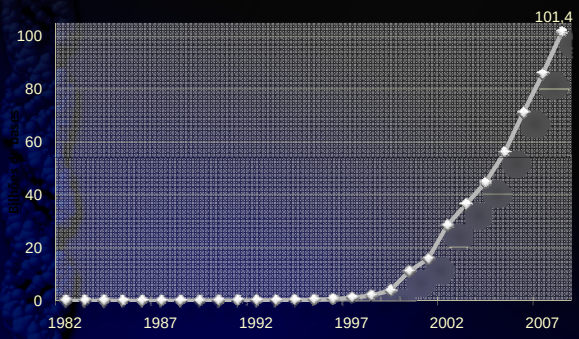
Projetos Genoma

MO640A - Biologia Computacional

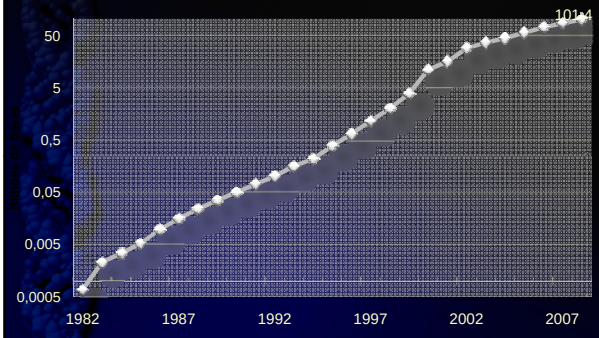
Felipe Rodrigues da Silva

Embrapa Recursos Genéticos e Biotecnologia

GenBank 2008

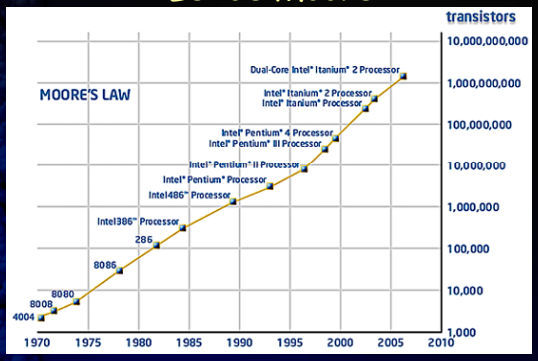


GenBank 2008



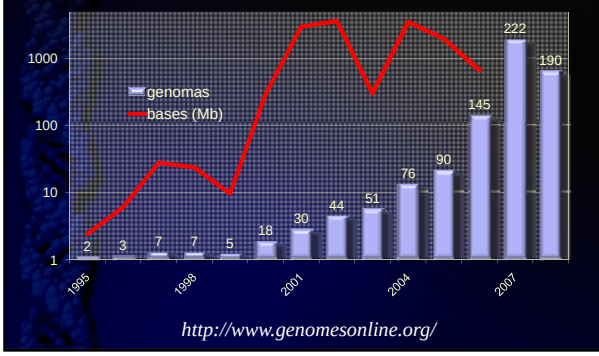
ftp://ftp.ncbi.nih.gov/genbank/release.notes/

Lei de Moore



http://www.intel.com/technology/mooreslaw/index.htm

Genomas Publicados



http://www.genomesonline.org/

Genomas Completos (jan 2009)

916

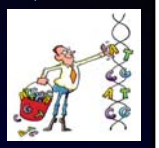
- Procariotos **761+55**
  - *Nanoarchaeum* 490 Kb*
  - Mycoplasma 580 Kb
  - *Bradyrhizobium* 9,105 Kb
- Eucariotos **100**

Genomas em Andamento

(jan 2009)

3.317

- **Procariotos** **2.306**
- **Eucariotos** **1.011**
  - Protozoa (*Trypanosoma cruzi*)
  - Algae (*Chlamydomonas*)
  - Fungy (*Candida albicans*)
  - Nematodes (*Ascaris suum*)
  - Plants (Maize, Wheat , Tomato, Cotto, Soy beam)
  - Insects (*Aedes egypti, Apis mellifera*)
  - Amphibious (*Xenopus*)
  - Birds (chicken)
  - Mammals (cow, dog, pig)



Mais 304 ESTs, 75 RSTs
e 137 Metagenomas…...

