

# 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 Scientific 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 / Variant discovery

## ❖ 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

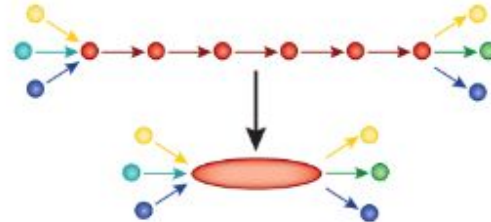
1. Fragment DNA and sequence



2. Find overlaps between reads

...AGCCTAGACCTACA**GGATGCGCGACACGT**  
**GGATGCGCGACACGT**CGCATATCCGGT

3. Assemble overlaps into contigs



4. Assemble contigs into scaffolds



[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 + 1g0 unanchored sequences**



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

# Bix @FEM - Examples

## SNP-Chips development for GWAS

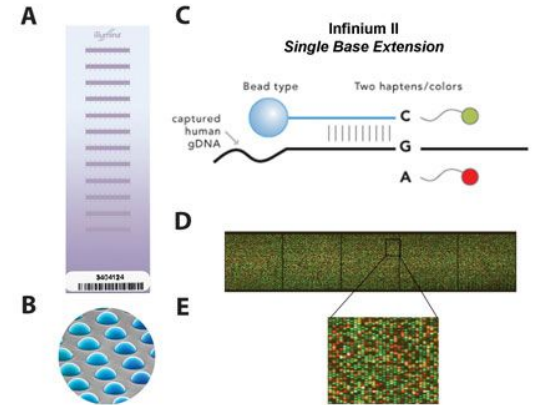
20K SNP Illumina Infinium II Array (reseq of 16 Apple cultivars, Illumina 30x)

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!**



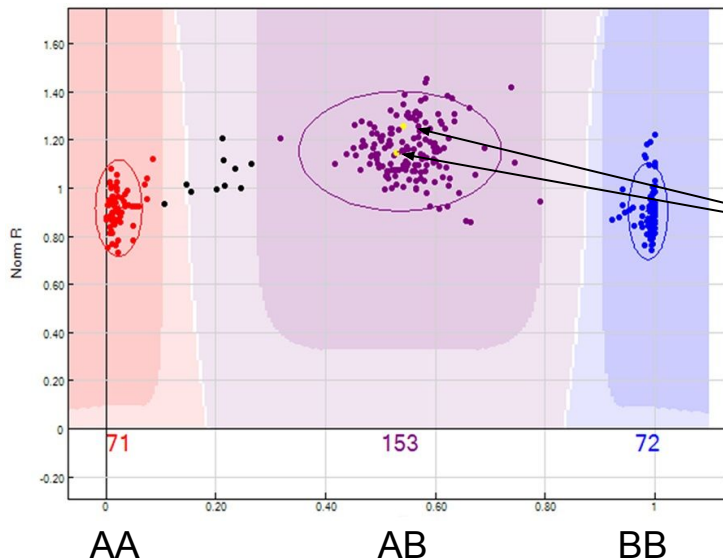


# Bix @FEM - Examples

## 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)



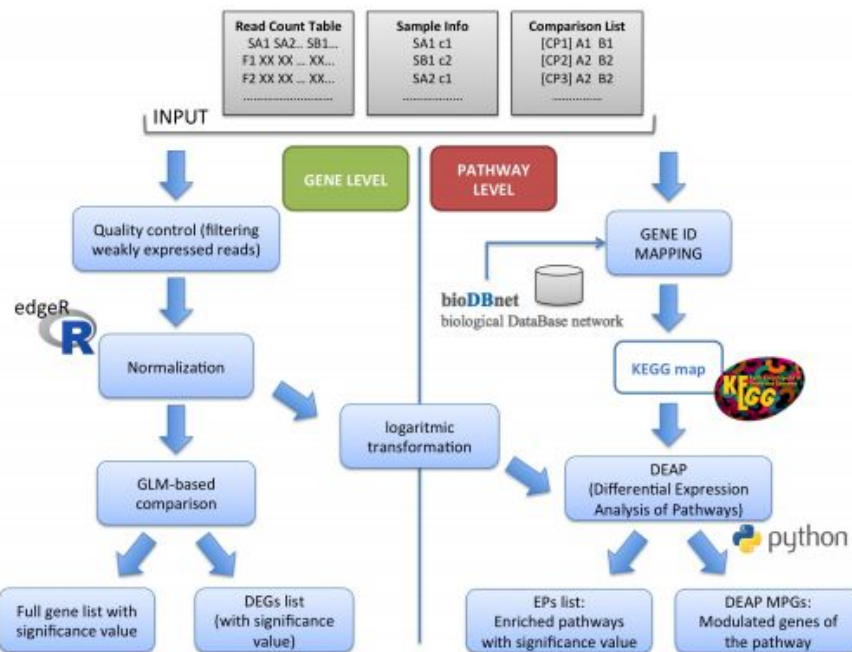
### Task:

