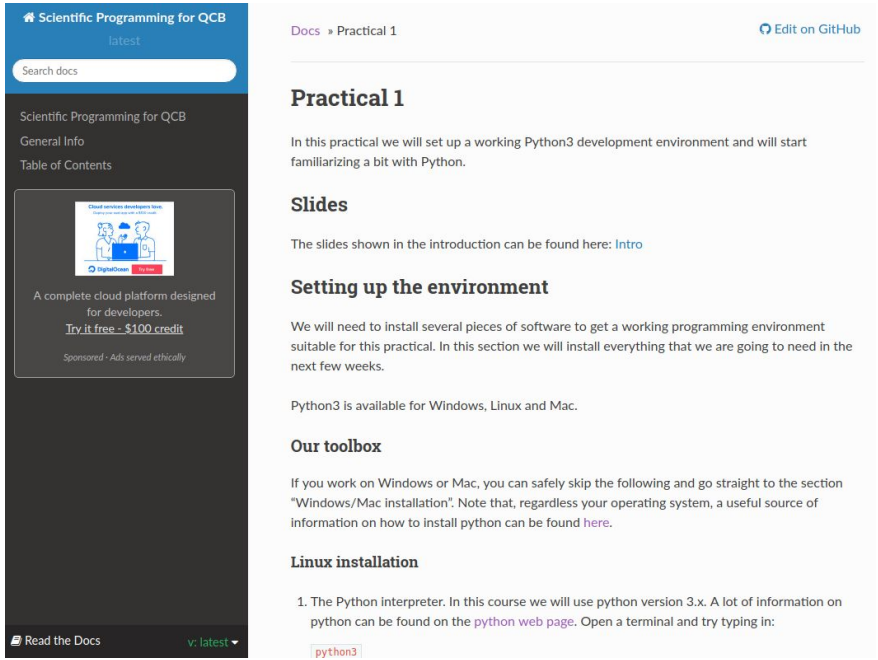
Deadline Sunday, September 30th



Any questions?

If not, go to:

https://qcbsciprolab.readthedocs.io/en/latest/introduction.html



@
luca.bianco@fmach.it