Projeto Genoma

Estrutural

Sequenciamento Completo do Genoma

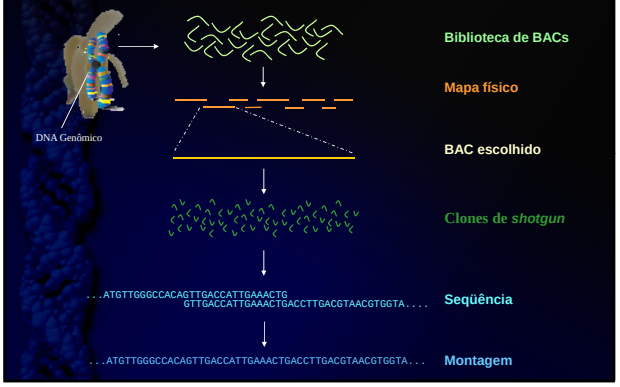
– Região Gênica e Região Intergênica

Funcional

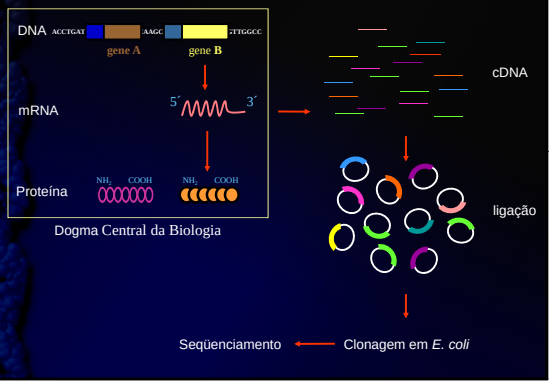
EST – *Expressed Sequence Tag*

–Regiões que codificam proteínas (Genes)

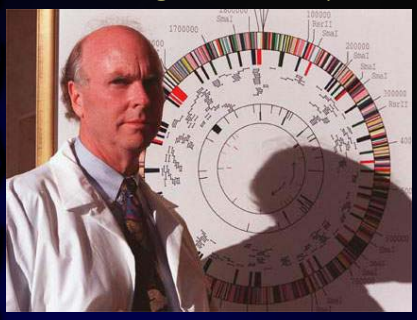
Seqüenciamento Completo



EST - *Expressed Sequence Tag*

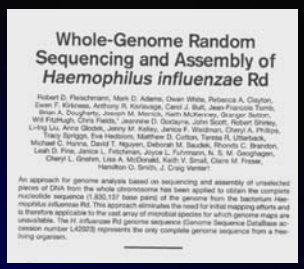


1995: TIGR
Primeiro genoma completo



1995: TIGR
Primeiro genoma completo

Bactéria
1,6 Mb - 1.700 genes

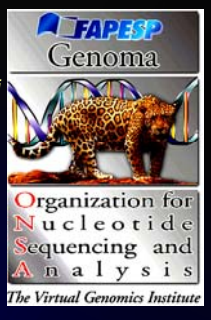


28 de Julho de 1995

Science 269:496

A Rede ONSA

- *Xylella fastidiosa*
- *Xanthomonas axonopodis pv citri*
- Sugar Cane EST
- Human Cancer EST
- *Xanthomonas campestris*
- AEG
  - *X. fastidiosa* / Pierce's Disease
  - *Leifsonia xyli subsp. Xyli*
  - *eucalipto*



