

Scientific Programming

Practical 1

Introduction

Luca Bianco - Academic Year 2017-18
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, MR Maid, X-Tracker...)

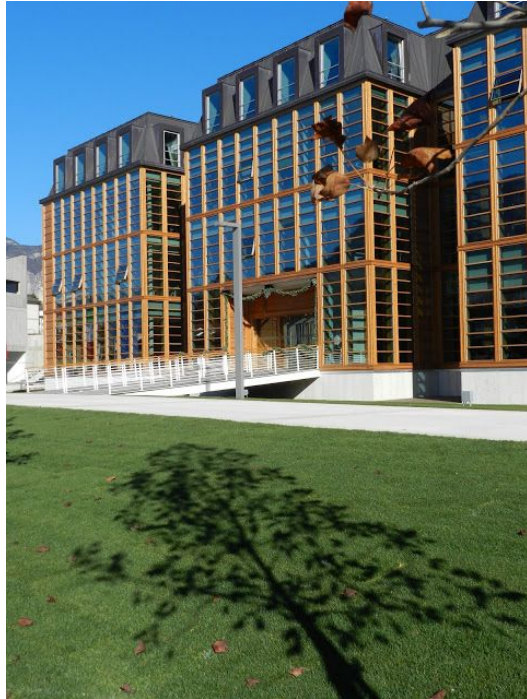
Module manager and lecturer in several courses of the MSc in Bioinformatics

Bioinformatician at IASMA – FEM

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

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

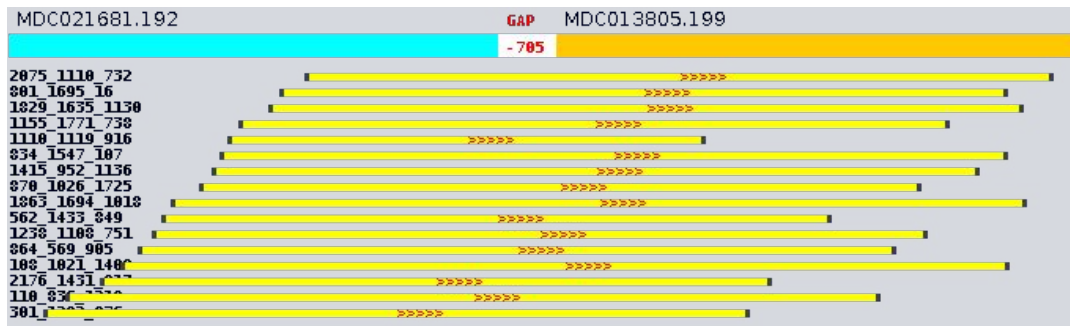
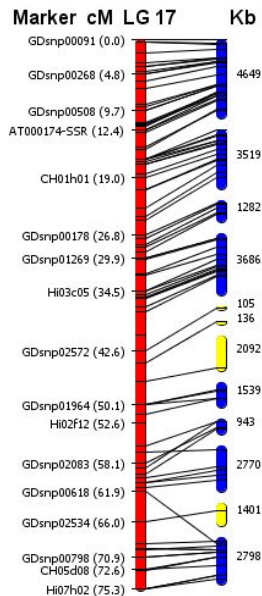
❖ Statistical data analysis

