

# Mineração de parâmetros físicos de DNA/RNA em bases dados biológicas

Gerald Weber

Departamento de Física, UFMG

INCT Sistemas Complexos, Maio 2012

# Bioinformática

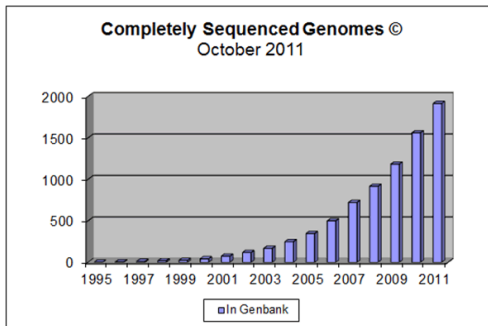
O que é?

## Bioinformática ou Biologia Computacional

É o uso de técnicas de matemática aplicada, física, bioquímica, estatística, engenharia e ciência da computação para resolver problemas em biologia.

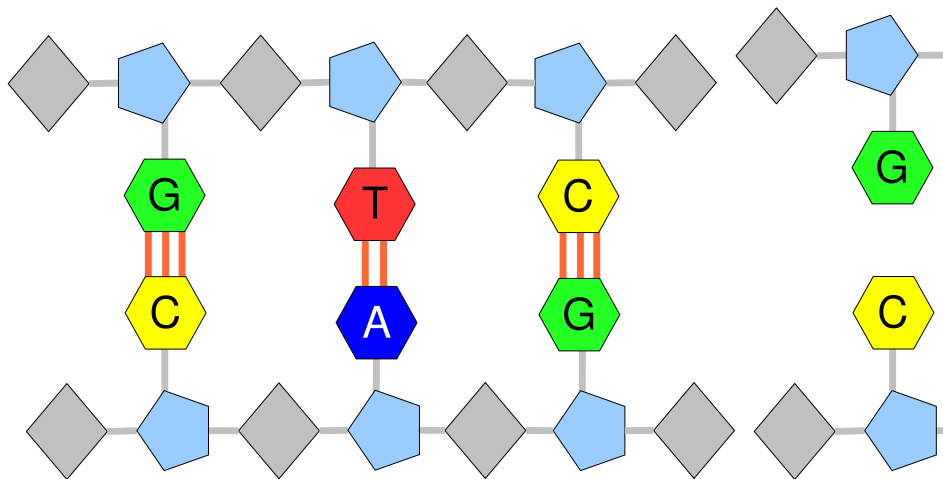
# Genbank

<http://www.genomesonline.org/>



# DNA

## A sequência



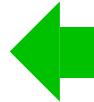
# Sequenciamento tipo *Shotgun*

ATGCTACATGA  
ATGATACATTC  
ATTCTACACTGG



ATGCTACATGATACATTCTACACTGG

Reassemble Fragments



ATGCTACATGA  
CTCCTACATGA  
ATCGTACATGA  
AAGATTACTTA  
CCTTAGAGATT  
TATGGATTAAA  
GATTAGGATTA

Digital Set of Fragments

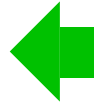
# Sequenciamento tipo *Shotgun*

ATGCTACATGA  
ATGATACATTC  
ATTCTACACTGG



ATGCTACATGATACATTCTACACTGG

Reassemble Fragments

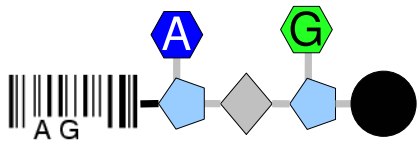
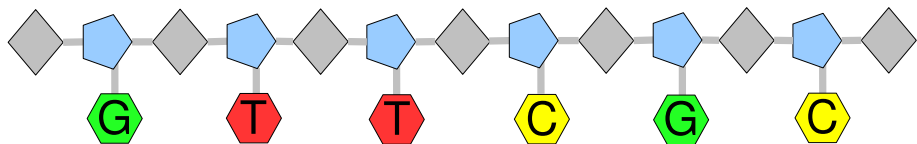


