Gene finding for the helical cytokines

Darrell Conklin 1,*, Betty Haldeman 2 and Zeren Gao 2
1Department of Computing, City University London, United Kingdom
2ZymoGenetics Inc. Seattle, USA
*To whom correspondence should be addressed.

For comparison purposes, three other gene finding methods were also run on the same dataset: Fgenes v1.6 (Solovyev et al., 1994), Genscan v1.0 (Burge and Karlin ...
Shaw potassium channel genes in *Drosophila*

James J. L. Hodge¹,², James C. Choi², Cahir J. O'Kane¹*, Leslie C. Griffith²

¹Department of Genetics, University of Cambridge, Downing Site, Cambridge CB2 3EH, United Kingdom
²Department of Biology and Volen Center for Complex Systems, Brandeis University MS008, 415 South Street, Waltham, Massachusetts 02454-9110

... Intron/exon structure of Shawl was initially predicted using the **FGENES** program (Salamov and Solovyev, 2000) on the Baylor College of Medicine Genefinder ...


**Annotation of the Human Genome by High-Throughput Sequence Analysis of Naturally Occurring Proteins**

**Authors:** McGowan S.J.¹; Terrett J.¹; Brown C.G.¹; Adam P.J.¹; Aldridge L.¹; Allen J.C.¹; Amess B.¹; Andrews K.A.¹; Barnes M.¹; Barnwell D.E.¹

**Affiliations:** ¹: Oxford GlycoSciences plc, The Forum, 86 Milton Park, Abingdon, OX14 4RY, UK.

... The polymorphic ‘hypothetical transcriptome’ was cons- tructed from transcripts predicted by **FGENES, FGENESH** (Softberry Inc, Mount Kisco, NY, USA), GENSCAN ...

Nature Biotechnology 22, 1146 – 1149

**5'-end SAGE for the analysis of transcriptional start sites**

Hashimoto et al. (2004)...

...from expressed sequence tag (EST) maps, analysis of full-length cDNAs and computational annotation by Genscan, Genie, **FGENES** and other programs.

**Hemoglobin** Volume 28, Number 3 / 2004 255 – 259

**An a-Thalassemia Phenotype in a Dutch Hindustani, Caused by a New Point Mutation that Creates an Alternative Splice Donor Site in the First Exon of the a2-Globin Gene**

Cornelis L. Harteveld A¹, Pierre W. Wijermans A², Peter van Delft A¹, Ellen Rasp A², Hans L. Haak A², Piero C. Giordano A¹

¹Hemoglobinopathies Laboratory, Leiden University Medical Center, PO Box 9503, 2300, Leiden, The Netherlands
²Leyenburg Hospital, The Hague, The Netherlands

... and **GRAIL/ gap2** (Gene Recognition and Assembly Internet Link, Sequence exploration and gene discovery Version 1.3), Genefinder, GENSCAN, **Fgenes** and HMMGene ...

Journal of Medical Genetics 2004;41:e52

**Genomic organisation of the UDP-N-acetylglucosamine-1-phosphotransferase gamma subunit (GNPTAG) and its mutations in mucolipidosis III**
... third intron is 9 kb. This structure was consistent with the exon prediction of \textit{fgenes} and Genscan. Each splice donor and acceptor ...
FGENESH

European Journal of Plant Pathology  Issue:  Volume 112, Number 1 Date:  May 2005 Pages: 23 - 29
DOI: 10.1007/s10658-004-7088-7

Leptosphaeria maculans, a fungal pathogen of Brassica napus, secretes a subtilisin-like serine protease

Leanne M. Wilson\textsuperscript{1} and Barbara J. Howlett\textsuperscript{1}
\textsuperscript{1} School of Botany, The University of Melbourne, Parkville, Victoria, 3010, Australia

... DNA and cDNA sequences were compared to identify intron positions, which confirmed those predicted by FGENESH gene prediction software (www.softberry.com). ...

Current Genetics  Issue:  Volume 47, Number 5 Date:  May 2005 Pages: 307 – 315
DOI: 10.1007/s00294-004-0559-8

During attachment Phytophthora spores secrete proteins containing thrombospondin type 1 repeats

Andrea V. Robold\textsuperscript{1} and Adrienne R. Hardham\textsuperscript{1}
\textsuperscript{1} Plant Cell Biology Group, Research School of Biological Sciences, The Australian National University, Canberra, ACT 2601, Australia

... info.html). The DNA sequence was searched for introns using the soft- ware program FGENESH (http://www.softberry.com/berry.phtml ...

Microbiology 151 (2005), 1499-1505; DOI 10.1099/mic.0.27759-0

Overproduction, purification and characterization of FgaPT2, a dimethylallyltryptophan synthase from Aspergillus fumigatus

Inge A. Unsöld and Shu-Ming Li
Pharmazeutische Biologie, Pharmazeutisches Institut, Eberhard-Karls-Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

Correspondence Shu-Ming Li shuming.li@uni-tuebingen.de

... FGENESH (Softberry; www.softberry.com/berry.phtml) and the DNASIS software package (version 2.1; Hitachi Software Engineering) were used for intron prediction ...


Sex-Specific Homeodomain Proteins Sxi1\textsuperscript{1,2} and Sxi2a Coordinately Regulate Sexual Development in Cryptococcus neoformans

Christina M. Hull,\textsuperscript{1}\textsuperscript{*}, Marie-Josee Boily,\textsuperscript{1} and Joseph Heitman\textsuperscript{1,2*}
Department of Molecular Genetics and Microbiology,\textsuperscript{1} the Howard Hughes Medical Institute, Duke University Medical Center, Durham, North Carolina\textsuperscript{2}
Received 14 January 2005/ Accepted 14 January 2005
* Corresponding author. Mailing address: 322 CARL Building, Box 3546, Department of Molecular Genetics and Microbiology, 100 Research Dr., Duke University Medical Center, Durham, NC 27710. Phone: (919) 684-2824. Fax: (919) 684-5458. E-mail: heitm001@duke.edu
\textsuperscript{1}Present address: Departments of Biomolecular Chemistry and Medical Microbiology & Immunology, University of Wisconsin—Madison, Madison, WI 53706.
Sequence manipulations. Splice predictions of candidate gene sequences for SXI2a were facilitated with a **Softberry** algorithm (www.softberry.com). ...
BIOINFORMATICS-PLANT DATABASES
The Institute for Genomic Research Osa1 Rice Genome Annotation Database

Qiaoping Yuan, Shu Ouyang, Aihui Wang, Wei Zhu, Rama Maiti, Haining Lin, John Hamilton, Brian Haas, Razvan Sultana, Foo Cheung, Jennifer Wortman and C. Robin Buell*

The Institute for Genomic Research, Rockville, Maryland 20850

1 This work (on rice genome annotation) was supported by the National Science Foundation (grant no. DBI–0321538 to C.R.B.) and the U.S. Department of Agriculture (grant no. 2003–35317–13173 to C.R.B.).

2 Present address: Laboratory of Neurogenetics, NIAAA, NIH, 5625 Fishers Lane, Suite 3532, MSC 9412, Bethesda, MD 20892.

* Corresponding author; e-mail rbuell@tigr.org; fax 301–838–0208.

Received December 31, 2004; returned for revision February 24, 2005; accepted March 21, 2005.

... The ab initio gene finders used in the rice EGC pipeline include FGENESH (monocot matrix; Salamov and Solovyev, 2000 ), GeneMark.hmm (rice matrix; Lukashin and ...
Finally, we compared TWINSCAN with two other gene-prediction systems that have recently been developed for nematodes—FGENESH (Salamov and Solovyev 2000...
predict the genes between D9S277 and D9S127 in 9q31.1, the gene prediction tools also in the UCSC Genome Browser having Acembly, Ensembl, Fgenesh , GenScan and ...

