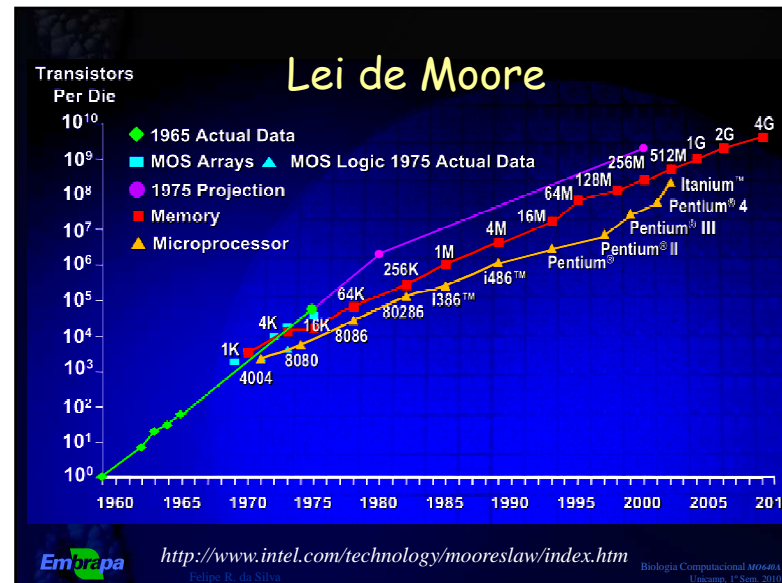
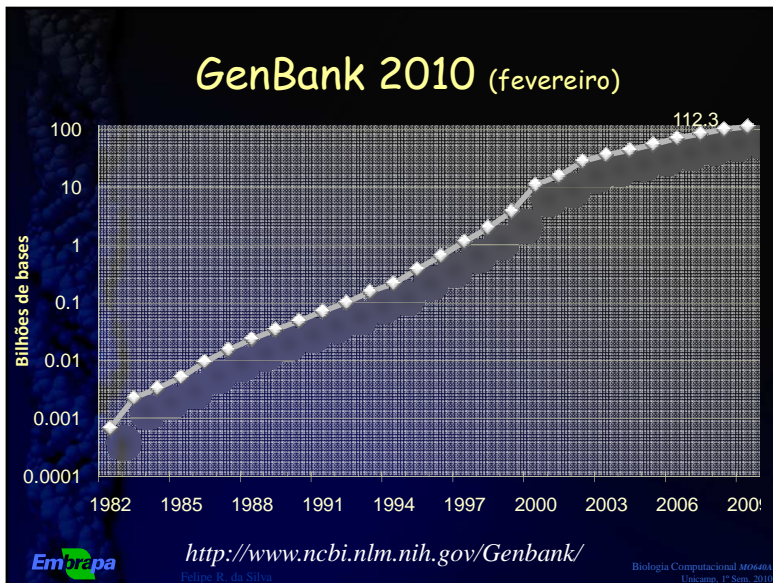
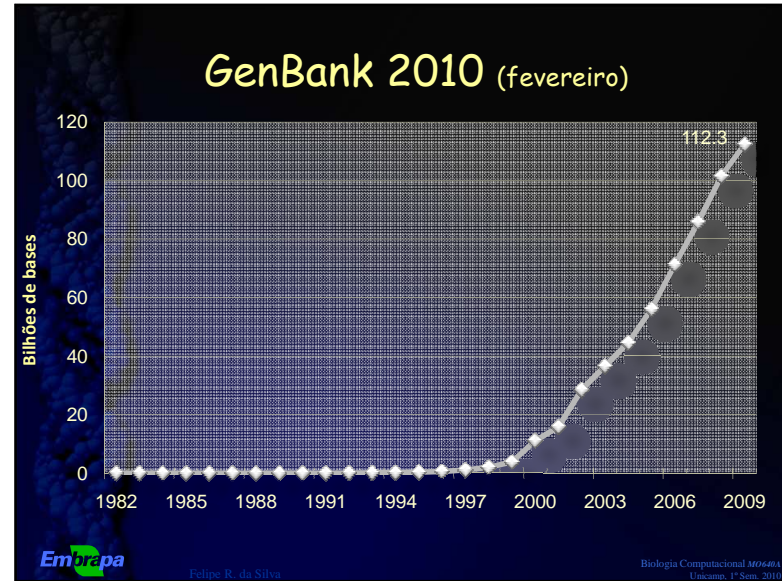


# Projetos Genoma

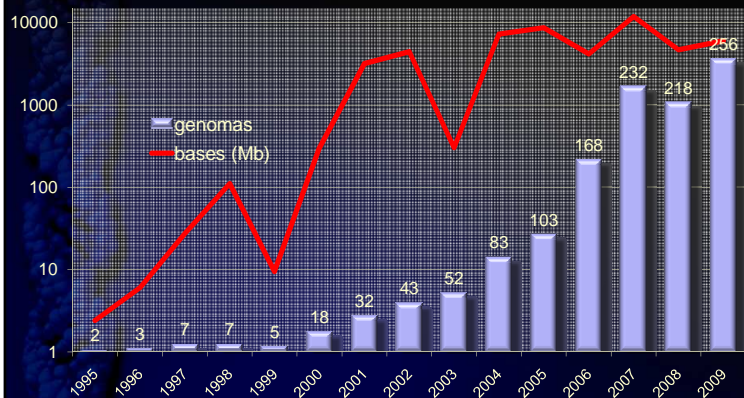
MO640A - Biologia Computacional

Felipe Rodrigues da Silva

Embrapa Recursos Genéticos e Biotecnologia

## Genomas Completos Publicados



Embrapa

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

Biologia Computacional M06604  
Unicamp, 1º Sem. 2010

## Genomas Completos (fev 2010)

1.214

• Procariotos 1.010 + 80

- *Hodgkinia cicadicola* 143 Kb\*
- *Mycoplasma genitalium* 580 Kb
- *Sorangium cellulosum* 13.034 Kb