- Integration of -omic data and analysis of correlation networks

Bix @FEM -- Examples

Improving the Apple Assembly with SOLiD&Illumina data

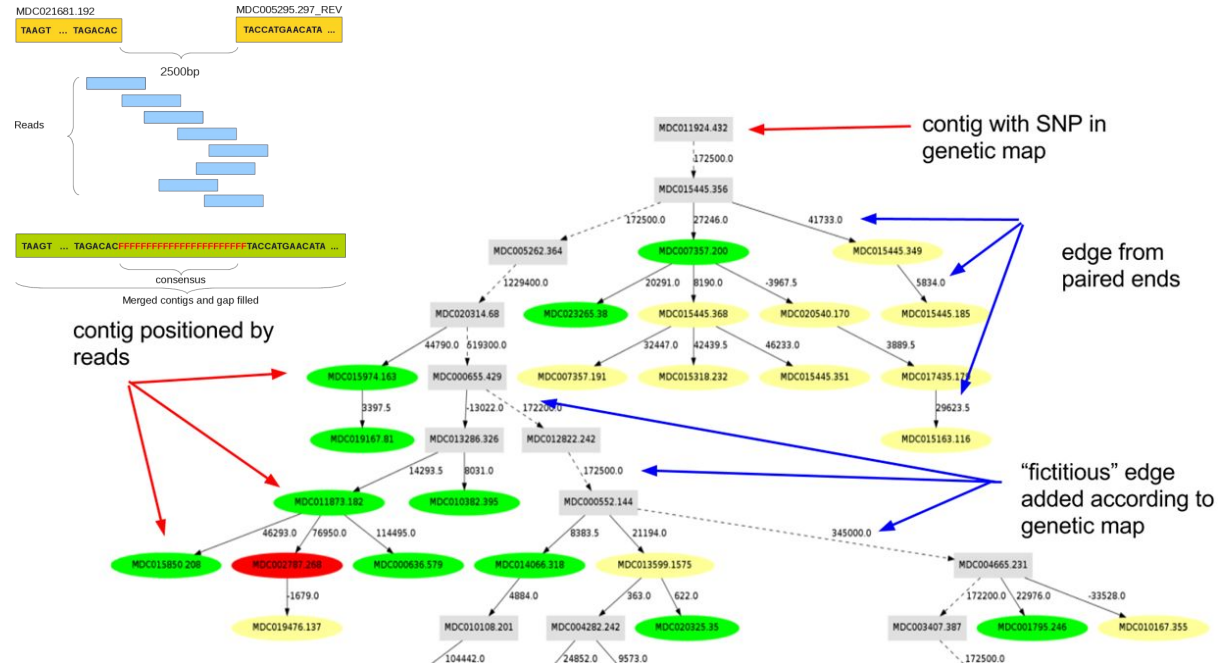
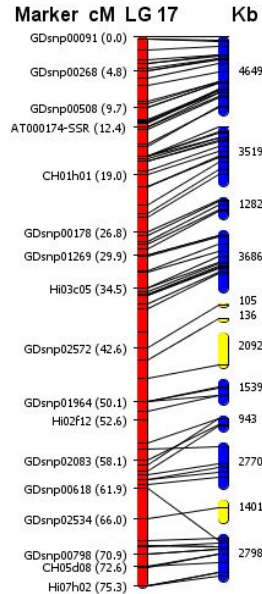
Spatial relationships with long jump libraries (5K, 7.5K and 10K mate pairs)

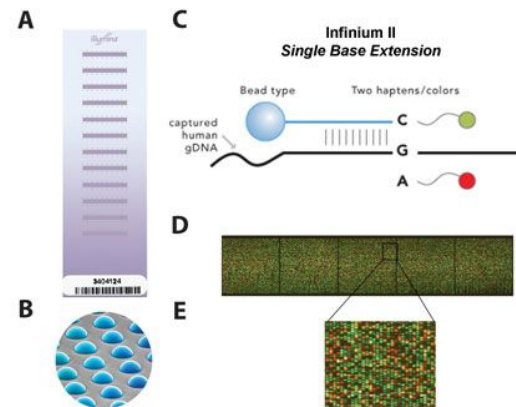


Bix @FEM -- Examples

Improving the Apple Assembly with SOLiD&Illumina data

Spatial relationships with long jump libraries (5K, 7.5K and 10K mate pairs)



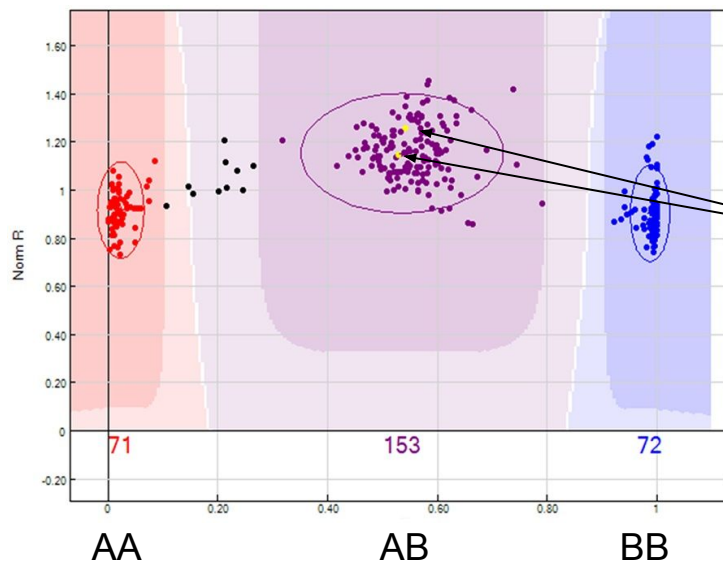


Bix @FEM -- Examples

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 500k of these...
(1 x SNP)