**Genome Research** 15:54-66, 2005

Gene and alternative splicing annotation with AIR

Liliana Florea1,4,5, Valentina Di Francesco2, Jason Miller1, Russell Turner1, Alison Yao2, Michael Harris3, Brian Walenz1, Clark Mobyar1, Gennady V. Merkulov3, Rosane Charlab3, Ian Dew1, Zuoming Deng3, Sorin Istrail1, Peter Li4 and Granger Sutton1

1 *Informatics Research, Applied Biosystems, Rockville, Maryland 20850, USA;* 2 *Advanced Solutions, Celera Genomics, Rockville, Maryland 20850, USA;* 3 *Scientific Content and Applications, Celera Genomics, Rockville, Maryland 20850, USA*

... Ab initio prediction programs such as GenScan (Burge and Karlin 1997 ), FGenesh (Salamov and Solovyev 2000 ), Genie (Kulp et al. ...
and Mathematical Sciences, University of Aarhus; \textsuperscript{5} Aros Applied Biotechnology, Aarhus, Denmark; and \textsuperscript{6} Eos Biotechnology/Protein Design Labs, Fremont, California

Requests for reprints: Torben F. Ørntoft, Molecular Diagnostic Laboratory, Department of Clinical Biochemistry, Aarhus University Hospital, Skejby, DK-8200 Aarhus N, Denmark. Phone: 45-89495100; Fax: 45-89496018; E-mail: orntoft@ki.au.dk

... array comprising 59,619 probe sets representing 46,000 unique sequences, including known genes, expressed sequence tag clusters, and \textsc{FGENESH}-predicted exons ...


\textbf{Research article}

\textbf{The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants}

Yuanji Zhang* and Liangjiang Wang
Address: Plant Biology Division, The Samuel Roberts Noble Foundation, Ardmore, OK 73402, USA
Email: Yuanji Zhang* - yjzhang@noble.org; Liangjiang Wang - Kevinlwang@aol.com
* Corresponding author

... Despite minor differences in the gene structure prediction, both gene prediction programs \textsc{FGENESH} and GENSCAN agree on the major features of the protein ...

\textit{Nature Immunology}, 2005 v.6 n.3

\textbf{The immunoglobulin heavy-chain locus in zebrafish: identification and expression of a previously ...}

N Danilova, J Bussmann, K Jekosch, LA Steiner - - nature.com
Nature Immunology, Full text access provided to: Googlebot Access by: Web Services. ...

Published online 2005 May 24. doi: 10.1371/journal.pbio.0030181

\textbf{RAG1 Core and V(D)J Recombination Signal Sequences Were Derived from Transib Transposons}

Vladimir V Kapitonov\textsuperscript{\textcopyright i} and Jerzy Jurka\textsuperscript{\textcopyright i}
\textsuperscript{1}Genetic Information Research Institute, Mountain View, California, United States of America
David Nemazee, Academic Editor
Scripps Research Institute, United States of America
\textsuperscript{\textcopyright}Corresponding author.

Vladimir V Kapitonov: vladimir@girinst.org; Jerzy Jurka: jurka@girinst.org

... Using \textsc{FGENESH} [33], we detected that the RAG1 core–like open reading frame (ORF) in the contig 29068 forms a terminal exon (positions 1154–2947) of an ...

\textbf{Genetics:} Published Articles Ahead of Print, published on February 16, 2005 as 10.1534/genetics.104.036327

\textbf{Identification and Characterization of Regions of the Rice Genome Associated with Broad-Spectrum, Quantitative Disease Resistance}

Randall J. Wisser\textsuperscript{*}, Qi Sun\textsuperscript{†}, Scot H. Hultbert\textsuperscript{§}, Stephen Kresovich\textsuperscript{*}, and Rebecca J. Nelson\textsuperscript{*††, 1}
\textsuperscript{*}Department of Plant Breeding and Genetics, Institute for Genomic Diversity, Cornell University, Ithaca, New York 14853, \textsuperscript{†}Computational Biology Service Unit, Cornell Theory
Annotations and Functional Analyses of the Rice WRKY Gene Superfamily Reveal Positive and Negative Regulators of Abscisic Acid Signaling in Aleurone Cells

Zhen Xie, Zhong-Lin Zhang, Xiaolu Zou, Jie Huang, Paul Ruas, Daniel Thompson and Qingxi J. Shen

Department of Biological Sciences, University of Nevada, Las Vegas, Nevada 89154

This work was supported by the U.S. Department of Agriculture (grant no. 02–35301–12066), by the National Institutes of Health (Biomedical Research Infrastructure Network seed grant no. P20 RR16464), and by the University of Las Vegas, Nevada (start-up funds to Q.J.S.). X.Z. was supported by a National Science Foundation Experimental Program to Stimulate Competitive Research (EPSCoR), Integrative Approaches to Abiotic Stress (EPS–0132556), graduate assistantship and Z.-L.Z. by a National Science Foundation EPSCoR Advanced Computing in Environmental Sciences postdoctoral fellowship.

These authors contributed equally to the paper.

Present address: Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720 and Plant Gene Expression Center, U.S. Department of Agriculture, Albany, CA 94710.

The online version of this article contains Web-only data.

First of all, three genes (OsWRKY41, -43, and -44) were reannotated using FGENESH (www.softberry.com), because the first introns of these genes were too small ...


RAG1 Core and V(D)J Recombination Signal Sequences Were Derived from Transib Transposons.

Vladimir V. Kapitonov*, Jerzy Jurka*

*To whom correspondence should be addressed. E-mail: vladimir@girinst.org (VVK), E-mail: jurka@girinst.org (JJ)

Using FGENESH [33], we detected that the RAG1 core–like open reading frame (ORF) in the contig 29068 forms a terminal exon (positions 1154–2947) of an ...


From Mapping to Sequencing, Post-sequencing and Beyond

Takuji Sasaki, Takashi Matsumoto, Baltazar A. Antonio and Yoshiaki Nagamura

National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan

* Corresponding author: E-mail, tsasaki@nias.affrc.go.jp
... The gene predictions by programs such as Genescan (Burge and Karlin 1997 ), FGENESH [see Appendix 1 (4)] and Genemark [see Appendix 1 (5)], BLAST (Altschul et ...

**Improving the nutritional value of Golden Rice through increased pro-vitamin A content**

JA Paine, CA Shipton, S Chaggar, RM Howells, MJ … - Nature Biotechnology, 2005 - nature.com

... Arabidopsis thaliana psy and rice psy (AY024351) genes identified genomic sequences of similarity in which genes were predicted using FGENESH algorithm with ...

---

**Genetics**: Published Articles Ahead of Print, published on January 16, 2005 as 10.1534/genetics.104.035543

**THE GENETIC BASIS FOR INFLORESCENCE VARIATION BETWEEN FOXTAIL AND GREEN MILLET (POACEAE)**

Andrew N. Doust*, Katrien M. Devos†, Mike D. Gadberry*, Mike D. Gale†, & Elizabeth A. Kellogg*

*University of Missouri-St. Louis, Department of Biology, One University Boulevard, St. Louis, MO 63121, USA
†John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK

Current address: University of Georgia-Athens, Departments of Crop and Soil Sciences and Plant Biology, Miller Plant Sciences Building, Athens, GA 30602, USA

... Each of these contigs was scanned using FgeneSH (S ALAMOV and S OLOVYEV 2000), and open reading frames (ORFs) were translated and ... PLoS Biol 3(1): e13 January 2005

**Sorghum Genome Sequencing by Methylation Filtration**

Joseph A. Bedell*, Muhammad A. Budiman†, Andrew Nunberg‡, Robert W. Citek‡, Dan Robbins‡, Joshua Jones‡, Elizabeth Flick‡, Theresa Rohlf‡, Jason Fries‡, Kourtney Bradford‡, Jennifer McMenamy‡, Michael Smith‡, Heather Holeman‡, Bruce A. Roe‡, Graham Wiley‡, Ian F. Korf‡, Pablo D. Rabinowicz‡, Nathan Lakey‡, W. Richard McCombie‡, Jeffrey A. Jeddeloh‡, Robert A. Martienssen‡

1 Bioinformatics, Orion Genomics, Saint Louis, Missouri, United States of America,, 2 Library Construction, Orion Genomics, Saint Louis, Missouri, United States of America,, 3 Sequencing, Orion Genomics, Saint Louis, Missouri, United States of America,, 4 Biomarkers, Orion Genomics, Saint Louis, Missouri, United States of America,, 5 Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma, United States of America,, 6 Genome Center, University of California, Davis, California, United States of America,, 7 The Institute for Genomic Research, Rockville, Maryland, United States of America,, 8 Business, Orion Genomics, Saint Louis, Missouri, United States of America,, 9 Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States of America

... additional parameters: wordmask=seg; icmask; M=1; N=–1; Q=3; R=3; kap; E=1e-10; hspmax=0. To look for potentially novel genes, we used FgenesH (http://www ...


**FAM20: an evolutionarily conserved family of secreted proteins expressed in hematopoietic cells**

Demet Nalbant‡, Hyewon Youn‡, S Isil Nalbant‡, Savitha Sharma‡, Everardo Cobos‡, Elmus G Beale‡, Yang Du‡ and Simon C Williams‡

1Department of Cell Biology and Biochemistry, Texas Tech University Health Sciences Center,
These results were compared against genes assembled by two gene prediction programs, FGENESH: http://www.softberry.com/berry.phtml and GENSCAN: http://genes.mit ...