• Eucariotos 124



Embrapa

Felipe R. da Silva

Biologia Computacional M06604  
Unicamp, 1º Sem. 2010

## Genomas em Andamento

(fev 2010)

5.488

- Procariotos 3.963 + 202
- Eucariotos 1.323
  - Protozoa (*Trypanosoma cruzi*)
  - Algae (*Chlamydomonas*)
  - Fungy (*Candida albicans*)
  - Nematodes (*Ascaris suum*)
  - Plants (Maize, Wheat, Tomato, Cotto, Soy beam)
  - Insects (*Aedes aegypti*, *Apis mellifera*)
  - Amphibious (*Xenopus*)
  - Birds (chicken)
  - Mammals (cow, dog, pig)



Mais 342 ESTs e 149 Metagenomas.....

Embrapa

Felipe R. da Silva

Biologia Computacional M06604  
Unicamp, 1º Sem. 2010



Embrapa

Felipe R. da Silva

Biologia Computacional M06604  
Unicamp, 1º Sem. 2010



Embrapa

Felipe R. da Silva

Biologia Computacional M06694  
Unicamp, 1º Sem. 2010



Embrapa

Felipe R. da Silva

Biologia Computacional M06694  
Unicamp, 1º Sem. 2010

## Projeto Genoma

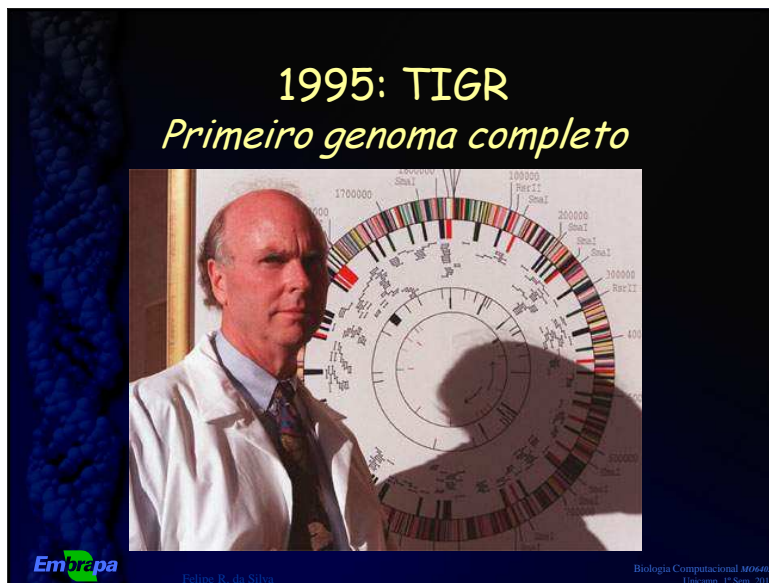
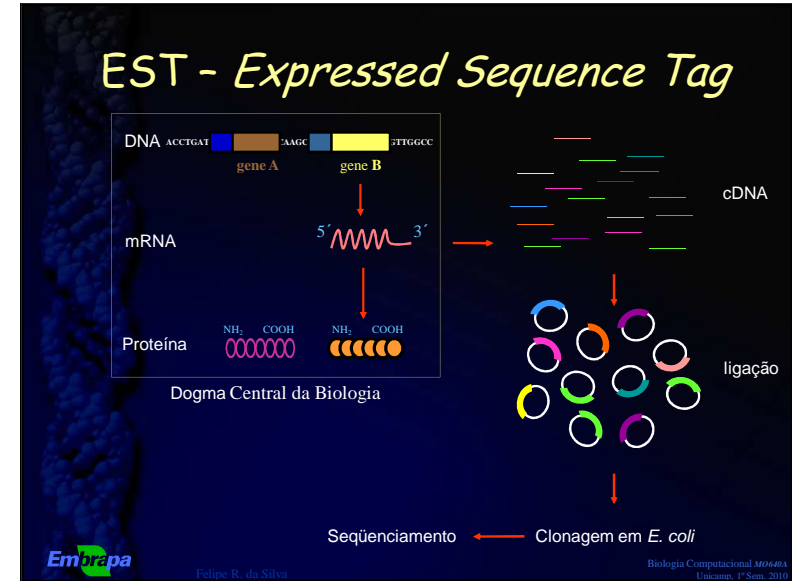
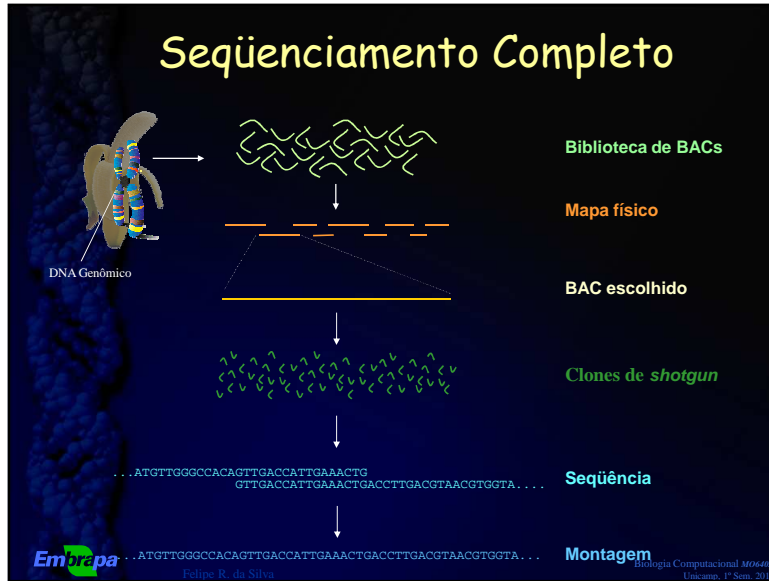
**Estrutural**  
**Sequenciamento Completo do Genoma**  
– Região Gênica e Região Intergênica