parents

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

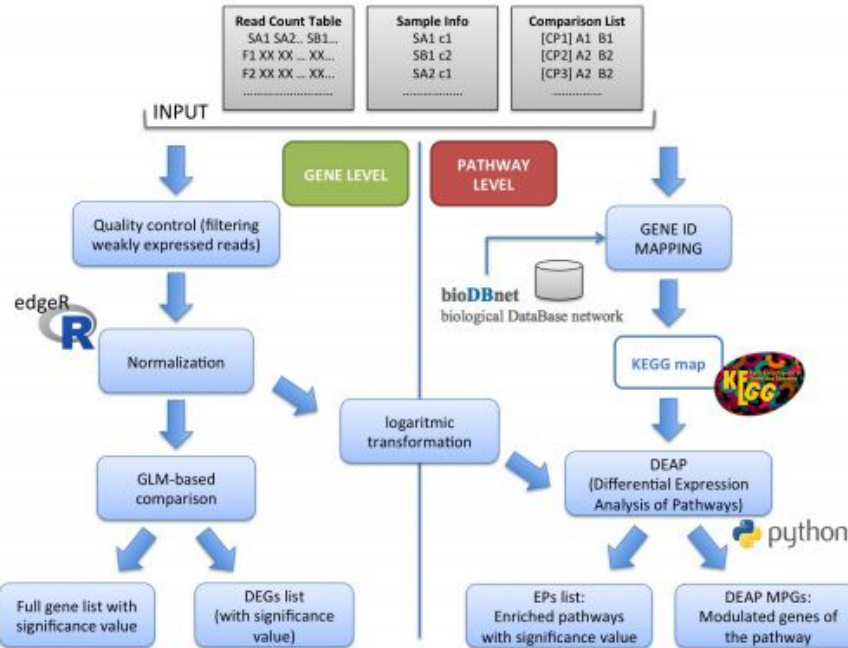
17 chromosomes + Ig0 unanchored sequences



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

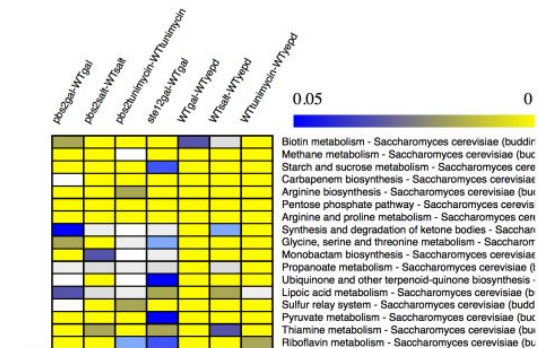
Bix @FEM -- Examples

RNAseq data analysis with Pathway Inspector



Bix @FEM -- Examples

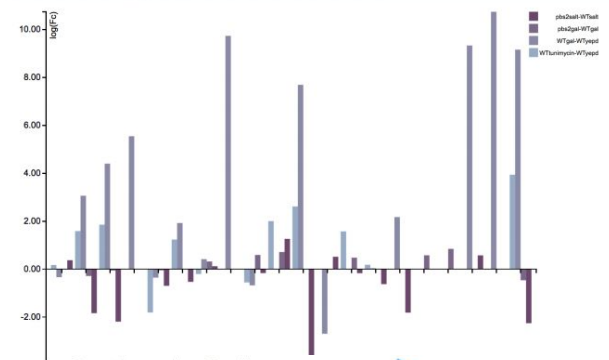
RNAseq data analysis with Pathway Inspector



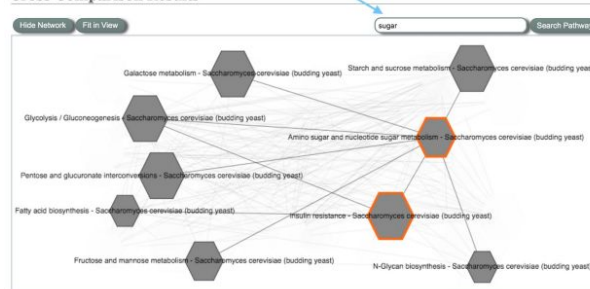
Intersection pbs2gal-WTsalt WTtunimycin-WTyeptd pbs2gal-WTgal WTgal-WTyeptd

