8 Janvier 2018

M1BI – Outils d'analyse des génomes

Introduction au TP "analyse statistique de données de transcriptome issues de puces d'expression"

Claire Vandiedonck

Associate professor at Paris Diderot

claire.vandiedonck@univ-paris-diderot.fr

University Paris Diderot
School of medicine

INSERM UMR-S-958
Genetics of diabetes

Dataset used in this course

Guida *et al. BMC Genomics* 2011, **12**:628
http://www.biomedcentral.com/1471-2164/12/628

RESEARCH ARTICLE Open Access

Using RNA-seq to determine the transcriptional landscape and the hypoxic response of the pathogenic yeast *Candida parapsilosis*

Alessandro Guida[1], Claudia Lindstädt[2], Sarah L Maguire[2], Chen Ding[2,3], Desmond G Higgins[1], Nicola J Corton[4], Matthew Berriman[4] and Geraldine Butler[2*]

Candida species

The most common cause of fungal infection worldwide

8-10% of hospital-acquired bloodstream

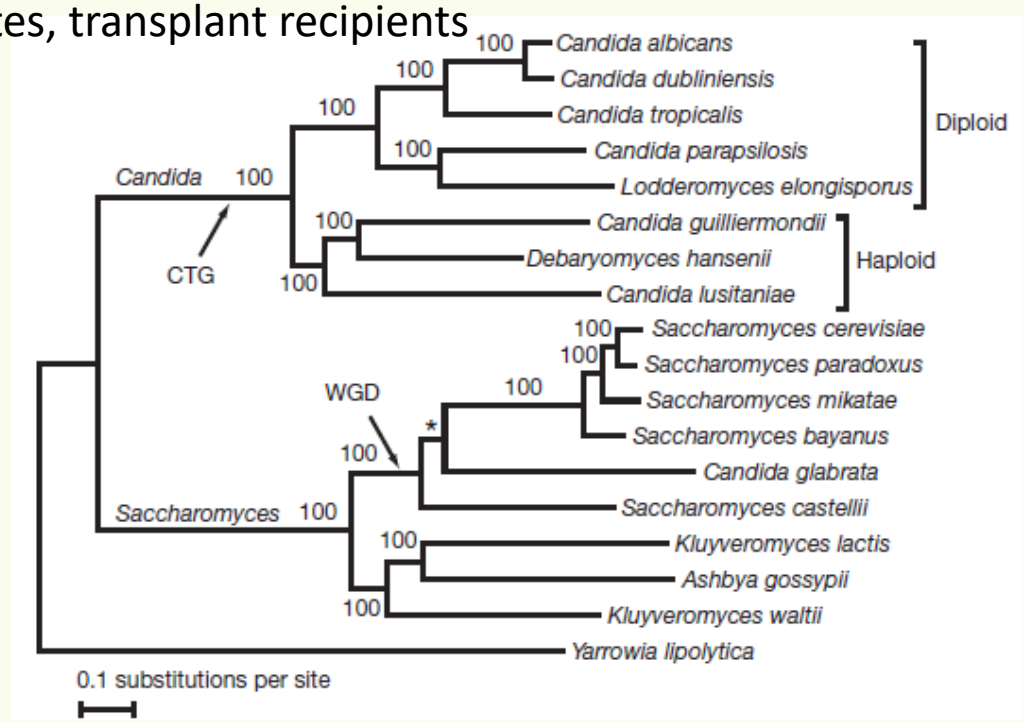
4 Candida species account for ~ 95% of Candida infections

- *C. albicans*: the most common causative agent
- *C. glabrata:* mainly non-pathogenic but high mortality rate when disseminate
- *C. tropicalis:* associated with neutropenia and malignancy
- *C parapsilosis:* health issue for neonates, transplant recipients