Analyze 500,000 of these...  
(1 x SNP)

parents

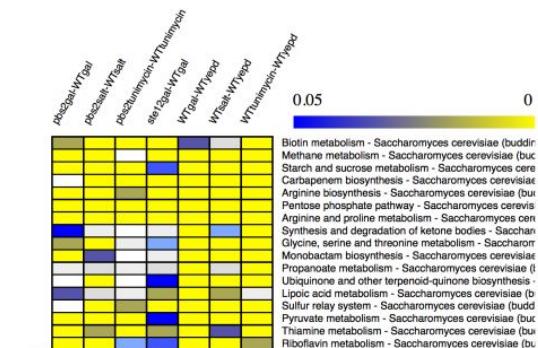
# Bix @FEM - Examples

## RNAseq data analysis with Pathway Inspector



# Bix @FEM - Examples

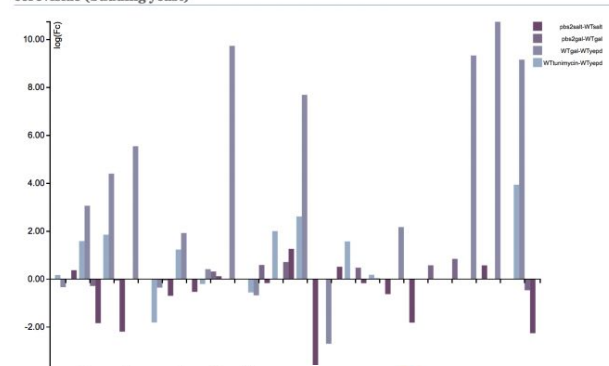
## RNAseq data analysis with Pathway Inspector



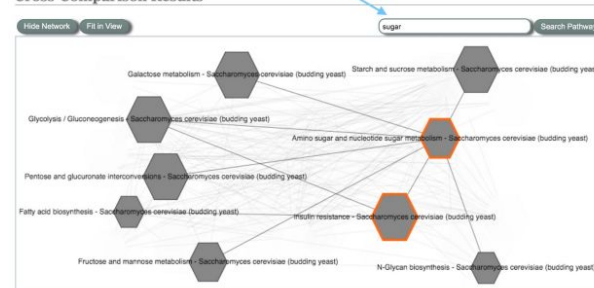
Intersection pbs2salt-WTsalt WTtunimycin-WTtpepd pbs2gal-WTgal WTgal-WTtpepd