---


Complex Organization and Evolution of the Tomato Pericentromeric Region at the FER Gene Locus

Romain Guyot, Xudong Cheng, Yan Su, Zhukuan Cheng, Edith Schlagenhauf, Beat Keller and Hong-Qing Ling  
Institute of Plant Biology, University of Zurich, 8008 Zurich, Switzerland (R.G., E.S., B.K., H.-Q.L.); and Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Chaoyang District, Beijing 100101, China (X.C., Y.S., Z.C., H.-Q.L.)

... Putative genes were determined by a combination of coding region prediction software (GENSCAN, FGENESH, and MZEF with Arabidopsis and/or monocot matrix ...

---

**J Gen Virol** 86 (2005), 973-983; DOI 10.1099/vir.0.80833-0

Cloning, characterization and analysis by RNA interference of various genes of the *Chelonus inanitus* polydnavirus

Marianne Bonvin, Dorothee Marti, Stefan Wyder, Dejan Kojic, Marc Annaheim and Beatrice Lanzrein  
Institute of Cell Biology, University of Berne, Baltzerstrasse 4, CH-3012 Bern, Switzerland  
**Correspondence** Beatrice Lanzrein beatrice.lanzrein@izb.unibe.ch

... 12g1forw (5'-GAGTCCATGCCGAATGTCAC-3') and 12g1rev (5'-CTTCTTGCACAGCGACGAAC-3') were set to amplify the middle region of 12g1, as predicted with FGENESH 1.0 and ...

---

**The Plant Cell** 17:343-360 (2005)

Evolution of DNA Sequence Nonhomologies among Maize Inbreds

Stephan Brunner, Kevin Fengler, Michele Morgante, Scott Tingey and Antoni Rafalski  
DuPont Crop Genetics Research, Wilmington, Delaware 19880-353  
Università degli Studi di Udine, Dipartimento di Scienze Agrarie ed Ambientali, 33100 Udine, Italy

1 To whom correspondence should be addressed. E-mail stephan.brunner@cgr.dupont.com; fax 302-695-2726

---

**PNAS** | February 1, 2005 | vol. 102 | no. 5 | 1566-1571  
Published online before print January 24, 2005, 10.1073/pnas.0409421102

A computational and experimental approach to validating annotations and gene predictions in the *Drosophila melanogaster* genome
Detection and analysis of alternative splicing in the silkworm by aligning expressed sequence tags with the genomic sequence

Correspondence: Dr Qing-You Xia, The Key Sericultural Laboratory of Agricultural Ministry, Southwest Agricultural University, Chongqing 400716, China. Tel.: +86 23 68250748; fax: +86 23 68251128; e-mail: xiaqy@swau.cq.cn

... previously predicted silkworm genes in the genomic sequences by BGF, a newly developed program based on GENSCAN (Burge & Karlin, 1997) and Fgenes (Salamov & ...
Position Effects Due to Chromosome Breakpoints that Map \(\sim 900\) Kb Upstream and \(\sim 1.3\) Mb Downstream of SOX9 in Two Patients with Campomelic Dysplasia


Departments of Pathology and Pediatrics, University of Texas Medical Branch, Galveston; Departments of Molecular and Human Genetics and Pediatrics, Baylor College of Medicine, and Texas Children's Hospital, Houston; and Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN

...In an effort to identify possible transcripts that may be responsible for the CD phenotype, we used several gene-prediction programs and identified seven hypothetical transcripts in the region that spans 100 kb in either direction from the breakpoint on chromosome 17—\textit{Ecgene}\textsuperscript{6} H17C12306.1 and H17C12308.1, SGP genes Chr17_1538.1 and Chr17_1539.1, \textit{Fgenesh++} gene C17001650, and \textit{Genscan} genes NT_010641.44 and NT_010641.45...
....The gene annotations used to classify the constrained elements contain nearly 40,000 genes, including RefSeq genes and gene predictions; they are based on annotations for the human, mouse, and rat genomes made by Fgenesh++ software developed by Softberry Inc. (Solovyev 2002; http://www.softberry.com)...
For human genome applications, the **FGENESH++** software was first used to map known human genes using sequences available from the Reference Sequence (RefSeq) Project at the National Center for Biotechnology Information (NCBI; Bethesda, MD; [www.ncbi.nlm.nih.gov/LocusLink/refseq.html](http://www.ncbi.nlm.nih.gov/LocusLink/refseq.html)).

REFERENCES

**Genome Research** 14:539-548, 2004
ISSN 1088-9051/ $5.00

**Letter**
Characterization of Evolutionary Rates and Constraints in Three Mammalian Genomes

Gregory M. Cooper¹, Michael Brudno², Eric A. Stone³, Inna Dubchak⁴, Serafim Batzoglou² and Arend Sidow¹,⁵,⁶

¹ Department of Genetics, Stanford University, Stanford, California 94305, USA; ² Department of Computer Science, Stanford University, Stanford, California 94305, USA; ³ Department of Statistics, Stanford University, Stanford, California 94305, USA; ⁴ Department of Pathology, Stanford University, Stanford, California 94305, USA; ⁵ Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

The gene annotations used to classify the constrained elements contain nearly 40,000 genes, including RefSeq genes and gene predictions; they are based on annotations for the human, mouse, and rat genomes made by **Fgenesh++** software developed by Softberry Inc. (Solovyev 2002; [http://www.softberry.com](http://www.softberry.com)).

Published online before print June 12, 2003, 10.1101/gr.529803

**Genome Research** 13:1765-1774, 2003
ISSN 1088-9051/ $5.00

Identification of Promoter Regions in the Human Genome by Using a Retroviral Plasmid Library-Based Functional Reporter Gene Assay

Shirin Khambata-Ford¹,⁵, Yueyi Liu², Christopher Gleason¹, Mark Dickson³, Russ B. Altman², Serafim Batzoglou⁴ and Richard M. Myers¹,³,⁶

¹ Department of Genetics, Stanford University School of Medicine, Stanford, California 94305, USA; ² Stanford Medical Informatics, Stanford University School of Medicine, Stanford, California 94305, USA; ³ Stanford Human Genome Center, Stanford University School of Medicine, Stanford, California 94305, USA; ⁴ Department of Computer Science, Stanford University, Stanford, California 94305, USA

…Of 858 sequences, 9% of GFP+ low clones and 8% of GFP+ high clones aligned to the 2-kb segment upstream of the transcription start site of a predicted gene in at least two of four data sets of predicted genes from Genscan, Ensembl, Softberry (**Fgenesh++**), and Acembly (category B in [Table 1](#)).


HIV-1 Integration in the Human Genome
Favors Active Genes and Local Hotspots
Astrid R.W. Schroeder, Paul Shinn, Huaming Chen, Charles Berry, Joseph R. Ecker, and Frederic Bushman

Infectious Disease Laboratory
Genomic Analysis Laboratory The Salk Institute 10010 North Torrey Pines Road La Jolla, California 92037

Department of Family/Preventive Medicine School of Medicine University of California, San Diego San Diego, California 92093

An integration target sequence was scored as a part of a transcription unit if it was (1) a member of the Refseq set of well-studied genes (http://www.ncbi.nlm.nih.gov/LocusLink/refseq.html) or (2) if, it was predicted to be a transcription unit by the ENSEMBLE (http://www.ensembl.org) or Fgenesh++ (http://www.softberry.com/Help/fgeneshplus2.htm) programs and if that assignment was supported by mRNA or spliced EST sequence evidence.

Genome Research
13:313–322 ©2003
Received March 27, 2002; accepted in revised form December 3, 2002.
Article and publication are at http://www.genome.org/cgi/doi/10.1101/gr.313703.

A Complexity Reduction Algorithm for Analysis and Annotation of Large Genomic Sequences

Trees-Juen Chuang, 1 Wen-Chang Lin, 1 Hurng-Chun Lee, 2 Chi-Wei Wang, 2 Keh-Lin Hsiao, 2 Zi-Hao Wang, 2 Danny Shieh, 2 Simon C. Lin, 2 and Lan-Yang Ch’ang 1,3
1 Bioinformatics Research Center, Institute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan; 2 Academia Sinica Computing Center, Academia Sinica, Taipei 11529, Taiwan
3 Corresponding author.
E-MAIL lychang@ibms.sinica.edu.tw; FAX 886-2-27858594.