Gene	Comparison	P-value	Fold Change	FDR
YOR04W	pbs2gal-WTsalt	0.02848082028946	0.2420313772254	0.0481361748189205
	WTtunimycin-WTyeptd	2.38460349315057e-07	0.524172138345184	4.474873128663e-07
	pbs2gal-WTgal	0.00087890548504029	-0.30100531814261	0.0207830810500887
	WTgal-WTyeptd	2.86788386434103e-15	0.83218983220708	7.191248826014e-15
YIL140W	pbs2gal-WTsalt	3.85247540741733e-08	0.833589168058476	1.85448937803052e-07
	WTtunimycin-WTyeptd	1.58817870171733e-24	1.001649102893	5.82391638177541e-24
	WTgal-WTgal	1.87826251172733e-09	0.30764051192865	1.49862020124624e-08
	WTgal-WTyeptd	0.0034912942367044	0.276548277440217	0.0548210627660931
YLR130C	pbs2gal-WTsalt	0.301481010108873	0.31882150138088	0.03038287774872549
	WTtunimycin-WTyeptd	3.4140342077802e-06	-1.5876578914284	3.8332176378679e-05
	pbs2gal-WTgal	2.7670331688143e-09	0.498214139134541	2.3895374291845e-08
	WTgal-WTyeptd	8.8037203388083e-18	-0.8312188730747	1.8224549793064e-17
YOR011W	pbs2gal-WTsalt	1.00574728111917e-16	0.721830534688741	8.2817487473501e-16
	WTtunimycin-WTyeptd	0.010862387903186	0.21641810221331	0.0147157873255527

Comparative analysis for Amino sugar and nucleotide sugar metabolism - *Saccharomyces cerevisiae* (budding yeast)



Cross-Comparison Results



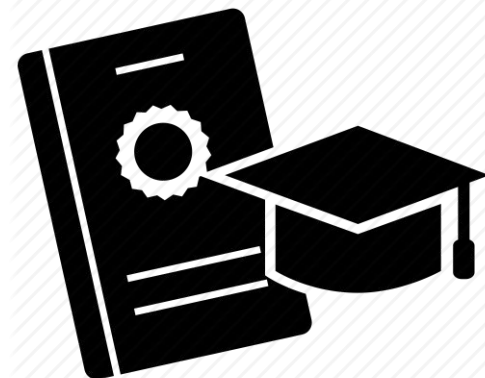
Opportunities @FEM

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

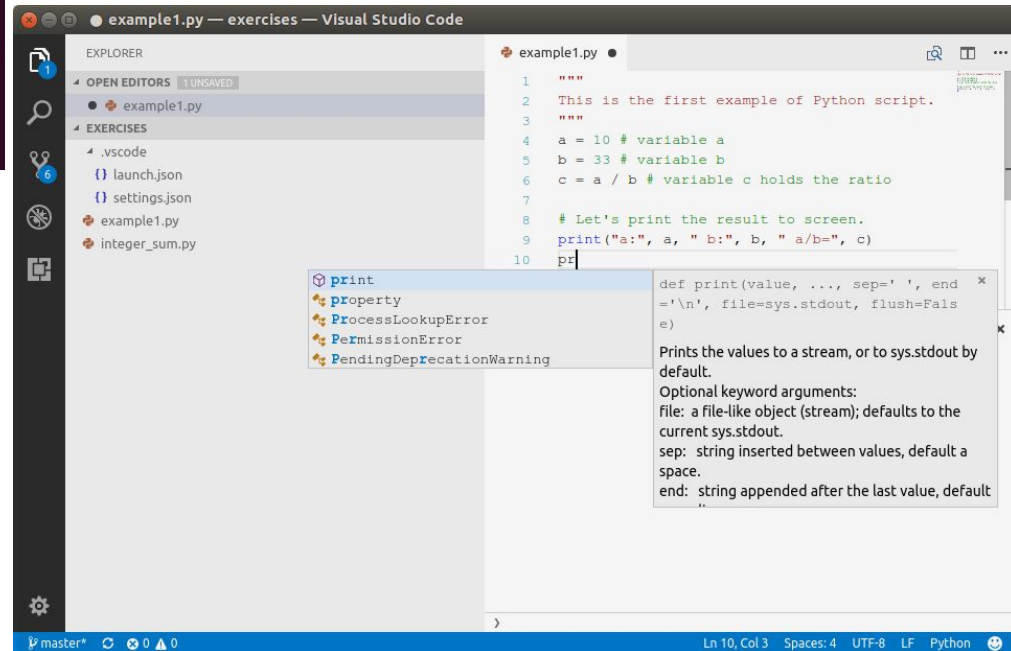
1. Install Python 3.x (and pip)
2. Install Visual Studio Code
3. Get familiar with the Python console
4. Start using Visual Studio Code and advanced features (like debugging)
5. End the session with some exercises