Gene	Comparison	P-value	Fold Change	FDR
YOR04W	pbs2salt-WTsalt	0.028480382028946	0.2420313772254	0.0481361748189205
	WTtunimycin-WTtpepd	2.38460349315057e-07	0.524172138345184	4.474873128693e-07
	pbs2gal-WTgal	0.000867890485354028	-0.301105531814261	0.0207830810500887
	WTgal-WTtpepd	2.86788386434163e-15	0.83218803220708	7.191248826014e-15
YIL140W	pbs2salt-WTsalt	3.85247940741733e-08	0.83358916808476	1.88548837803025e-07
	WTtunimycin-WTtpepd	1.58817878171733e-24	1.001649102893	5.8239103817754e-24
	pbs2gal-WTgal	1.87826251172733e-09	0.30764051192865	1.4986202101462e-08
	WTgal-WTtpepd	0.3034912942367544	0.276548277440217	0.0548210627660931
YLR130C	pbs2salt-WTsalt	0.301481019188815	0.31882150138888	0.0308287774872549
	WTtunimycin-WTtpepd	3.0140342877803e-06	-1.58765788614284	3.8332178378679e-05
	pbs2gal-WTgal	2.7870331888813e-09	0.488214138134541	2.3895374291845e-08
	WTgal-WTtpepd	8.8037203388888e-18	-0.8312188730247	1.8224549730366e-17
YOR011W	pbs2salt-WTsalt	1.00574728111917e-16	0.72183031688741	8.2817487473501e-16
	WTtunimycin-WTtpepd	0.010862387803186	0.21641810221331	0.014715787255527