**Funcional**  
**EST – Expressed Sequence Tag**  
– Regiões que codificam proteínas (Genes)

Embrapa

Felipe R. da Silva

Biologia Computacional M06694  
Unicamp, 1º Sem. 2010



## 1995: TIGR Primeiro genoma completo

**Whole-Genome Random Sequencing and Assembly of *Haemophilus influenzae* Rd**

Robert D. Fleischmann, Mark D. Adams, Owen White, Rebecca A. Clayton, Even F. Kirovski, Anthony R. Kerlavage, Carl J. Sali, Jean-François Tomb, Brian A. Dougherty, Joseph M. Merrick, Keith McKenney, Granger Sutton, Will Fitzhugh, Chris Fields, Jeanine D. Godoyne, John Scott, Robert Shirley, Li-Hing Liu, Amy Glöck, Jeremy M. Kelly, Janice F. Wiedman, Cheryl A. Phillips, Tracy Spriggs, Eva Hedstrom, Matthew D. Cotton, Teresa R. Urtaback, Michael C. Harris, David T. Nguyen, Deborah M. Soudik, Ronald C. Branton, Leah D. Fine, Janice L. Fritchman, Joyce L. Fuhrmann, K. S. M. Geochagan, Cheryl L. Gnehm, Lisa A. McDonald, Keith V. Small, Claire M. Fraser, Hamilton O. Smith, J. Craig Venter

An approach for genome analysis based on sequencing and assembly of unselected pieces of DNA from the whole chromosome has been applied to obtain the complete nucleotide sequence (1,230,137 base pairs) of the genome from the bacterium *Haemophilus influenzae* Rd. This approach eliminates the need for initial mapping efforts and is therefore applicable to the vast array of microbial species for which genome maps are unavailable. The *H. influenzae* Rd genome sequence (Genome Sequence Database accession number L42023) represents the only complete genome sequence from a free-living organism.

**28 de Julho de 1995**

**Bactéria**  
1,6 Mb - 1.700 genes

**Embrapa** Science 269:496  
Biologia Computacional M06494  
Unicamp, 1º Sem. 2010

<p><b>1997</b> Eukaryote 13 Mb - 6.000 genes [Nature 387:1]</p> 	<p><b>1998</b> Animal 100 Mb - 18.000 genes [Science 282:1945]</p> 	<p><b>2000</b> Insect 130 Mb - 13.000 genes [Science 287:2185]</p> 	<p><b>2000</b> Plant 120 Mb - 26.000 genes [Nature 408:791]</p> 	<p><b>2000</b> Bacteria - Plant pathogen 2,6 Mb - 3.000 genes [Nature 406:151]</p> 
<p><b>2001</b> Human 3,2 Gb - 40.000 genes [Science 291:1304]</p> 	<p><b>2001</b> Human 3,2 Gb - 40.000 genes [Nature 409:745]</p> 	<p><b>2002</b> Plant 420 Mb - 35.000 genes [Science 296:79]</p> 	<p><b>2002</b> Parasite - Host 23 Mb - 5.300 genes [Nature 419:498] 278Mb - 14.000 genes [Science 298:129]</p> 	<p><b>2002</b> Mouse 2,5 Gb - 30.000 genes [Nature 420:520]</p> 

Embrapa Felipe R. da Silva Biologia Computacional M06604  
Unicamp, 1º Sem, 2010

## 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




Organization for  
**N**ucleotide  
**S**equencing and  
**A**nalysis  
*The Virtual Genomics Institute*