FAPESP

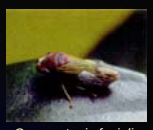
Genoma Xylella

CVC
clorose variegada dos citros

Alguns vetores de *X. fastidiosa*


Dilobopterus costalimai


Acrogonia terminalis

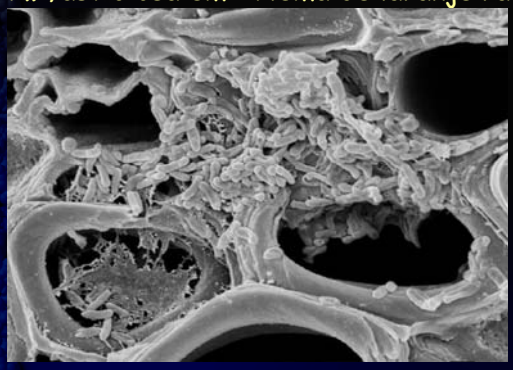

Oncometopia facialis


Homolodisca ignorata


Bucephalogonia xanthopis


Macugonalia sp

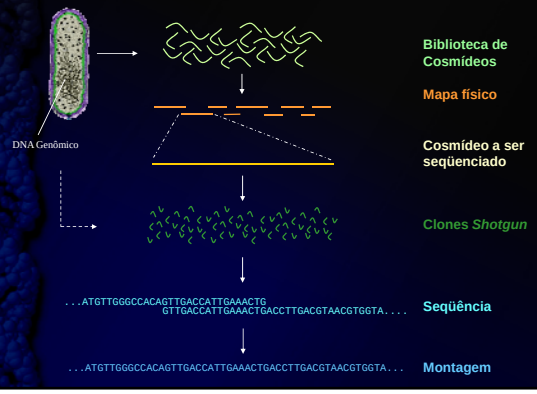
X. fastidiosa em xilema de laranjeira



Genomas

- Seqüenciamento
- Montagem
- Análise
- Anotação
- Pós-genoma

Seqüenciamento Completo



Estratégia de seqüenciamento

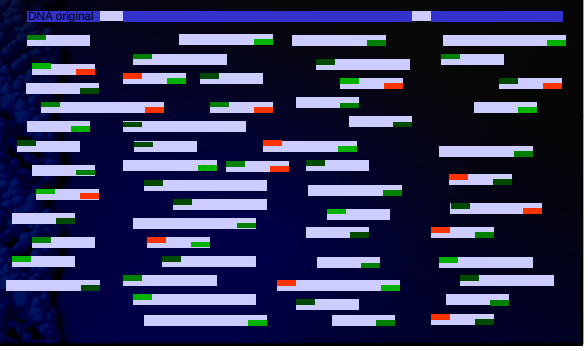
- Pequena escala
  - Leitura única
  - Deleções sucessivas
  - Subclonagem
  - *Primer walking*
  - *shotgun*

Shotgun

- Amostrar fragmentos da seqüência-alvo da maneira mais aleatória possível.
- Determinar a maior porção possível das seqüências das extremidades destes fragmentos

Sanger F, Coulson AR, Hong GF, Hill DF, Petersen GB. (1982) **Nucleotide sequence of bacteriophage lambda DNA**. *J Mol Biol* 162(4): 729-73.

Montagem shotgun



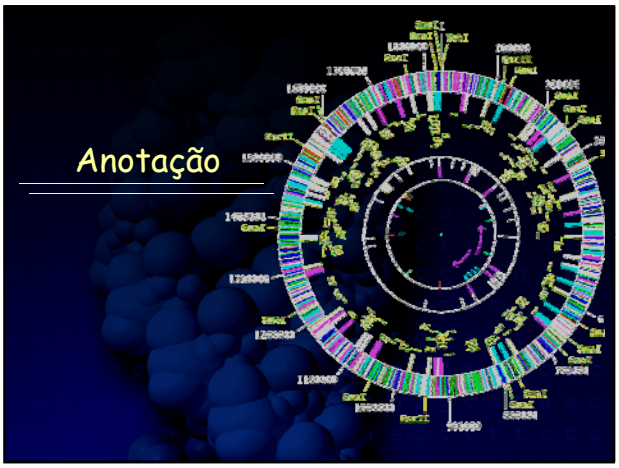
Montagem

- *Trimming* (corte)
  - fixo
  - por ambiguidade
  - por qualidade
- Consenso
  - inclusivo (código de ambiguidades)
  - por freqüencia
  - por qualidade

Montagem com *phrap*

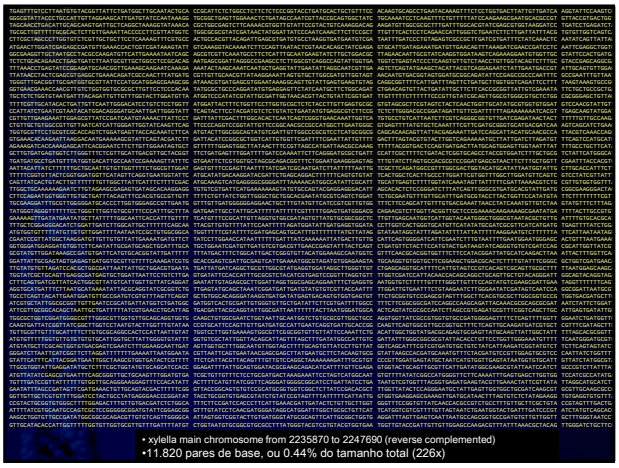
Análise

- Reads
  - sem dados
  - ruins
  - errados
  - vetor
  - hospedeiro
- Montagem
  - gaps
  - conflitos
  - quimeras/deleções
- Checagem de co-linearidade



Anotação

- ORFs
  - localização
  - identificação
- RNAs
- Seqüências regulatórias



- xylella main chromosome from 2235870 to 2247690 (reverse complemented)
- 11.820 pares de base, ou 0.44% do tamanho total (226x)

Predição de genes

- Esta seqüência é codificadora?
- Como o gene X está organizado?