Numerous ab initio prediction programs have been used extensively in genome annotation, including FGENES H (Solovyev et al. 1995; Salamov and Solovyev 2000),… Successful implementation of this method includes AAT (Huang et al. 1997), FGENES H+ and FGENES H++ (Salamov and Solovyev 2000),… Among these programs FGENES H+ (and FGENES H++), GenomeScan, GeneWise, and Procrustes are combined tools of sequence homology and ab initio annotation.
FGENESH+

Automated Whole-Genome Multiple Alignment of Rat, Mouse, and Human

Michael Brudno¹, Alexander Poliakov², Asaf Salamov³,⁴, Gregory M. Cooper⁵, Arend Sidow⁵,⁶, Edward M. Rubin²,³, Victor Solovyev³,⁴, Serafim Batzoglou¹,⁵ and Inna Dubchak²,³,⁷

¹ Department of Computer Science, Stanford University, Stanford, California 94305, USA; ² Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA; ³ U.S. Department of Energy Joint Genome Institute, Walnut Creek, California 94598, USA; ⁴ Softberry Inc., Mount Kisco, New York 10549, USA; ⁵ Department of Genetics, Stanford University, Stanford, California 94305-5324, USA; ⁶ Department of Pathology, Stanford University, Stanford, California 94305-5324, USA

... Fgenesh+ gene prediction is conducted on sequences with protein homology ...

A Complexity Reduction Algorithm for Analysis and Annotation of Large Genomic Sequences

Trees-Juen Chuang,¹ Wen-Chang Lin,¹ Hürn-Chun Lee,² Chi-Wei Wang,² Keh-Lin Hsiao,² Zi-Hao Wang,² Danny Shieh,² Simon C. Lin,² and Lan-Yang Ch’ang¹,³
¹Bioinformatics Research Center, Institute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan; ²Academia Sinica Computing Center, Academia Sinica, Taipei 11529, Taiwan
³Corresponding author.
E-MAIL lychang@ibms.sinica.edu.tw; FAX886-2-27858594.

…Numerous ab initio prediction programs have been used extensively in genome annotation, including FGENESH (Solovyev et al. 1995; Salamov and Solovyev 2000), …Successful implementation of this method includes AAT (Huang et al. 1997), FGENESH+ and FGENESH++ (Salamov and Solovyev 2000),… Among these programs FGENESH+ (and FGENESH++), GenomeScan, GeneWise, and Procrustes are combined tools of sequence homology and ab initio annotation.

DATABASES AND TOOLS FOR BROWSING GENOMES

Ewan Birney,¹ Michele Clamp,² and Tim Hubbard²
¹European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, United Kingdom; e-mail: birney@ebi.ac.uk
²Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, United Kingdom; michele@sanger.ac.uk th@sanger.ac.uk
... Another predicted gene track on the UCSC browser comes from Softberry (http://www.softberry.com) and uses a program Fgenesh+, which is based on HMMs and...
Oxygen-Controlled Bacterial Growth in the Sponge *Suberites domuncula*: toward a Molecular Understanding of the Symbiotic Relationships between Sponge and Bacteria†

Werner E. G. Müller,* Vladislav A. Grebenjuk, Narsinh L. Thakur, Archana N. Thakur, Renato Batel, Anatoli Krasko, Isabel M. Müller, and Hans J. Breter
Institut für Physiologische Chemie, Abteilung Angewandte Molekularbiologie, Universität Mainz, D-55099 Mainz, Germany
*Corresponding author. Mailing address: Institut für Physiologische Chemie, Abteilung Angewandte Molekularbiologie, Universität Mainz, Duesbergweg 6, 55099 Mainz, Germany. Phone: 6131-3925910. Fax: 6131-3925243. E-mail: wmueller@mail.uni-mainz.de

... For genes and potential promoter prediction, we used the FGENESB-Pattern/Markov chain-based bacterial operon and gene prediction program from the SoftBerry ...

Computational prediction of conserved operons and phylogenetic footprinting of transcription regulatory elements in the metal-reducing bacterial family Geobacteraceae

Bin Yan, Barbara A. Methe, Derek R. Lovley, Julia Kruskal
Department of Preventive Medicine, Center of Genomics and Bioinformatics, University of Tennessee Health Science Center, 66 N. Pauline St., Ste. 633, Memphis, TN 38163, USA
The Institute for Genomic Research, Rockville, MD, USA
Department of Microbiology, Morrill Science Center IV North, University of Massachusetts, 659 North Pleasant Str., Amherst, MA 01003, USA

... the conserved nature of the operons 2. Operons in Geobacter sulfurreducens were predicted ab initio by the public version of program FGENESB (V. Solovyev and V ...

Gene conversion: a mechanism for generation of heterogeneity in the tprK gene of Treponema pallidum during infection

Arturo Centurion-Lara, Rebecca E. LaFond, Karin Hevner, Charmie Godornes, Barbara J. Molini, Wesley C. Van Voorhis and Sheila A. Lukehart

... Using the fgenesb program, which identifies putative operons and genes in microbial genomes (Softberry; http://www.softberry.com/ berry.phtml ), the tprK ORF ...

Proteorhodopsin genes are distributed among divergent marine bacterial taxa

José R. de la Torre, Lynne M. Christianson, Oded Béjà, Marcelino T. Suzuki, David M. Karl, John Heidelberg, and Edward F. DeLong
Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA 95039;
Department of Biology, Technion-Israel Institute of Technology, Haifa 32000, Israel;
… Analysis of the potential genes and protein-coding regions was performed by using a combination of the BLAST (11), GLIMMER 2.02 (TIGR) (12, 13), FGENESB (Softberry, Mount Kisco, NY), and ARTEMIS (Sanger Center, Cambridge University, U.K.) (14) software packages. 

Different SAR86 subgroups harbour divergent proteorhodopsins

Gazalah Sabehi1; Oded Béjà1; Marcelino T. Suzuki2; Christina M. Preston3; Edward F. DeLong4

Affiliations: 1: Department of Biology, Technion-Israel Institute of Technology, Haifa 32000, Israel. 2: Chesapeake Biological Laboratory, University of Maryland Center for Environmental Sciences, Solomons, MD 20688, USA. 3: Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA 95039, USA. 4: Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

… program. FGENESB (Softberry), and the annotation was subsequently refined and curated manually using ARTEMIS (Sanger Center). Fig. …
Functional Genomics Analysis of Singapore Grouper Iridovirus: Complete Sequence Determination and Proteomic Analysis

Wen Jun Song, Qi Wei Qin, Jin Qiu, Can Hua Huang, Fan Wang, and Choy Leong Hew
Department of Biological Sciences, Tropical Marine Science Institute, National University of Singapore, Singapore

... A total of 162 ORFs, predicted by the FGENESV program (available through: http://www.softberry.com), supplemented with Vector NTI suite 7.1, are indicated ...

doi:10.1023/B:VIRU.0000025771.48128.f8
Virus Genes 28 (3): 239-246, April 2004
Article ID: 5269250

Complete Nucleotide Sequence of a Strawberry Isolate of Beet Pseudoyellows Virus

Ioannis E. Tzanetakis
Molecular and Cellular Biology Program, Department of Botany and Plant Pathology, Oregon State University, Corvallis 97331, USA
Robert R. Martin
Horticultural Crops Research Laboratory, USDA-ARS, Corvallis OR 97330, USA; E-mail: martinrr@science.oregonstate.edu


Genome Organization of the SARS-CoV

Jing Xu1*, Jianfei Hu2;1*, Jing Wang2;1*, Yujun Han1*, Yongwu Hu1;3, Jie Wen1, Yan Li1, Jia J1, Jia Ye1;4, Zizhang Zhang5, Wei Wei4, Songgang Li1;2, Jun Wang1, Jian Wang1;4, Jun Yu1;4#, and Huanming Yang1;4#
1 Beijing Genomics Institute, Chinese Academy of Sciences, Beijing 101300, China; 2 College of Life Sciences, Peking University, Beijing 100871, China; 3 Wenzhou Medical College, Wenzhou 325003, China; 4 James D. Watson Institute of Genome Sciences, Zhejiang Campus, Zhejiang University and Hangzhou Genomics Institute, Hangzhou 310008, China; 5 College of Materials Science and Chemical Engineering, Yuquan Campus, Zhejiang University, Hangzhou 310027, China.
* These authors contributed equally to this work.
# Corresponding authors.
E-mail: junyu@genomics.org.cn; yanghm@genomics.org.cn

... FGENESV, a program for gene prediction provided by Softberry Inc. (Mount Kisco, USA) through a web-based interface, has been specially modified and trained with parameters for virus (http://www.softberry.com/berry.phtml?topic= gfindv). ...
The hypothetical minus sense ORF identified by FGENESV (from 48 to 203 nt on the minus strand or 29,523 to 29,678 nt on the plus strand) may be fake, but we should not absolutely deny the probability of the existence of minus ORFs.