Embrapa Felipe R. da Silva Biologia Computacional M06604  
Unicamp, 1º Sem, 2010




# Genoma Xylella

Embrapa Felipe R. da Silva Biologia Computacional M06604  
Unicamp, 1º Sem, 2010

## CVC clorose variegada dos citros





Embrapa Felipe R. da Silva Biologia Computacional M06604  
Unicamp, 1º Sem, 2010

## Alguns vetores de *X. fastidiosa*



*Dilobopterus costalimai*



*Acrogonia terminalis*



*Oncometopia facialis*



*Homolodisca ignorata*

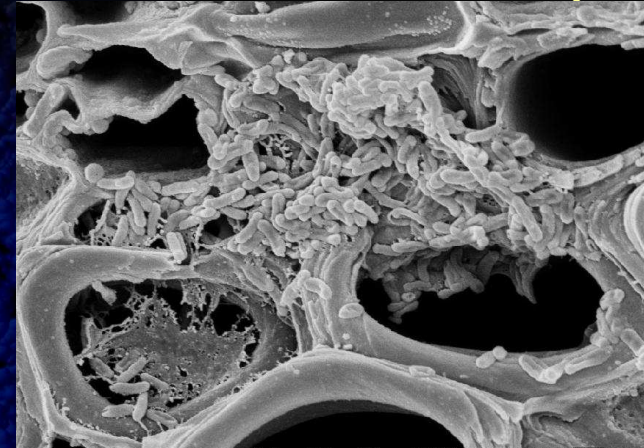


*Bucephalogonia xanthopis*



*Macugonalla sp*

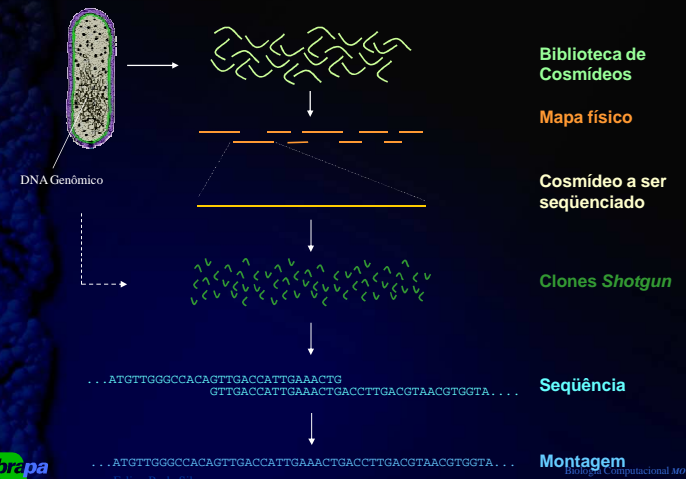
## *X. fastidiosa* em xilema de laranja



## 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.

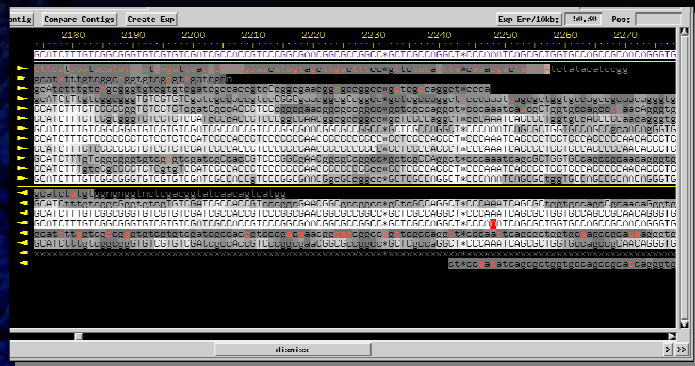
## Montagem shotgun



## Montagem

- *Trimming* (corte)
  - fixo
  - por ambigüidade
  - por qualidade
- *Consenso*
  - inclusivo (código de ambigüidades)
  - por freqüência
  - por qualidade

## Montagem com phrap

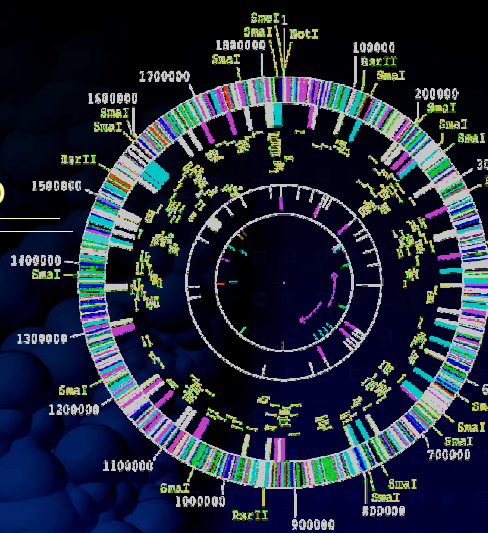


The screenshot shows the phrap software interface with a window titled 'Compare Contigs'. The main area displays a grid of sequence alignments with various colored markers indicating overlaps and gaps. The status bar at the bottom indicates '100%' completion.

## 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)

Embrapa  
Felipe R. da Silva  
Biologia Computacional M06404  
Unicamp, 1º Sem. 2010

# Predição de genes

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

Embrapa  
Felipe R. da Silva  
Biologia Computacional M06404  
Unicamp, 1º Sem. 2010

# Anotação

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

Embrapa  
Felipe R. da Silva  
Biologia Computacional M06404  
Unicamp, 1º Sem. 2010

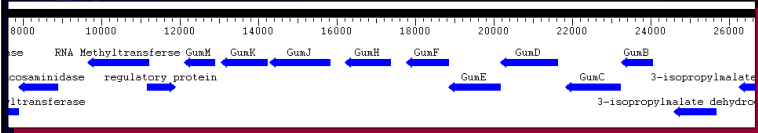
# Predição de genes é difícil

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

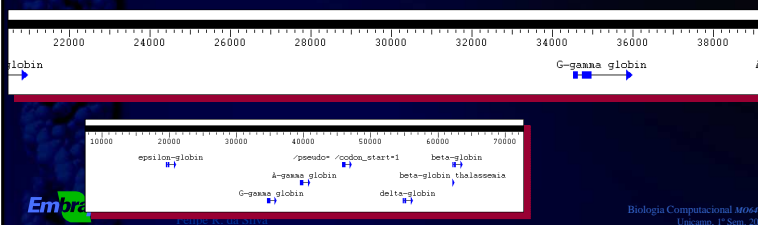
Embrapa  
Felipe R. da Silva  
Biologia Computacional M06404  
Unicamp, 1º Sem. 2010