Anotação

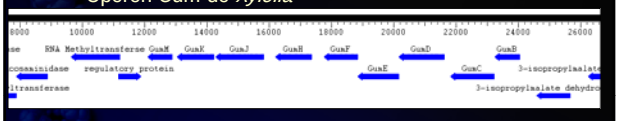
- Que vias metabólicas estão presentes neste organismo?
- O que ele é capaz de fazer?
- O que ele <u>não</u> é capaz de fazer?

Predição de genes é difícil

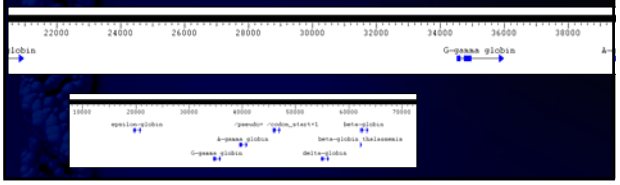
- Sinais nas seqüências de DNA tem baixo conteúdo de informação
  - Degenerado, altamente inespecífico
- Discriminção dos sinais verdadeiros é complicado
- Erros de seqüenciamento

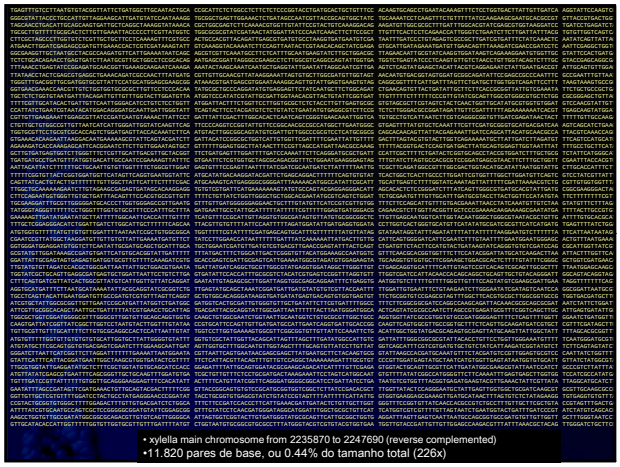
Organização dos genomas

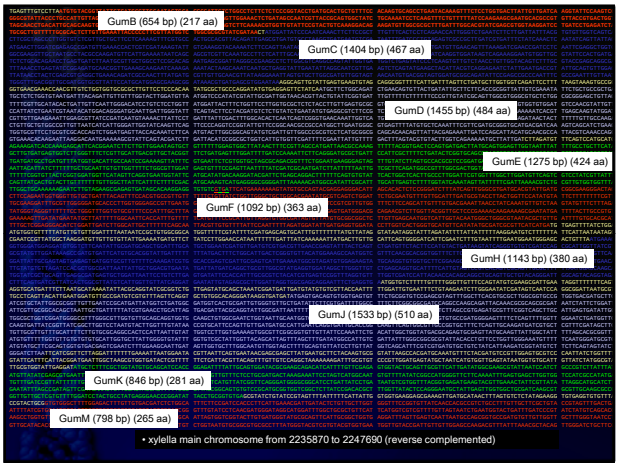
- Operon Gum de *Xylella*



- Região da beta globina humana no cromossomo 11







O genoma publicado de *Xylella fastidiosa*

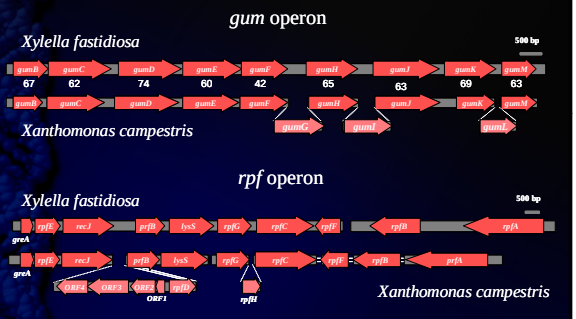
- Mar/98 – Jan/00
- 2,7 Mbp
- 2,904 ORFs

Nature, (July 13, 2000) 406:151-157.