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



Cross-Comparison Results



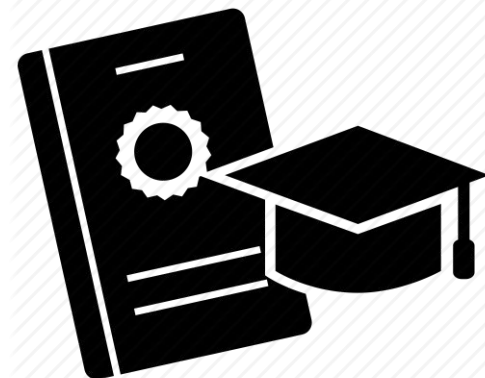
# Opportunities @FEM

## 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](mailto: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

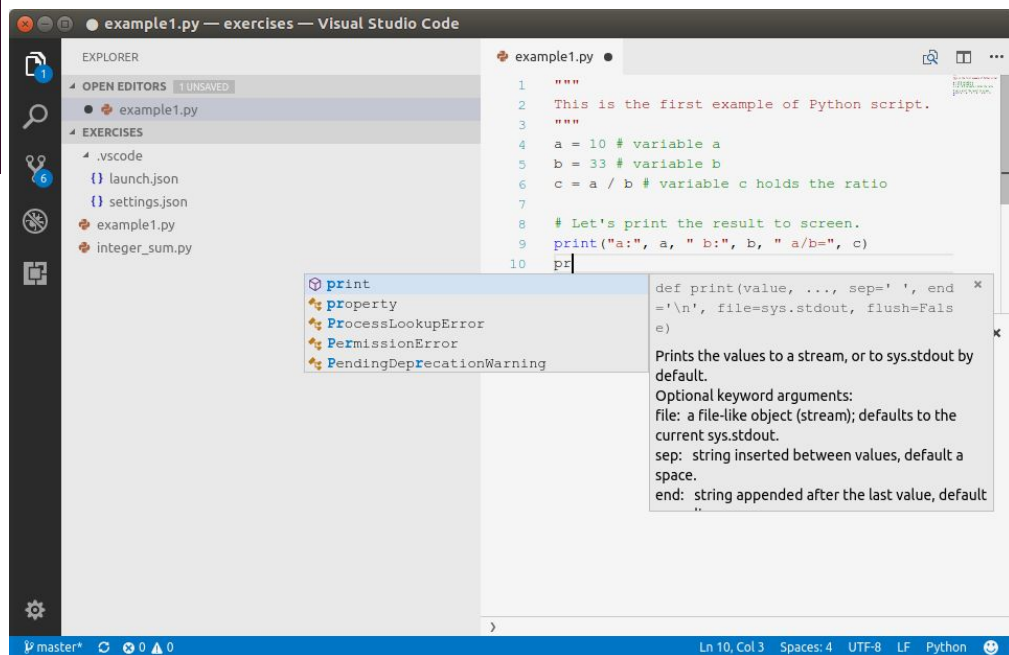


# Scientific Programming Practical

## 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.
>>>
```

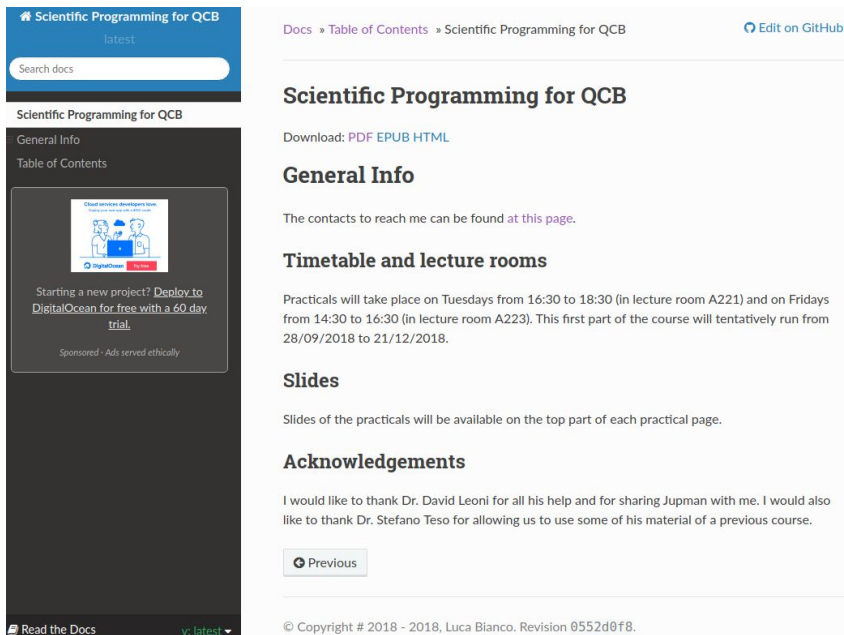
```
>>> print("Hi there")
Hi there
>>> print("{} + {} = {}".format(10,20, 10+20))
10 + 20 = 30
>>>
```



# Resources

All material regarding practicals will be found here:

**<http://qcbsciprolab.readthedocs.io>**



The screenshot shows the Read the Docs interface for the 'Scientific Programming for QCB' project. The left sidebar contains a search bar, the project name, and navigation links for 'General Info' and 'Table of Contents'. The main content area has a breadcrumb trail 'Docs » Table of Contents » Scientific Programming for QCB' and an 'Edit on GitHub' link. Below this is the title 'Scientific Programming for QCB' and download links for 'PDF', 'EPUB', and 'HTML'. The 'General Info' section includes a link to contact information. The 'Timetable and lecture rooms' section describes the course schedule. The 'Slides' section mentions their availability. The 'Acknowledgements' section thanks Dr. David Leoni and Dr. Stefano Teso. A 'Previous' button is at the bottom of the main content area. The footer shows the copyright notice: '© Copyright # 2018 - 2018, Luca Bianco. Revision 0552d0f8.'



luca.bianco@fmach.it



# Timetable

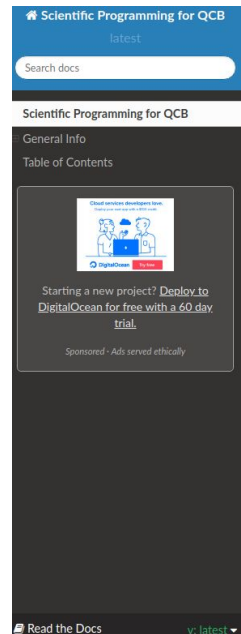
Tuesdays:

A221: 16,30 - 18,30

Fridays:

A223: 14,30 - 16,30

Most probably we will skip October 30th practical.



[Docs](#) » [Table of Contents](#) » Scientific Programming for QCB

[Edit on GitHub](#)

## Scientific Programming for QCB

Download: [PDF](#) [EPUB](#) [HTML](#)

### General Info

The contacts to reach me can be found [at this page](#).

### Timetable and lecture rooms

Practicals will take place on Tuesdays from 16:30 to 18:30 (in lecture room A221) and on Fridays from 14:30 to 16:30 (in lecture room A223). This first part of the course will tentatively run from 28/09/2018 to 21/12/2018.

### Slides

Slides of the practicals will be available on the top part of each practical page.