# Organização dos genomas

- Operon Gum de *Xylella*

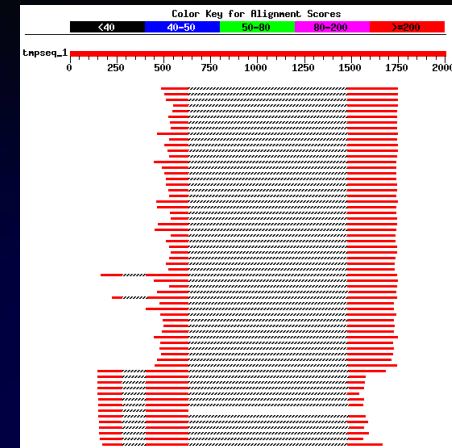


- Região da beta globina humana no cromossomo 11



Biologia Computacional M06604  
Unicamp, 1º Sem. 2010

# Blastn β-globina (62000 to 64000)



Embrapa

Felipe R. da Silva

Biologia Computacional M06604  
Unicamp, 1º Sem. 2010

# Blastn β-globina (62000 to 64000)

Sequences producing significant alignments:

	Score	E
	(bits)	Value
emb F24593.1 F24593 HSPD11021 HM3 Homo sapiens cDNA clone s...	523	e-146
emb F26056.1 F26056 HSPD13352 HM3 Homo sapiens cDNA clone s...	523	e-146
gb T41162 T41162 ya31f08.s3 Homo sapiens cDNA clone 62247 3...	523	e-146
emb F27352.1 F27352 HSPD15247 HM3 Homo sapiens cDNA clone s...	523	e-146
emb F31173.1 F31173 HSPD26430 HM3 Homo sapiens cDNA clone s...	521	e-145
emb F21808.1 F21808 HSPD06401 HM3 Homo sapiens cDNA clone 0...	519	e-145
emb F21643.1 F21643 HSPD06227 HM3 Homo sapiens cDNA clone 0...	519	e-145
emb F20795.1 F20795 HSPD05201 HM3 Homo sapiens cDNA clone 0...	519	e-145
emb F31552.1 F31552 HSPD22843 HM3 Homo sapiens cDNA clone s...	517	e-144
gb R06757 R06757 yf11d02.s1 Homo sapiens cDNA clone 126531 ...	517	e-144
emb F26092.1 F26092 HSPD13402 HM3 Homo sapiens cDNA clone s...	515	e-143
emb F28594.1 F28594 HSPD17840 HM3 Homo sapiens cDNA clone s...	515	e-143
emb F24982.1 F24982 HSPD11692 HM3 Homo sapiens cDNA clone s...	515	e-143
emb F37427.1 F37427 HSPD36142 HM3 Homo sapiens cDNA clone s...	513	e-143
gb R91762 R91762 yp98b11.s1 Homo sapiens cDNA clone 195453 ...	513	e-143
gb AA296097 AA296097 EST10430 Adipose tissue, white I Homo ...	513	e-143
gb R93060 R93060 yq13c07.s1 Homo sapiens cDNA clone 196812 ...	513	e-143
gb R91514 R91514 yq09c11.s1 Homo sapiens cDNA clone 196436 ...	513	e-143
emb F31377.1 F31377 HSPD22507 HM3 Homo sapiens cDNA clone s...	513	e-143
emb F36284.1 F36284 HSPD33731 HM3 Homo sapiens cDNA clone s...	511	e-142
emb F24336.1 F24336 HSPD10580 HM3 Homo sapiens cDNA clone s...	509	e-142
emb F29589.1 F29589 HSPD19534 HM3 Homo sapiens cDNA clone s...	509	e-142
gb H95645 H95645 yv17b07.s1 Soares fetal liver spleen INFLS...	509	e-142

Embrapa

Felipe R. da Silva

Biologia Computacional M06604  
Unicamp, 1º Sem. 2010

# Blastn β-globina (62000 to 64000)

emb|F26056.1|F26056 HSPD13352 HM3 Homo sapiens cDNA clone s4000047F08, mRNA sequence  
Length = 430

Score = 523 bits (264), Expect = e-146  
Identities = 264/264 (100%)  
Strand = Plus / Plus

Query: 1482 gctcctgggcaacgctgctgtgtgtgtgctggccatcactttggcaagaattcaccoc 1541  
Sbjct: 155 gctcctgggcaacgctgctgtgtgtgtgctggccatcactttggcaagaattcaccoc 214

Query: 1542 accagtcaggctgcctatcagaaaagtggctggctggctgactaatgccctggccacaa 1601  
Sbjct: 215 accagtcaggctgcctatcagaaaagtggctggctggctgactaatgccctggccacaa 274

Query: 1602 gtatcactaagctcgtttcttctgtgtccaatttctattaagggttcccttctccctaa 1661  
Sbjct: 275 gtatcactaagctcgtttcttctgtgtccaatttctattaagggttcccttctccctaa 334

Query: 1662 gtccactactaaactggggatattatgaaggcctgagcatctggattctgcttaatt 1721  
Sbjct: 335 gtccactactaaactggggatattatgaaggcctgagcatctggattctgcttaatt 394

Query: 1722 aaaaaacattattttcattgcaa 1745  
Sbjct: 395 aaaaaacattattttcattgcaa 418

Embrapa

Felipe R. da Silva

Biologia Computacional M06604  
Unicamp, 1º Sem. 2010

## Blastn $\beta$ -globina (62000 to 64000)

Score = 283 bits (143), Expect = 7e-74  
 Identities = 143/143 (100%)  
 Strand = Plus / Plus

Query: 490 accctaaggtgaaggctcatggcaagaagtgcggtgccttagtgatggcctggctc 549  
 |||  
 Sbjct: 13 accctaaggtgaaggctcatggcaagaagtgcggtgccttagtgatggcctggctc 72

Query: 550 acctggacaacctcaagggcacctttgccacaactgagtgagctgcaactgtgacaagctgc 609  
 |||  
 Sbjct: 73 acctggacaacctcaagggcacctttgccacaactgagtgagctgcaactgtgacaagctgc 132