Furthermore, we employed FGENESV to explore the sequences of MHV (NC 001846 in NCBI) and AIBV (NC 001451 in NCBI), and compared the results with their previous annotations, respectively.

Rapport de stage de DEA Juin 2003

Analyse du génome du virus de l’archée Pyrococcus abyssi (PAV1)

ROUAULT Karen
Laboratoire de Microbiologie et Biotechnologie des Extrêmophiles IFREMER- Centre de Brest et Equipe Microbiologie LEMAR – Institut Universitaire Européen de la Mer

... [14]. FGENESV http://www.softberry.com/berry.phtml?topic=gfindv Virus (>10 kb) Modèles de Markov Forme du génome Code génétique [40]. ...
This approach has been applied eg. in HMMgene or in **FGENES-M** (unpub.)...

---

**DNA Sequence - The Journal of Sequencing and Mapping**

**Volume 15, Number 1 / February 2004 39-43**

**Isolation and Characterization of Two cDNAs Encoding Translation Initiation Factor 1A from Rice (Oryza sativa L.)**

Ji Huang, Jianfei Wang, Shengping Qiu, Hongsheng Zhang
State Key Laboratory of Crop Genetics and Germplasm Enhancement Rice Research Institute Nanjing Agricultural University 210095 Nanjing People's Republic of China

... Then the ORF without introns was assembled based on the sequence of each scaffold using the program **FGENES-M** (http://www. softberry.com). ...
Sex-Specific Homeodomain Proteins Sxi1p and Sxi2a Coordinately Regulate Sexual Development in Cryptococcus neoformans

Christina M. Hull,† Marie-Josee Boily, and Joseph Heitman

Department of Molecular Genetics and Microbiology, the Howard Hughes Medical Institute, Duke University Medical Center, Durham, North Carolina

Received 14 January 2005/ Accepted 14 January 2005

* Corresponding author. Mailing address: 322 CARL Building, Box 3546, Department of Molecular Genetics and Microbiology, 100 Research Dr., Duke University Medical Center, Durham, NC 27710. Phone: (919) 684-2824. Fax: (919) 684-5458. E-mail: heitm001@duke.edu

†Present address: Departments of Biomolecular Chemistry and Medical Microbiology & Immunology, University of Wisconsin—Madison, Madison, WI 53706.

... Sequence manipulations. Splice predictions of candidate gene sequences for SXI2a were facilitated with a Softberry algorithm (www.softberry.com). ... We utilized the BESTORF gene prediction algorithm from Softberry, Inc., to electronically produce predicted spliced cDNA products encoded by a 10-kb region....
PROTCOMP

Genes and Immunity, 2005, v.5, n.4 - nature.com

Immune response in silico(IRIS): immune-specific genes identified from a compendium of microarray …

AR Abbas, D Baldwin, Y Ma, W Ouyang, A Gurney, F …

... The Protcomp algorithm (Softberry, Inc.) predicts for the 1589 IRIS genes with ORFs that 24% of the encoded proteins are in the plasma membrane, 13% are ...


Lotus japonicus LjKUP Is Induced Late During Nodule Development and Encodes a Potassium Transporter of the Plasma Membrane

Guilhem Desbrosses, Claudia Kopka, Thomas Ott, and Michael K. Udvardi
Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Golm, Germany

…Both PSORT and Protcomp predicted a PM location for LjKUP…

Planta
DOI: 10.1007/s00425-003-1182-5
Issue: Volume 218, Number 6 Date:  April 2004  Pages: 965 - 975

Biochemical and immunological characterization of pea nuclear intermediate filament proteins

Sonal S. D. Blumenthal1, Gregory B. Clark1 and Stanley J. Roux1
(1) School of Biological Sciences, Section of Molecular Cell and Developmental Biology, The University of Texas, Austin, TX 78712, USA
Stanley J. Roux   Email: sroux@uts.cc.utexas.edu

… html), BCM Search Launcher (Protein structure prediction, http://searchlauncher.bcm.tmc.edu/), SoftBerry (Protein subcellular. localization ...

Comparative and Functional Genomics Volume 5, Issue 4 , Pages 342 - 353
Published Online: 20 May 2004

Research Paper
Investigation into the use of C- and N-terminal GFP fusion proteins for subcellular localization studies using reverse transfection microarrays

Ella Palmer, Tom Freeman∗
MRC Rosalind Franklin Centre for Genomics Research (formerly the HGMP-Resource Centre), Genome Campus, Hinxton, Cambridge CB10 1SB, UK
email: Tom Freeman (tfreeman@rfcgr.mrc.ac.uk)
∗Correspondence to Tom Freeman, RFCGR, Hinxton, Cambridge/CB10 1SB, UK.
... ProtComp version 4 (Softberry) combines results with proteins of known subcellular localization and assumed subcellular localization (based on theoretical ... 


RHM2 Is Involved in Mucilage Pectin Synthesis and Is Required for the Development...

Usadel et al.

... Tentative subcellular localization prediction by TargetP (Emanuelsson et al., 2000) or ProtComp (http://www.softberry.com), a prediction software trained on ...

**Journal of Cellular Biochemistry**

*Volume 90, Issue 2*, Pages 361 - 378

*Published Online:* 3 Sep 2003

A proteomic study of the arabidopsis nuclear matrix

Tomasz T. Calikowski 1,3, Tea Meulia 2, Iris Meier 1,*
1Department of Plant Biology and Plant Biotechnology Center, Ohio State University, Columbus, Ohio 43210
2Molecular and Cellular Imaging Center, Ohio Agricultural and Research Development Center, Ohio State University, Columbus, Ohio 43210
3Institute of Biochemistry and Biophysics, Polish Academy of Sciences, UL. Pawinskiego, 5A, 02-106, Warszawa, Poland

**email:** Iris Meier (meier.56@osu.edu)

*Correspondence to Iris Meier, Department of Plant Biology and Plant Biotechnology Center, Ohio State University, 244 Rightmire Hall, 1060 Carmack Rd., Columbus, OH 43210.

... For prediction of subcellular localization, ProtComp 4 (Softberry, Inc., Mount Kisco, NY; http://www.softberry.com/berry.phtml?topic¼ proteinloc), PSORT v.6.4 ...

**Cellular Molecular Life Sciences**, 2003, in press

**Automatic prediction of protein function**

Burkhard Rost 1,2,3,* Jinfeng Liu 1,3,4, Rajesh Nair 1,5, Kazimierz O. Wrzeszczynski 1 and Yanay Ofran 1,6
1 CUBIC, Dept. of Biochemistry and Molecular Biophysics, Columbia University, 650 West 168th Street BB217, New York, NY 10032, USA
2 Columbia University Center for Computational Biology and Bioinformatics (C2B2), Russ Berrie Pavilion, 1150 St. Nicholas Avenue, New York, NY 10032, USA
3 North East Structural Genomics Consortium (NESG), Department of Biochemistry and Molecular Biophysics, Columbia University, 650 West 168th Street BB217, New York, NY 10032, USA
4 Dept. of Pharmacology, Columbia Univ., 630 West 168th Street, New York, NY 10032, USA
5 Dept. of Physics, Columbia Univ., 538 West 120th Street, New York, NY 10027, USA
6 Dept. of Medical Informatics, Columbia Univ., 630 West 168th Street, New York, NY 10032, USA

*Corresponding author: cubic@cubic.bioc.columbia.edu URL http://cubic.bioc.columbia.edu/ Tel: +1-212-305-4018, fax: +1-212-305-7932
An automated computational strategy was utilized to query each protein translation with the Signal Sensor, Sighmm, Tmdetect (T. Wu, unpubl.), hmmpfam (Eddy 1998•), and Protcomp (Softberry, Inc.) algorithms. The Protcomp algorithm predicts the subcellular localization of a protein, on the basis of homology to well-annotated proteins, a neural net, and various protein motifs. In this case, the Protcomp subcellular localization prediction was used to categorize these genes as "Other Secreted", "Other Transmembrane", or "Other Cytoplasmic or Nuclear".

**Gene prediction in eukaryota**

Samuel P. Hazen, John S. Scott-Craig, and Jonathan D. Walton*
Department of Energy-Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824
* Corresponding author; e-mail walton@msu.edu; fax 517-353-9168.