Grupo do genoma da *Xylella*

http://aeg.lbi.ic.unicamp.br/xf/

Mapa genético dos operons *gum* e *rpf*

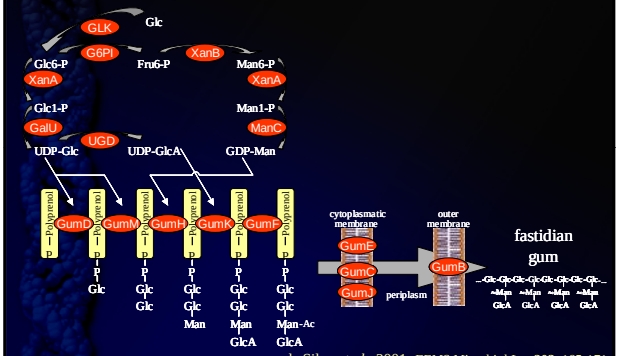


Via sintética da goma xantana

Katzen *et al.*, 1998. *J Bacteriol* **180** (7): 1607-1617

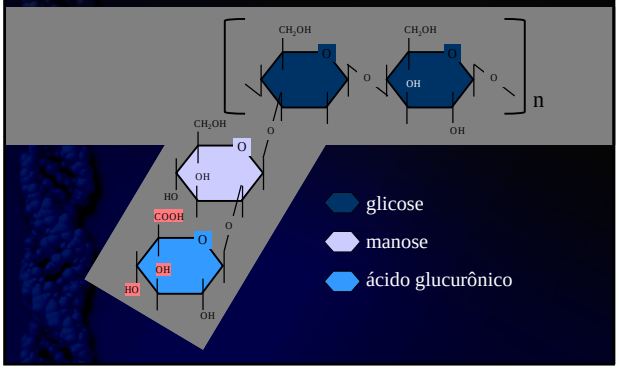
classe	gene ID	gene homologo	identitdade	função
Regulador	XF0287	*rpfB*	72.3	regulatory protein (DSF)
Regulador	XF0290	*rpfA*	80.0	aconitase
Regulador	XF1109	*rpfE*	65.2	regulatory protein
Regulador	XF1113	*rpfG*	77.0	two-component system, regulatory protein
Regulador	XF1114	*rpfC*	60.0	fused two-component sensor-regulator protein
Regulador	XF1115	*rpfF*	65.7	regulatory protein (DSF)
Precursor	XF0232	*pgi*	79.1	glucose-6-phosphate isomerase
Precursor	XF0259	*xanB*	84.5	phosphomannose isomerase-GDP-mannose pyrophosphorylase
Precursor	XF0260	*xanA*	84.8	phosphoglucomutase /phosphomannomutase
Precursor	XF1064	*glk*	41.4	glucose kinase
Precursor	XF1460	*glk*	32.7	glucose kinase
Precursor	XF1606	*algD*	66.1	UDP-glucose dehydrogenase
Precursor	XF2432	*gtaB*	81.8	UTP-glucose-1-phosphate uridylyltransferase
EPS-sint	XF2360	*gumM*	63.1	GumM protein
EPS-sint	XF2361	*gumK*	68.7	GumK protein
EPS-sint	XF2364	*gumH*	64.7	GumH protein
EPS-sint	XF2365	*gumF*	41.9	GumF protein
EPS-sint	XF2367	*gumD*	73.6	GumD protein
EPS-exp	XF2362	*gumJ*	62.7	GumJ protein
EPS-exp	XF2366	*gumE*	59.9	GumE protein
EPS-exp	XF2369	*gumC*	61.2	GumC protein
EPS-exp	XF2370	*gumB*	67.1	GumB protein

Via sintética da goma fastidiana

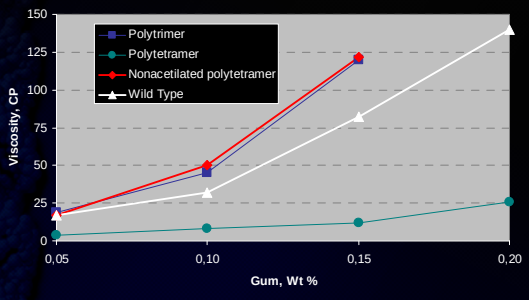


da Silva *et al.*, 2001. *FEMS Microbiol.Let.* **203**: 165-171

Possível estrutura da goma fastidiana



Viscosidades de polímeros derivados da goma xantana



Hassler & Doherty, 1990. *Biotechnol. Prog.* **6** (3): 182-187

Grupo Genoma - CBMEG

http://www.cenargen.embrapa.br/~felipes/

felipes@cenargen.embrapa.br