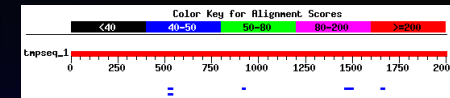
Query: 610 acgtggatcctgagaacttcagg 632  
 |||  
 Sbjct: 133 acgtggatcctgagaacttcagg 155



Felipe R. da Silva

Biologia Computacional M06604  
 Unicamp, 1º Sem. 2010

## Blastn Operon Gum (21000 - 23000)



Sequences producing significant alignments:

Sequence	Score (bits)	E Value
gb U22511 XCU22511 Xanthomonas campestris GumA, GumB, GumC,...	46	0.034
emb Z68013 CEW02H3 Caenorhabditis elegans cosmid W02H3, com...	42	0.54
gb AF025669 AF025669 Tribolium castaneum GABA receptor subu...	40	2.1
emb X00274 HSHL07 Human gene for HLA-DR alpha heavy chain a...	40	2.1
gb AF172352.1 AF172352 Ceratitis capitata GABA receptor sub...	40	2.1

gb|U22511|XCU22511 Xanthomonas campestris GumA, GumB, GumC, GumD, GumE, GumF, GumG, GumH, GumI, GumJ, GumK, GumL, and GumM genes, complete cds.  
 Length = 16075

Score = 46.1 bits (23), Expect = 0.034  
 Identities = 41/47 (87%)  
 Strand = Plus / Minus

Query: 1458 ctgtgaggagccagcaacgcattaagatcactcaggttctgtgcag 1504  
 |||  
 Sbjct: 2779 ctgtgaggatgccagcaacgcattgagatcgtcagattctgcgcag 2733



Felipe R. da Silva

Biologia Computacional M06604  
 Unicamp, 1º Sem. 2010

## Blastn Operon Gum (21000 - 23000)

emb|Z68013|CEW02H3 Caenorhabditis elegans coamid W02H3, complete sequence  
 [Caenorhabditis elegans]  
 Length = 17716

Score = 42.1 bits (21), Expect = 0.54  
 Identities = 21/21 (100%)  
 Strand = Plus / Plus

Query: 1653 acgtgctgattattttcaat 1673  
 |||  
 Sbjct: 9659 acgtgctgattattttcaat 9679

gb|AF025669|AF025669 Tribolium castaneum GABA receptor subunit (Rd1) gene, partial cds  
 Length = 836

Score = 40.1 bits (20), Expect = 2.1  
 Identities = 23/24 (95%)  
 Strand = Plus / Plus

Query: 520 gatacaaaacacgaagattagat 543  
 |||  
 Sbjct: 565 gatacaataacacgaagattagat 588

(...)

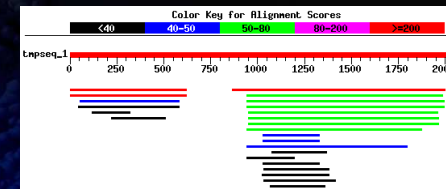
Database: Non-redundant GenBank+EMBL+DDBJ+PDB sequences  
 Posted date: Sep 22, 1999 2:18 PM  
 Number of letters in database: 1,306,157,801  
 Number of sequences in database: 477,403



Felipe R. da Silva

Biologia Computacional M06604  
 Unicamp, 1º Sem. 2010

## Blastx Operon Gum (21000 - 23000)



Sequences producing significant alignments:

Sequence	Score (bits)	E Value
pir  S67819 GumC protein - Xanthomonas campestris >gi 11720...	419	e-116
pir  S67820 GumD protein - Xanthomonas campestris >gi 73314...	263	5e-69
pir  S31721 xps2A protein - Xanthomonas campestris >gi 4854...	258	1e-67
gb AAD43842.1 AF076290.12 (AF076290) Bme12 [Brucella melite...	80	8e-14
sp P33698 EXOP_RHIME_SUCCINOGLYCAN_BIOSYNTHESIS_TRANSPORT_P...	76	9e-13
pir  S37031 exoP protein - Rhizobium meliloti >gi 742331 pr...	68	2e-10
gi 3493609 (AF067140) exopolysaccharide polymerization prot...	68	3e-10
gi 2921558 (AF039306) integral membrane protein [Bradyrhizo...	67	4e-10
gb AAD35728.1 AE001738.8 (AE001738) hypothetical protein [T...	63	6e-09
emb CAA75431  (Y15162) pTK [Acinetobacter johnsonii]	54	4e-06
gb AAD34733.1 AF131869.2 (AF131869) putative undecaprenylph...	43	0.011
gi 3449364 (AF032862) intracellular hyaluronic acid binding...	41	0.024
pir  JC5016 hyaluronan receptor - human	41	0.041
sp Q46631 AMSA_ERWAM_ AMYLOVORAN_BIOSYNTHESIS_PUTATIVE_MEMBR...	40	0.053
gb AAD32394.1 AAD32394 (AF065404) pXO1-90 [Bacillus anthracis]	40	0.070



Felipe R. da Silva

Biologia Computacional M06604  
 Unicamp, 1º Sem. 2010

# Blastx Operon Gum (21000 - 23000)

pir||867819 GumC protein - Xanthomonas campestris sgi|1172092 (U22511) GumC [Xanthomonas campestris]  
Length = 449

Score = 419 bits (1066), Expect = e-116  
Identities = 239/378 (63%), Positives = 310/378 (81%), Gaps = 6/378 (1%)  
Frame = -2