ATGCTACATGA  
CTCCTACATGA  
ATCGTACATGA  
AAGATTACTTA  
CCTTAGAGATT  
TATGGATTAAA  
GATTAGGATTA

Digital Set of Fragments

## DNA

## Sequenciamento



# Fluorescence Resonance Energy Transfer (FRET)

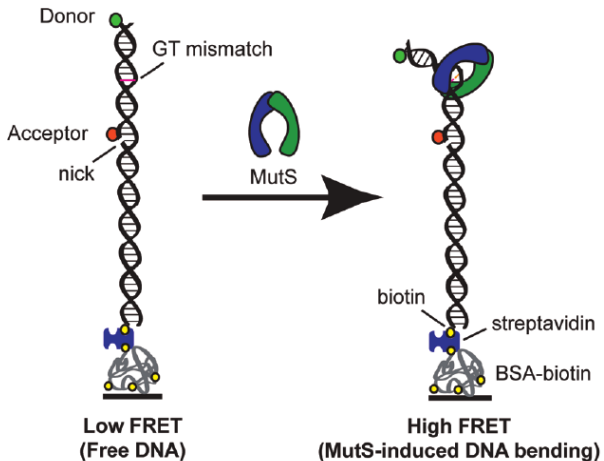
## FRET

Interação dependente de distância entre estados excitados de moléculas de corante. A escala de distância é comparável com as distâncias de interação entre macromoléculas.



# Fluorescence Resonance Energy Transfer (FRET)

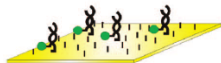
[Sass et al. *Biochem.* 2010]



# MicroArray



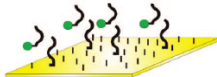
1) Disulphide modified oligonucleotide introduced to surface



4) Labeled target sequence hybridized placing label close to surface



2) Immobilization on surface



5) Temperature/potential ramp applied to determine melting temperature/potential



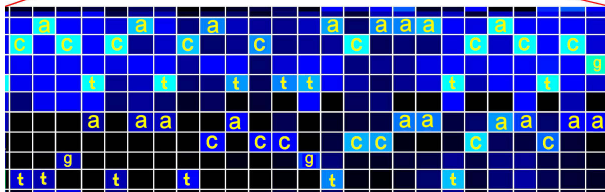
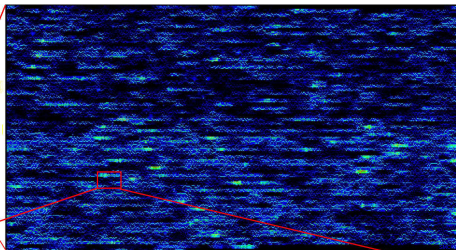
3) Surface treated with mercapto-hexanol to prevent non-specific binding



6) Surface washed to remove target molecules

[Mahajan et al. *J. Am. Chem. Soc.* 2008]

# Chip de DNA



# Mineração de dados

## Conceito



# Mineração de dados

## Proposta

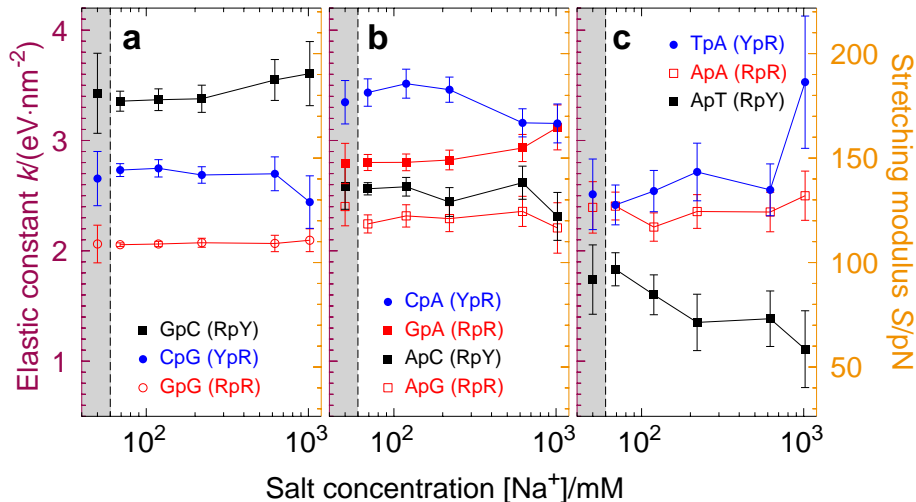
- Dados coletados em microarranjos
  - Stanford Microarray database: 97 mil experimentos depositados
- Processar estes dados para que possam ser analisados
- Processar em função do tipo de sonda, de temperatura, intensidade de luminescência etc
- Correlação flexibilidade de DNA ou RNA, hibridização de bases despareadas