### Acknowledgements

I would like to thank Dr. David Leoni for all his help and for sharing Jupman with me. I would also like to thank Dr. Stefano Teso for allowing us to use some of his material of a previous course.

[Previous](#)

© Copyright # 2018 - 2018, Luca Bianco. Revision 0552d0f8.



luca.bianco@fmach.it

Please, fill the form at

<https://tinyurl.com/spquest01>

Deadline Sunday, September 30th

## Scientific Programming 2018/19

\*Campo obbligatorio

**Indirizzo email \***

Il tuo indirizzo email

**First name**

La tua risposta

**Family name (surname) \***

La tua risposta

**What operating system do you use on your PC? \***

☐ GNU/Linux

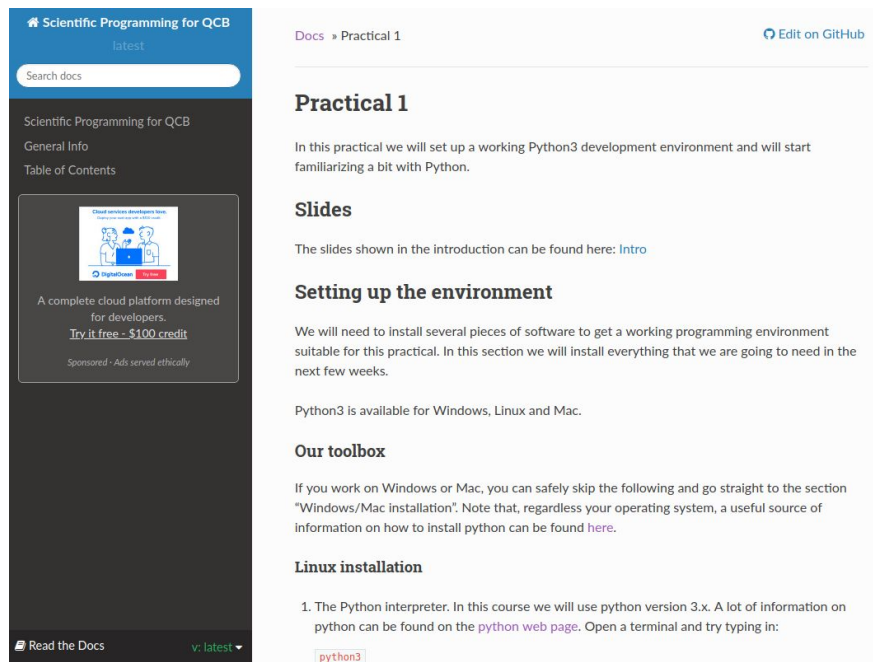
☐ MacOS

☐ Windows 10

# Any questions?

If not, go to:

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



The screenshot shows the 'Scientific Programming for QCB' documentation page on Read the Docs. The left sidebar contains a search bar, navigation links for 'General Info' and 'Table of Contents', and a promotional banner for a cloud development platform. The main content area is titled 'Practical 1' and includes sections for 'Slides', 'Setting up the environment', 'Our toolbox', and 'Linux installation'. The 'Setting up the environment' section mentions installing software for a Python3 development environment. The 'Linux installation' section lists the first step as installing the Python interpreter.

Scientific Programming for QCB  
latest

Search docs

Scientific Programming for QCB  
General Info  
Table of Contents

Cloud enables developers to  
develop code faster and more efficiently.  
Try it free - \$100 credit  
Sponsored - Ads served ethically

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

The slides shown in 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 are going to need in the next few weeks.

Python3 is available for Windows, Linux and Mac.

### Our toolbox

If you work on Windows or Mac, you can safely skip the following and go straight to the section "Windows/Mac installation". Note that, regardless your operating system, a useful source of information on how to install python can be found [here](#).

### Linux installation

1. The Python interpreter. In this course we will use python version 3.x. A lot of information on python can be found on the [python web page](#). Open a terminal and try typing in:

python3



luca.bianco@fmach.it