Query: 2000 IPVESPIRDRDFVQYQYQLQSSRLARAVIRKMNLDREPMKPLVKEVLSKVQNL----- 1839  
+EVESP DRDFVQYQYQLQSSRLARAVIR+ LD+EP K V++ L+K  
Sbjct: 58 MPVESPQRDRDFVQYQYQLQSSRLARAVIREAKLDQEPAFKQVEEALAKAEKNPEAGK 117

Query: 1838 FVNTIRTAIESALEMMLNSLIRPILNSRLVVFHDSHDFLSARVANAYAKMFINNQ 1659  
+RR +E +S + +E L +EPILNSRLVVF+ DS DP L+R+AN Y KFF + O  
Sbjct: 118 SLDSRQAVIERSLDTLTLAQLVPEPILNSRLVYVNYSDPPVLAARLANTYKPKVIFVSTQ 177

Query: 1658 QRRSNAPSFMKYLAGRLKQLRKYVDSERNVVAYSTDEKIVSVDKPSLSAQLNLDLN 1479  
+RR A SFA +LA +QLR KV+ SE+++V+YST+P+ IVSVDG+KPSL AQLN+LDN  
Sbjct: 178 ERMKASSFATQFLAERLKQLREKVEDEKDLVSYSTEQIVSVDGKPSLPAQLNLDLN 237

Query: 1478 ALLASAQNERIRAEASWRQASIGDGLSTIPQVLSNLLVQSLRFEQANNVNEYQKLSMFKP 1299  
ALLASAQ+ RI+AE+WRQAS GDG+S+PQVLS+ L+QSLR+EQ + +EYQKLS FKP  
Sbjct: 238 ALLASAQDARIKASAWRQASSGDGMSLPQVLSLPLIQSLRSEQVRLTSEYQKLSFKP 297

Query: 1298 EYFEMQRLKARKENTVQINAVLINRQSLKQVEATLROEMLNDR+AVLKEKDELKLT 1119  
+YFEMQRLKA+ +E +QIN EV+NRQSLK+ Y+A++ OE LINDRIA L +ELD++  
Sbjct: 298 DYPEMQRLKAQIESRRQINGEVINRQSLKATYDASVHQEGLNDR+AVLKEKDELKLT 357

Query: 1118 RLIRVNLLEBSAETTRQLDYALLQRYKEIVSLGDCSSNNVTVYDAADI PRSP FSNPLMN 939  
R IIRVNL+R+ +TRQLDYALLQRYKEI V +VGNVNFVD AD+P+ SD P+ N  
Sbjct: 358 RSIRVNLKRDVDTRNQLDYALLQRYKEIGVASNGANNVTVYDAVPTKSTPKLKN 417

Query: 938 TILGGIPGVGLGLVAMVRYMIG 867  
LG IPGVGLG+ VA+VRY + G  
Sbjct: 418 LALGIPGVGLGVAVLVRYFLRG 441

**Embrapa** Biologia Computacional M06404  
Felipe R. da Silva, Uncamp, 1º Sem. 2010

GumB (654 bp) (217 aa)

GumC (1404 bp) (467 aa)

GumD (1455 bp) (484 aa)

GumE (1275 bp) (424 aa)

GumF (1092 bp) (363 aa)

GumG (1143 bp) (380 aa)

GumH (1533 bp) (510 aa)

GumK (846 bp) (281 aa)

GumM (798 bp) (265 aa)

• xylella main chromosome from 2235870 to 2247690 (reverse complemented)

**Embrapa** Biologia Computacional M06404  
Felipe R. da Silva, Uncamp, 1º Sem. 2010



## O genoma publicado de *Xylella fastidiosa*

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

**Embrapa** Nature, (July 13, 2000) 406:151-157.  
Felipe R. da Silva

**Embrapa** Biologia Computacional M06404  
Uncamp, 1º Sem. 2010

# Grupo do genoma da *Xylella*



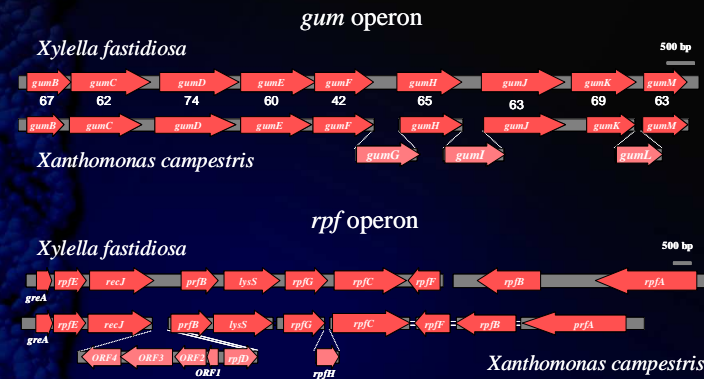
<http://ac>

Embrapa

Felipe R. da Silva

Biologia Computacional M06404  
Unicamp, 1º Sem. 2010

# Mapa genético dos operons *gum* e *rpf*

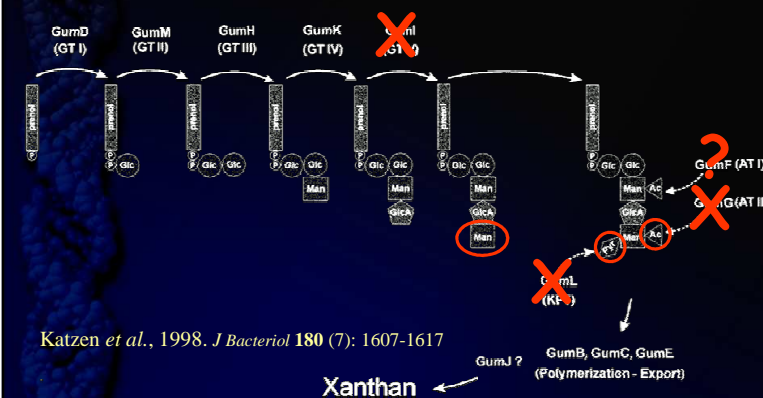


Embrapa

Felipe R. da Silva

Biologia Computacional M06404  
Unicamp, 1º Sem. 2010

# Via sintética da goma xantana



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

Xanthan

GumB, GumC, GumE  
(Polymerization - Export)