**The Cia5 gene controls formation of the carbon concentrating mechanism in Chlamydomonas reinhardtii**

Youbin Xiang, Jun Zhang, and Donald P. Weeks*
Department of Biochemistry and School of Biological Sciences, University of Nebraska, Lincoln, NE 68588-0664
Computer-assisted analysis of the CIA5 aa sequence (PROTCOMP, version 4, [http://www.softberry.com](http://www.softberry.com)) predicted a nuclear localization of the protein. Finally, computer program predictions (e.g., PROTCOMP version 4, [http://www.softberry.com](http://www.softberry.com)) for a nuclear localization of CIA5 and the clear-cut nuclear localization of CIA5 in onion epidermal cells (Fig. 3) provide additional weight to the argument that CIA5 may be a transcription factor.

Dissertation
zur Erlangung des akademischen Grades Dr. rer. nat.
der Fakultät der Naturwissenschaften
der Universität Ulm

**Untersuchungen zur Identifizierung von Faktoren und Mechanismen der mRNA 3' Prozessierung und Degradation in Chloroplasten höherer Pflanzen**

vorgelegt von
Michael Walter
aus Immenstadt i. Allgäu
Abteilung Molekulare Botanik, Universität Ulm
Ulm, November 2001
Tag der Promotion: 19 Feb. 2002

- Algorithmen zur Vorhersage der subzellulären Lokalisation:
  - Predotar: [http://www.inra.fr/Internet/Produits/Predotar/](http://www.inra.fr/Internet/Produits/Predotar/)
The genome of BCJA1c: a bacteriophage active against the alkaliphilic bacterium, *Bacillus clarkii*

Andrew M. Kropinski¹, Melissa Hayward¹, M. Dorothy Agnew¹ and Ken F. Jarrell¹
(1) Department of Microbiology and Immunology, Queens University, Kingston, ON, K7L 3N6, Canada

... al. 2002), Promoters were predicted using *Softberry's* BPROM program at http://www.softberry.com/berry.phtml?topic=promoter. ...

The Generalized Transducing *Salmonella* Bacteriophage ES18: Complete Genome Sequence and DNA Packaging Strategy

Sherwood R. Casjens,¹ ² * Eddie B. Gilcrease,¹  Danella A. Winn-Stampley,¹ ³* Petra Schicklmaier,³ ⁴* Horst Schmieger,³ Marisa L. Pedulla,² ⁴ Michael E. Ford,² ⁴ Jennifer M. Houtz,² ⁴ Graham F. Hatfull,² ⁴ and Roger W. Hendrix² ⁴

Department of Pathology, University of Utah Medical School, Salt Lake City, Utah,¹ Department of Biological Sciences,² Pittsburgh Bacteriophage Institute, University of Pittsburgh, Pittsburgh, Pennsylvania,² Institut für Genetik und Mikrobiologie, Universität München, Munich, Germany³

* Corresponding author. Mailing address: Department of Pathology, University of Utah Medical School, Salt Lake City, UT 84132. Phone: (801) 581-5980. Fax: (801) 581-3607. E-mail: sherwood.casjens@path.utah.edu

Present address: Biology Department, MIT, Cambridge, MA 02139.

Present address: Biogen Idec GmbH, D-85737 Ismaning, Germany.

Received 1 September 2004/ Accepted 3 November 2004

... The DNA sequence analysis software used was DNA Strider (24), GeneMark (5), Staden programs (78), BLAST (2), BPROM (http://www.softberry.com/berry.phtml?topic=...)

Infection and Immunity, May 2005, p. 2899-2909, Vol. 73, No. 5

Characterization of the Major Secreted Zinc Metalloprotease-Dependent Glycerophospholipid:Cholesterol Acyltransferase, PlaC, of *Legionella pneumophila*

Sangeeta Banerji,¹ Mayte Bewersdorff,¹,† Björn Hermes,¹,† Nicholas P. Cianciotto,² and Antje Flieger¹*

Robert Koch-Institut, Berlin, Germany,¹ Department of Microbiology-Immunology, Northwestern University Medical School, Chicago, Illinois²

Received 25 October 2004/ Returned for modification 18 November 2004/ Accepted 22 December 2004
... legion.) (12). Nucleotide sequences were also analyzed for promoters using the web-based program **BPROM** ([www.softberry.com](http://www.softberry.com)). Sequence ...

---


The Type III-Dependent Hrp Pilus Is Required for Productive Interaction of *Xanthomonas campestris* pv. vesicatoria with Pepper Host Plants

Ernst Weber,1 Tuula Ojanen-Reuhs,2,† Elisabeth Huguet,3‡ Gerd Hause,4 Martin Romantschuk,2§ Timo K. Korhonen,2 Ulla Bonas,1,3 and Ralf Koebnik1*

Institute of Genetics,1 Biozentrum, Martin Luther University, Halle, Germany, 4 General Microbiology, Faculty of Biosciences, University of Helsinki, Helsinki, Finland, 2 Institut des Sciences Végétales, CNRS, Gif-sur-Yvette, France

Received 10 November 2004/ Accepted 28 December 2004

* Corresponding author. Mailing address: Martin-Luther-Universität, Institut für Genetik, Weinbergweg 10, D-06120 Halle (Saale), Germany. Phone: 49 345 5526293. Fax: 49 345 5527151. E-mail: koebnik@gmx.de

† Present address: Purdue University, Department of Food Sciences, West Lafayette, IN 47907.

‡ Present address: Institut de Recherche sur la Biologie de l'Insecte, UMR, CNRS 6035, Faculté des Sciences, F-37200 Tours, France.

§ Present address: University of Helsinki, Department of Ecological and Environmental Sciences, FIN-15140 Lahti, Finland.

... The promoter recognition program **BPROM** ([www.softberry.com](http://www.softberry.com)) was used for prediction of bacterial sigma70 promoter motifs. RESULTS. ...

---


Identification of Operators and Promoters That Control SXT Conjugative Transfer

John W. Beaber and Matthew K. Waldor*

Department of Microbiology, Tufts University School of Medicine, and Howard Hughes Medical Institute, Boston, Massachusetts

*Corresponding author. Mailing address: Tufts University School of Medicine, 136 Harrison Ave., Jaharis 425, Boston, MA 02111. Phone: (617) 636-2730. Fax: (617) 636-2723. E-mail: matthew.waldor@tufts.edu.

Received April 1, 2004; Accepted May 24, 2004.

... Computer algorithms and 5′ random amplification of cDNA ends (RACE) were used to define the setR and s086 transcription start sites. Software for the identification of bacterial promoters ([http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb](http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb)) identified putative −10 and −35 elements for both PL and PR (Fig. 2) (23, 24)...

---

The pKO2 Linear Plasmid Prophage of *Klebsiella oxytoca*

Sherwood R. Casjens,1,2* Eddie B. Gilcrease,1 Wai Mun Huang,1 Kim L. Bunny,3
Marisa L. Pedulla,2,4 Michael E. Ford,2,4 Jennifer M. Houtz,2,4
Graham F. Hatfull,2,4 and Roger W. Hendrix;2,4

Department of Pathology, University of Utah Medical School, Salt Lake City, Utah 84132; Pittsburgh Bacteriophage Institute and Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260; and Section of Microbiology, University of California at Davis, Davis, California 95616.

* Corresponding author. Mailing address: Department of Pathology, University of Utah Medical School, Salt Lake City, UT 84132. Phone: (801) 581-5980. Fax: (801) 581-3607. E-mail: sherwood.casjens@path.utah.edu.

...The DNA sequence analysis software packages used were DNA Strider (27), GeneMark (8), the Staden programs (94), BLAST (3), BPROM http://www.softberry.com/berry.phtml?topic=bprom), and DNA Master (J. Lawrence [http://cobamide2.bio.pitt.edu/])....

*BMC Microbiology* 2004, 4:4

**Analysis of the lambdoid prophage element e14 in the *E. coli* K-12 genome**

Preeti Mehta1, Sherwood Casjens2 and Sankaran Krishnaswamy*1

Address: 1Bioinformatics Centre, School of Biotechnology, Madurai Kamaraj University, Madurai-625021, India and 2University of Utah Medical School, Department of Pathology, 90 North 1900 East, Salt Lake City UT 84132-2501, USA

Email: Preeti Mehta - mehta_p74@yahoo.com; Sherwood Casjens - sherwood.casjens@path.utah.edu;
Sankaran Krishnaswamy* - krishna@mrna.tn.nic.in

* Corresponding author

This article is available from: http://www.biomedcentral.com/1471-2180/4/4

... Putative promoters predicted using BPROM available at the website http://www.softberry.com. Scores are as given by BPROM. Promoters ...