Phylogeny

The CUG codon Candida clade
Ser instead of a Leu

Two separate subclades for haploid and diploid



Phenotypic variation

pathogenicity and mating

Genomic features

Variable genomic size : 10.6 to 15.4 Mb

Similar number of protein-coding genes

High conservation of protein-coding genes across Candida CUG clade

Candida parapsilosis

Pathogen responsible for bloodstream infections

specially in < 2 years old children
resistant to antifungal treatment's

Genomic features

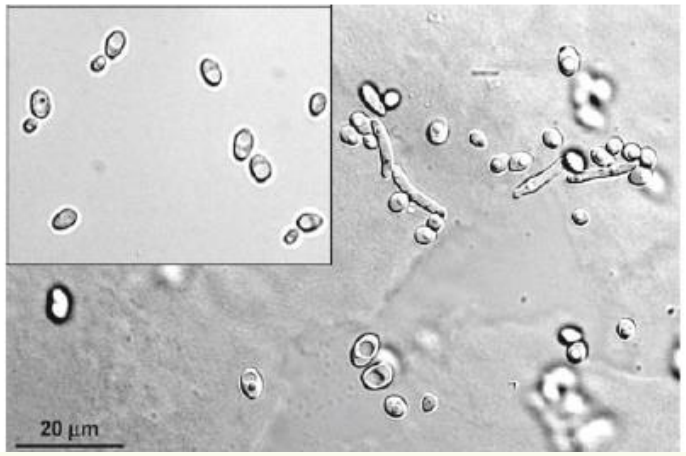
- Diploid
- Genome sequence published in 2009
- Candida databases:
  http://www.broadinstitute.org/annotation/genome/candida_albicans/MultiHome.html
  http://www.candidagenome.org/
- Genome size: 13.1 Mb
- GC content: 38.7%
- Number of genes: 5,733
- Average gene size: 1,533 bp
- Intergenic average size: 752 bp
- 8 chromosomes: 2 annotations (“cpag” and “cpar”)
- No introns annotated

Three groups on the basis of their mitochondrial DNA

- Group I: the “authentic” *C. parapsilosis*
- Groups II and III: *C. orthopsilosis and C. metapsilosis*

Pathogenicity of *Candida* parapsilosis

Polymorphic growth
- yeast growth (blastospores)
- filamentous growth (pseudo-hyphae) when invading host tissues



Pathogenicity facilitated by the ability to:

1. Adhere to medical devices and/or host cells
- cell hydrophobicity
- presence of cell-wall proteins like adhesins
- hydrolytic enzymes secretion: proteases, phospholipases and haemolysins

2. Form biofilm
- Discontinuous monolayer or multilayer
- High level of carbohydrates and low levels of proteins
- Depends on the medium, pH
- **In conditions with low oxygen levels = hypoxia**

The biological question

Determine how the transcriptome of *Candida* parapsilosis is modified under hypoxia conditions:

- to better explain its pathogenicity
- to help identify treatments against this infection

Material and Methods selected for the course

1 WT strain

2 growth conditions in rich media (YDP) at 30°C:

normoxia
oxygen : 21%

versus

hypoxia
oxygen : 1%

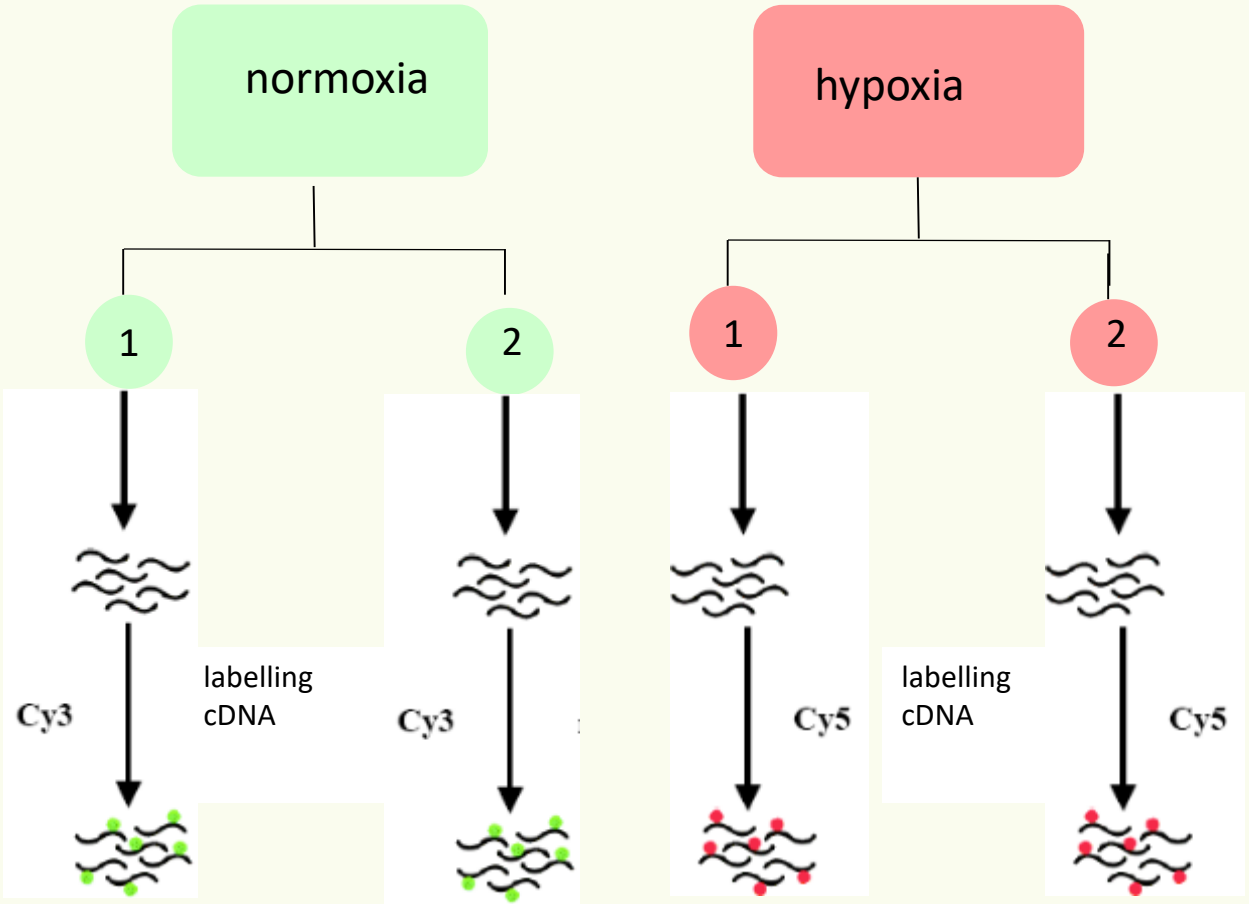
Material and Methods selected for the course