Scientific Programming Practical

Console VS. Integrated Development Environment (IDE)

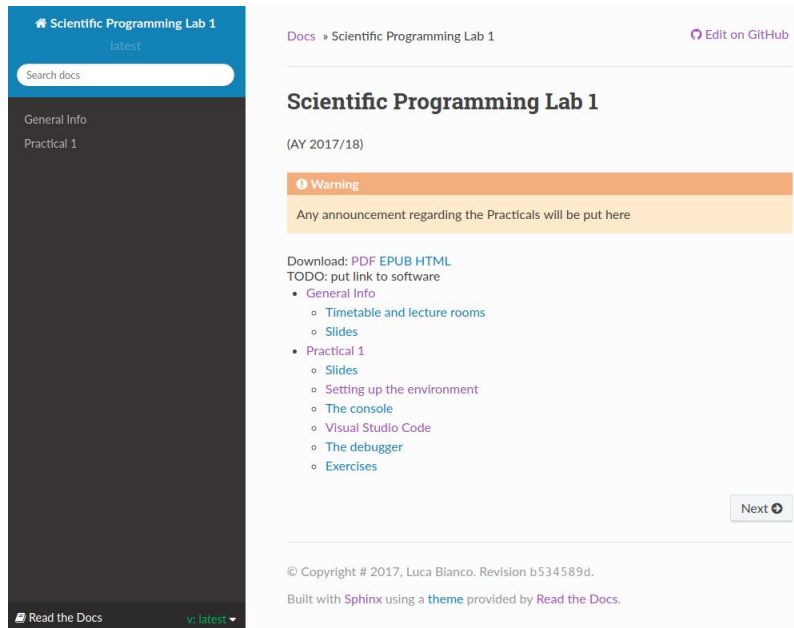
```
biancol@bluhp:~$ python3
Python 3.5.2 (default, Aug 18 2017, 17:48:00)
[GCC 5.4.0 20160609] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> 
```



Resources

All material regarding practicals can be found here:

<http://sciprolab1.readthedocs.io>



The screenshot shows the documentation page for 'Scientific Programming Lab 1' on the Read the Docs platform. The page has a dark blue header with the title and a 'latest' version indicator. Below the header is a search bar and a sidebar with links to 'General Info' and 'Practical 1'. The main content area features a warning box, download links for PDF, EPUB, and HTML, a TODO item, and a table of contents for 'Practical 1' including links to 'Timetable and lecture rooms', 'Slides', 'Setting up the environment', 'The console', 'Visual Studio Code', 'The debugger', and 'Exercises'. A 'Next' button is visible at the bottom right of the content area. The footer contains copyright information and mentions the use of Sphinx and a theme by Read the Docs.

Scientific Programming Lab 1
latest

Search docs

General Info
Practical 1

Docs » Scientific Programming Lab 1 [Edit on GitHub](#)

Scientific Programming Lab 1

(AY 2017/18)

Warning
Any announcement regarding the Practicals will be put here

Download: [PDF](#) [EPUB](#) [HTML](#)
TODO: put link to software

- General Info
 - Timetable and lecture rooms
 - Slides
- Practical 1
 - Slides
 - Setting up the environment
 - The console
 - Visual Studio Code
 - The debugger
 - Exercises

Next

© Copyright # 2017, Luca Bianco. Revision b534589d.
Built with Sphinx using a theme provided by Read the Docs.



luca.bianco@fmach.it

Any questions?

If not, go to:

<http://sciprolab1.readthedocs.io/en/latest/introduction.html>

Scientific Programming Lab 1

latest

Search docs

General Info

▢ Practical 1

- Slides
- ▢ Setting up the environment
 - The console
 - Visual Studio Code
 - The debugger
 - Exercises

Docs » Practical 1

 Edit on GitHub

Practical 1

In this practical we will set up a working Python3 development environment and will start familiarizing a bit with Python.

Slides

Slides of the introduction can be found here: [Intro](#)

Setting up the environment

We will need to install several pieces of software to get a working programming environment suitable for this practical. In this section we will install everything that we will need in the next few weeks.



luca.bianco@fmach.it