Embrapa

Felipe R. da Silva

Biologia Computacional M06404  
Unicamp, 1º Sem. 2010

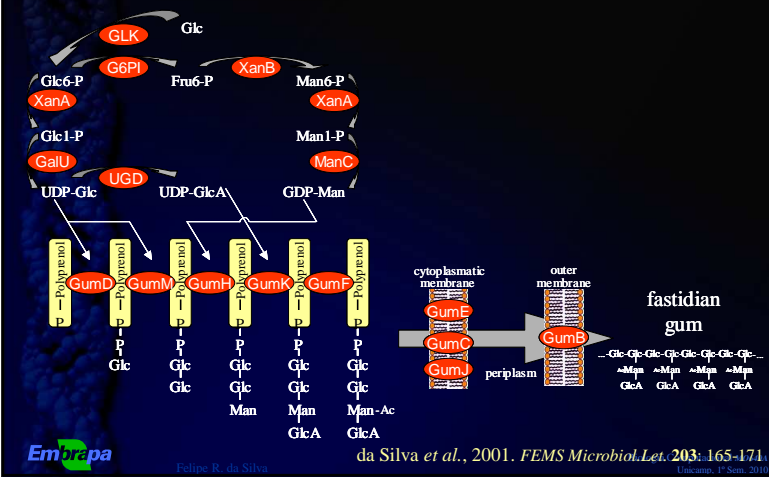
classe	gene ID	gene homologo	identidade	função
Regulador	XF0287	rpfB	72.3	regulatory protein (DSF)
Regulador	XF0290	rpfA	80.0	aconitase
Regulador	XF1109	rpfE	65.2	regulatory protein
Regulador	XF1113	rpfG	77.0	two-component system, regulatory protein
Regulador	XF1114	rpfC	60.0	fused two-component sensor-regulator protein
Regulador	XF1115	rpfF	65.7	regulatory protein (DSF)
Precursor	XF0232	pgl	79.1	glucose-6-phosphate isomerase
Precursor	XF0259	xanB	84.5	phosphomannose isomerase-GDP-mannose pyrophosphorylase
Precursor	XF0260	xanA	84.8	phosphoglucmutase /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	gtbB	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

Embrapa

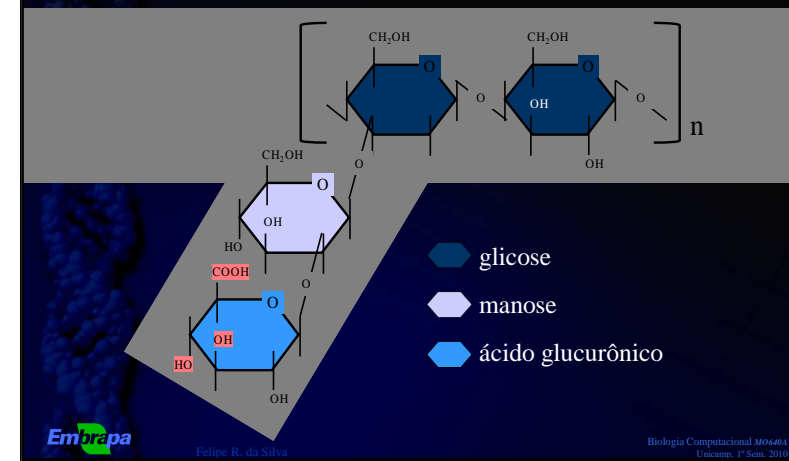
Felipe R. da Silva

Biologia Computacional M06404  
Unicamp, 1º Sem. 2010

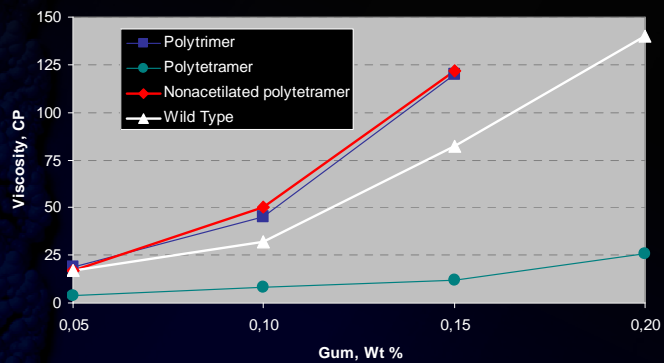
## Via sintética da goma fastidiana



## Possível estrutura da goma fastidiana



## Viscosidades de polímeros derivados da goma xantana



Hassler & Doherty, 1990. *Biotechnol. Prog.* 6 (3): 182-187.  
Unicamp, 1º Sem. 2010

## Grupo Genoma - CBMEG



Embrapa

Felipe R. da Silva

Biologia Computacional MO6404  
Unicamp, 1º Sem. 2010



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

[felipes@cnptia.embrapa.br](mailto:felipes@cnptia.embrapa.br)



Felipe R. da Silva

Biologia Computacional M06603  
Unicamp, 1º Sem. 2010