*Plant Molecular Biology* 53 (6): 865-876, December 2003

**Prokaryotic orthologues of mitochondrial alternative oxidase and plastid terminal oxidase**

Allison E. McDonald
Department of Life Sciences and Department of Botany, University of Toronto at Scarborough, 1265 Military Trail, Scarborough, Ontario, M1C 1A4 Canada

Sasan Amirsadeghi
Department of Life Sciences and Department of Botany, University of Toronto at Scarborough, 1265 Military Trail, Scarborough, Ontario, M1C 1A4 Canada

Greg C. Vanlerberghe
Department of Life Sciences and Department of Botany, University of Toronto at Scarborough, 1265 Military Trail, Scarborough, *Ontario, M1C 1A4 Canada* (e-mail: gregv@utsc.utoronto.ca)

... The A. variabilis PTOX sequence was analyzed in the upstream region of the start codon with Softberry’s BPROM software (http://www.softberry.com). ...
Evaluation of five \textit{ab initio} gene prediction programs for the discovery of maize genes

Hong Yao$^{1,4}$, Ling Guo$^{1,6}$, Yan Fu$^{1,4}$, Lisa A. Borsuk$^{1,6}$, Tsui-Jung Wen$^2$, David S. Skibbe$^{1,5}$, Xiangqin Cui$^{1,4,9}$, Brian E. Scheffler$^8$, Jun Cao$^{1,4}$, Scott J. Emrich$^6$, Daniel A. Ashlock$^{3,6}$ and Patrick S. Schnable$^{1,2,4,5,6,7}$

(1) Department of Genetics, Development, and Cell Biology, Iowa State University, Ames, Iowa, 50011-3650
(2) Department of Agronomy, Iowa State University, Ames, Iowa, 50011-3650
(3) Department of Mathematics, Iowa State University, Ames, Iowa, 50011-3650
(4) Interdepartmental Graduate Programs in Genetics, Iowa State University, Ames, Iowa, 50011-3650
(5) Department of Molecular, Cellular and Developmental Biology, Iowa State University, Ames, Iowa, 50011-3650
(6) Department of Electrical and Computer Engineering and Department of Bioinformatics and Computational Biology, Iowa State University, Ames, Iowa, 50011-3650
(7) Center for Plant Genomics, Iowa State University, Ames, Iowa, 50011-3650
(8) Mid South Area Genomics Facility, USDA-ARS, Stoneville, MS 38776-0038, USA
(9) Present address: Department of Biostatistics, Birmingham, AL 35294, USA

... model FGENESH http://www.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind
Monocots Yes Yes GHMM a GeneMark ...

Finding short DNA motifs using permuted markov models

X Zhao, H Huang, TP Speed

... The data are human donor sequences from SpliceDB [9], a recently developed database of known mammalian splice site sequences (http://www.softberry.com/spldb ...

Current Opinion in Structural Biology 2004, 14:273–282

The evolving roles of alternative splicing

Liana F Lareau, Richard E Green1, Rajiv S Bhatnagar2,3 and Steven E Brenner1,2
Departments of Molecular and Cell Biology, and Plant and Microbial Biology, University of California, Berkeley, California 94720, USA
Department of Dermatology, University of California, San Francisco, California 94143, USA
_e-mail: brenner@compbio.berkeley.edu

... [79] SpliceDB http://www.softberry.com/berry.phtml?topic¼splicedb Database and composition statistics for mammalian splice sites inferred from ESTs [80] ...

Yearbook of Medical Informatics. Review Paper. 2004 121-136

Curated databases and their role in clinical bioinformatics

CC Englbrecht, M Han, MT Mader, A Osanger KFX Mayer
MIPS, Institute for Bioinformatics

Address of the authors:
Claudia C. Englbrecht, Michael Han,
SpliceDB: database of canonical and non-canonical mammalian splice sites

M. Burset, I. A. Seledtsov and V. V. Solovyev
The Sanger Centre, Hinxton, Cambridge CB10 1SA, UK and Softberry Inc., 108 Corporate Park Drive, Suite 120, White Plains, NY 10604, USA

* To whom correspondence should be addressed at present address: EOS Biotechnology, 225A Gateway Boulevard, South San Francisco, CA 94080, USA. Tel: +1 650 246 2331; Fax: +1 650 583 3881; Email: solovyev@eosbiotech.com Present address: M. Burset, Institut Municipal d’Investigació Mèdica (IMIM), C/Dr Aiguader 80, 08003 Barcelona, Spain
Minireview

**Web-based bioinformatic resources for protein and nucleic acids sequence alignment**

Kamel A. Abd-Elsalam
Molecular Markers Lab., Plant Pathology Research Institute, Agricultural Research Center, Orman 12619, Giza, Egypt.
E-mail: kaabdelsalam@msn.com

... 16-SCAN2:: program for aligning two multimegabyte-size sequences. http://www.softberry.com/berry.phtml?topic=scanh&prg=SCAN2. derived ...

**PromH: promoters identification using orthologous genomic sequences**

V. V. Solovyev* and I. A. Shahmuradov
Softberry Inc., 116 Radio Circle, Suite 400, Mount Kisco, NY 10549, USA
1 Institute of Botany, Azerbaijan National Academy of Sciences, 370073 Baku, Azerbaijan
*To whom correspondence should be addressed. Tel: +1 914 242 3592; Fax: +1 914 242 3593; Email: victor@softberry.com

Present address: I. A. Shahmuradov, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK

... The full-length sequences of gene pairs have been aligned by the **SCAN2** program (http://softberry.com/berry.phtml?topic=scanh&prg=SCAN2), which can align ...
Research article
Proteins with two SUMO-like domains in chromatin-associated complexes: The RENi (Rad60-Esc2-NIP45) family

Maria Novatchkova*1, Andreas Bachmair3, Birgit Eisenhaber2 and Frank Eisenhaber2

Address: 1Gregor Mendel-Institut GMI, Austrian Academy of Sciences, Vienna Biocenter, A-1030 Vienna, Austria, 2Research Institute of Molecular Pathology, Dr. Bohr-Gasse 7, A-1030 Vienna, Austria and 3Max Planck Institute for Plant Breeding Research, Carl-von-Linné-Weg 10, D-50829 Cologne, Germany

Email: Maria Novatchkova* - maria.novatchkova@gmi.oeaw.ac.at; Andreas Bachmair - bachmair@mpiz-koeln.mpg.de; Birgit Eisenhaber - b_eisen@imp.univie.ac.at; Frank Eisenhaber - Frank.Eisenhaber@imp.univie.ac.at

* Corresponding author

…Initial analysis of its sequence complexity shows that the disordered N-terminal half of the protein is followed by a likely globular segment (predicted using Pdisorder by Softberry, Inc)…
Molecular evolution of the insect chemoreceptor gene superfamily in Drosophila melanogaster

Hugh M. Robertson,*† Coral G. Warr,‡§ and John R. Carlson§

*Department of Entomology, University of Illinois, 505 South Goodwin Avenue, Urbana, IL 61801; ‡School of Biological Sciences, Monash University, Clayton VIC 3800, Australia; and §Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06520

The genes were reconstructed manually in the PAUP editor (23) by using the expected exon/intron structures as guides and the SPL program (Softberry, www.softberry.com/berry.phtml) to locate predicted introns.
NSITE

GENETIC ANALYSES OF BOVINE CARD15, A PUTATIVE DISEASE RESISTANCE GENE
A Dissertation by KRISTEN HAWKINS TAYLOR
Submitted to Texas A&M University in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY
May 2004
Major Subject: Genetics

…Short motifs identified as being conserved between the three species in these intronic regions as well as in the 5’UTR and 3’UTR were then analyzed using the TFSCAN (http://zeon.well.ox.ac.uk/git-bin/tfscan) and NSITE (available through SoftBerry http://www.softberry.com/berry.phtml?topic=promoter) programs to identify putative regulatory motifs. Motifs selected for analysis required homology consisting of 6 or more bases with no more than 2 substitutions among the 3 species…

…Sequence that included the SNPs located within intronic regions and in the 3’ and 5’UTRs were analyzed using NSITE (available through SoftBerry at http://www.softberry.com/berry.phtml?topic=promoter) to identify putative regulatory motifs…
Cis-regulatory element based targeted gene finding: genome-wide identification of abscisic acid- and abiotic stress-responsive genes in Arabidopsis thaliana

Weixiong Zhang $^{1,2,*}$, Jianhua Ruan $^1$, Tuan-hua David Ho $^3$, Youngsook You $^3$, Taotao Yu $^1$ and Ralph S. Quatrano $^3$

$^1$Department of Computer Science and Engineering, Washington University in Saint Louis Saint Louis, MO 63130, USA  
$^2$Department of Genetics, Washington University in Saint Louis Saint Louis, MO 63130, USA  
$^3$Department of Biology, Washington University in Saint Louis Saint Louis, MO 63130, USA  

*To whom correspondence should be addressed.