Expression arrays

2 conditions

x 2 biological replicates

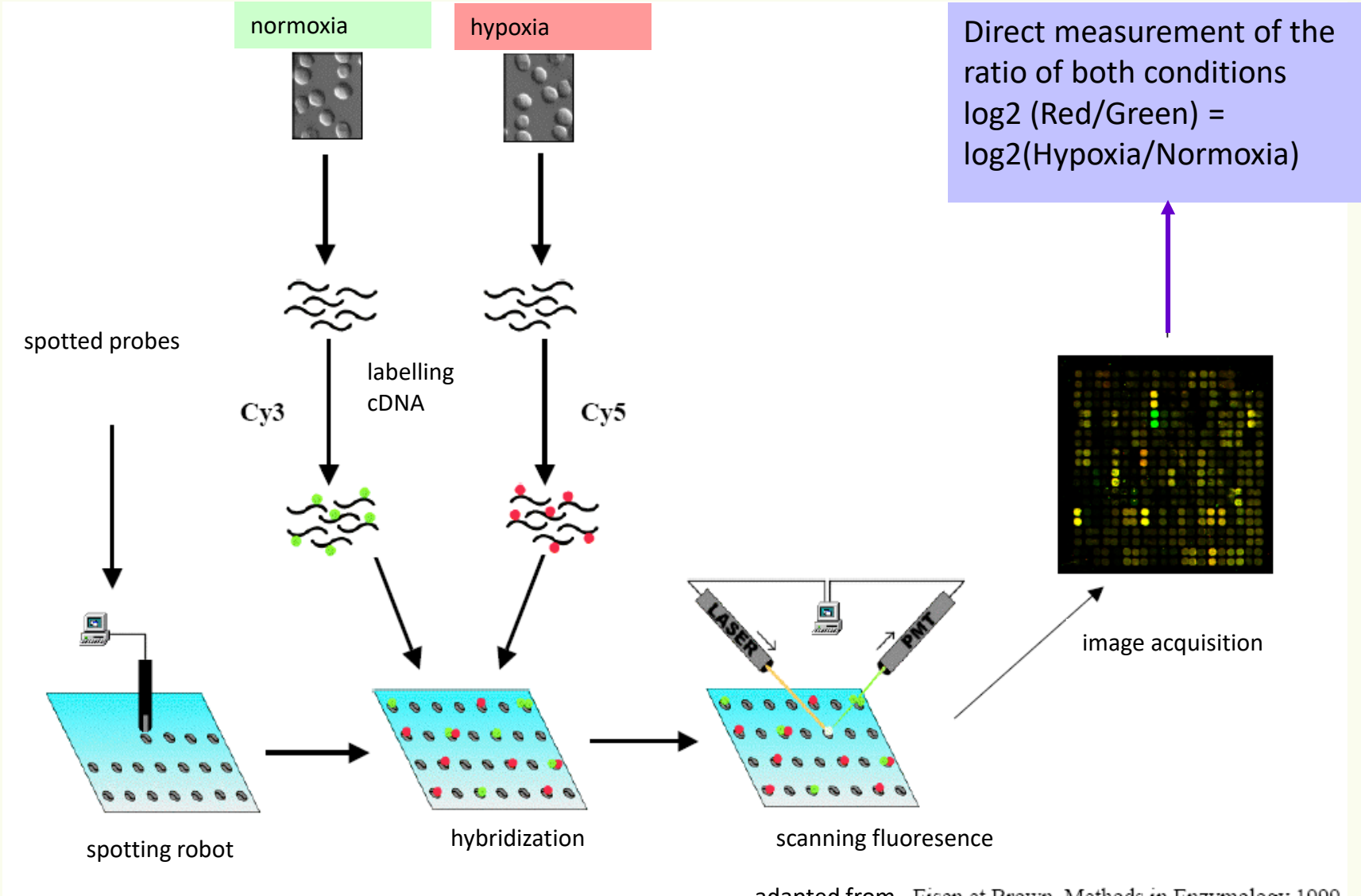
= 4 samples processed



Two-color spotted arrays

two colors = 2 samples are simultaneously hybridized

-> here N1 with H1, N1 with H2, N2 with H1 and N2 with H2 = 4 arrays



adapted from Eisen et Brown, Methods in Enzymology 1999