# Constantes el sticas obtidas de temperatura

- Dados de temperatura de denatura o  
[Owczarzy et al. *Biochem.* 2004]
- Modelo para denatura o de DNA, contendo
  - 2 par metros para as pontes de hidrog nio (AT e CG)
  - 10 par metros para as constantes el sticas
- M todo variacional para obten o dos par metros que ajustam as temperaturas

# Constantes elásticas $k$

[Weber et al. *Nature Physics* 2009]



Flexibilidade

$S = 91$  pN dados de raio-x de baixo ângulo

# RNA hydrogen bonds

Morse potential  $D$

base pair	DNA	RNA
CG	70(2) meV	67(4) meV
AT or AU	33(2) meV	39(3) meV

Experimental data from [Xia et al. *Biochem.* 1998]



# RNA hydrogen bonds

## NMR measurements



Published on Web 04/21/2004

### **N1...N3 Hydrogen Bonds of A:U Base Pairs of RNA Are Stronger than Those of A:T Base Pairs of DNA**

Ioannis Vakonakis and Andy C. LiWang\*

*Department of Biochemistry & Biophysics, Texas A&M University, College Station, Texas 77843-2128*

*“Although it is well accepted that hydrogen bonds are central to the structure and function of nucleic acids, it is not yet clear whether significant differences exist between RNA and DNA hydrogen bonds.”*

# Estudantes

Denise Fagundes Lima, MS Biotecnologia

Miriam C. S. Nunes, MS Biotecnologia

Guilherme Bicalho Saturnino, Física, IC

# Colaborações

DF/UFMG Ana Maria de Paula

UFOP Alcides Castro e Silva, Romuel Machado, Renata Guerra-Sá




CEFET-MG Elizabeth F. Wanner

Leeds Sarah A. Harris




Obrigado pela atenção!

[gweberbh@gmail.com](mailto:gweberbh@gmail.com)

## Referências I

-  Mahajan, S., Richardson, J., Brown, T., & Bartlett, P. N. (2008). SERS-melting: A new method for discriminating mutations in DNA sequences. *J. Am. Chem. Soc.*, 130(46), 15589–15601.
-  Mathew-Fenn, R. S., Das, R., & Harbury, P. A. B. (2008). Remeasuring the double helix. *Science*, 322, 446–449.
-  Owczarzy, R., You, Y., Moreira, B. G., Manthey, J. A., Huang, L., Behlke, M. A., & Walder, J. A. (2004). Effects of sodium ions on DNA duplex oligomers: Improved predictions of melting temperatures. *Biochem.*, 43, 3537–3554.

## Referências II

-  Sass, L. E., Lanyi, C., Weninger, K., & Erie, D. A. (2010). Single-molecule FRET TACKLE reveals highly dynamic mismatched DNA- MutS complexes. *Biochem.*, 49(14), 3174–3190.
-  Weber, G., Essex, J. W., & Neylon, C. (2009). Probing the microscopic flexibility of DNA from melting temperatures. *Nature Physics*, 5, 769–773.
-  Xia, T., SantaLucia, Jr., J., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., & Turner, D. H. (1998). Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochem.*, 37, 14719–14735.