To predict TSSs, we combined an A.thaliana cDNA database and a software, TSSP (SoftBerry, http://www.softberry.com). As ...
PLANTPROM


Genome wide analysis of Arabidopsis core promoters

Carlos Molina$^{1,2}$ and Erich Grotewold$^2$
1Department of Plant Cellular and Molecular Biology and Plant Biotechnology Center, The Ohio State University, Columbus, OH 43210
2Departamento de Informática, Universidad Técnica Federico Santa María, Valparaíso, Chile

... search for TATA elements is carried out on the 12,749 [-500, -1] regions, 6,316 sequences (using the MEME NFM) or 8,776 (using the expanded PlantProm NFM) are ...

Nucleic Acids Research, 2005, Vol. 33, No. 3 1069–1076
doi:10.1093/nar/gki247

Plant promoter prediction with confidence estimation

I. A. Shahmuradov$^1$, V. V. Solovyev$^{1,2,*}$ and A. J. Gammerman$^1$
1Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK and 2Softberry Inc., 116 Radio Circle, Suite 400, Mount Kisco, NY 10549, USA

... MATERIALS AND METHODS Training and testing sequences For training and testing procedures, we used 301 promoters with annotated TSS from PlantProm DB (22). ...

DOI: 10.1093/nar/gkh017

AthaMap: an online resource for in silico transcription factor binding sites in the Arabidopsis thaliana genome

Nils Ole Steffens, Claudia Galuschka, Martin Schindler, Lorenz BuÈlow and Reinhard Hehl$^*$
Institut fuÈ r Genetik, Technische UniversitaÈ t Braunschweig, Spielmannstraûe 7, D-38106 Braunschweig, Germany

... Shahmuradov,IA, Gammerman,AJ, Hancock,JM, Bramley,PM and Solovyev,VV (2003) PlantProm: a database of plant promoter sequences. ...


GENOME ANALYSIS
Utility of Different Gene Enrichment Approaches Toward Identifying and Sequencing the Maize Gene Space$^{1,[w]}

Nathan Michael Springer$^*$, Xiequn Xu and W. Brad Barbazuk$^*$
Center for Plant and Microbial Genomics, Department of Plant Biology, University of Minnesota, St. Paul, Minnesota 55108 (N.M.S.); and Donald Danforth Plant Sciences Center, St. Louis, Missouri 63132 (X.X., W.B.B.)

... Text], Shahmuradov IA, Gammerman AJ, Hancock JM, Bramley PM, Solovyev VV (2003) PlantProm: a database of plant promoter sequences. ...
Characterization of Three Functional High-Affinity Ammonium Transporters in Lotus japonicus with Differential Transcriptional Regulation and Spatial Expression

Enrica D’Apuzzo2, Alessandra Rogato2, Ulrike Simon-Rosin, Hicham El Alaoui3, Ani Barbulova, Marco Betti, Maria Dimou, Panagiotis Katinakis, Antonio Marquez, Anne-Marie Marini, Michael K. Udvardi, and Maurizio Chiurazzi*

Institute of Genetics and Biophysics, Via Marconi 12, 80125 Napoli, Italy (E.D., A.R., H.E.A., A.B., M.C.); Molecular Plant Nutrition Group, Max Planck Institute of Molecular Plant Physiology, Am Muhlenberg 1, 14476 Golm, Germany (U.-S.-R., M.K.U.); Université Libre de Bruxelles, Institut de Biologie de Médecine Moléculaire, 6041 Gosselies, Belgium (A.-M.M.); Agricultural University of Athens, Department of Agricultural Biotechnology, 11855 Athens, Greece (M.D., P.K.); and Departamento de Bioquímica Vegetal y Biología Molecular, Facultad de Química, 41080 Seville, Spain (M.B., A.M.)


PlantProm: a database of plant promoter sequences

Authors: Bramley P.M.1; Solovyev V.V.2; Shahmuradov I.A.; Gammerman A.J.; Hancock J.M.
Affiliations: Department of Computer Science, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK: 1: School of Biological Sciences, Royal Holloway, University of London, UK 2: Softberry Inc., 116 Radio Circle, Suite 400, Mount Kisco, NY 10549, USA 3: *To whom correspondence should be addressed. Email: victor@softberry.com, Present address: John M. Hancock, MRC Mammalian Genetics Unit, Harwell, Oxfordshire, UK
PromH: promoters identification using orthologous genomic sequences

V. V. Solovyev* and I. A. Shahmuradov
Softberry Inc., 116 Radio Circle, Suite 400, Mount Kisco, NY 10549, USA
Institute of Botany, Azerbaijan National Academy of Sciences, 370073 Baku, Azerbaijan
*To whom correspondence should be addressed. Tel: +1 914 242 3592; Fax: +1 914 242 3593; Email: victor@softberry.com
Present address: I. A. Shahmuradov, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK
Received February 15, 2003; Revised and Accepted March 21, 2003
Gene conversion: a mechanism for generation of heterogeneity in the \textit{tprK} gene of \textit{Treponema pallidum} during infection

Arturo Centurion-Lara*, Rebecca E. LaFond, Karin Hevner, Charmie Godornes, Barbara J. Molini, Wesley C. Van Voorhis and Sheila A. Lukehart

Affiliations: Departments of Medicine and Pathobiology, University of Washington, Harborview Medical Center, Box 359779, 325 Ninth Ave., Seattle, WA 98104, USA.

*E-mail acentur@u.washington.edu; Tel. (+1) 206 341 5364; Fax (+1) 206 341 5363.

... Using the fgenesb program, which identifies putative operons and genes in microbial genomes (\textit{Softberry}; http://www.softberry.com/ berry.phtml ), the tprK ORF ...

---

**Microbiology** 150 (2004), 518-520; DOI 10.1099/mic.0.26871-0

IVET experiments in \textit{Pseudomonas fluorescens} reveal cryptic promoters at loci associated with recognizable overlapping genes

Mark W. Silby\textsuperscript{1}, Paul B. Rainey\textsuperscript{2,3} and Stuart B. Levy\textsuperscript{1,4}

\textsuperscript{1} Center for Adaptation Genetics and Drug Resistance, Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, MA 02111, USA
\textsuperscript{2} Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK
\textsuperscript{3} School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand
\textsuperscript{4} Department of Medicine, Tufts University School of Medicine, Boston, MA 02111, USA

Correspondence Stuart B. Levy (stuart.levy@tufts.edu)

...Using SoftBerry software (http://www.softberry.com/berry.phtml), -35 and -10 boxes and a transcriptional start site were predicted 84, 60 and 44 bp upstream of the \textit{iiv5} ORF, respectively...

---

DOI: 10.1128/JVI.78.22.12576-12590.2004

**Functional Genomics Analysis of Singapore Grouper Iridovirus: Complete Sequence Determination and Proteomic Analysis**

Wen Jun Song,\textsuperscript{1} Qi Wei Qin,\textsuperscript{2} Jin Qiu,\textsuperscript{1} Can Hua Huang,\textsuperscript{1} Fan Wang,\textsuperscript{1} and Choy Leong Hew\textsuperscript{1,*}

Department of Biological Sciences,\textsuperscript{1} Tropical Marine Science Institute, National University of Singapore, Singapore

* Corresponding author. Mailing address: Department of Biological Sciences, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260, Singapore. Phone: 65-68742692. Fax: 65-67795671. E-mail: dbshewcl@nus.edu.sg or dbshead@nus.edu.sg.

Received 19 March 2004/ Accepted 29 June 2004
The whole genome was also submitted to [http://www.softberry.com](http://www.softberry.com) (Softberry Inc., Mount Kisco, N.Y.) for identification of all potential ORFs....

Coding capacity of the viral genomic DNA sequence. Prediction of presumptive genes was carried out by using the viral gene prediction program under the website [http://www.softberry.com](http://www.softberry.com) supplemented with Vector NTI suite 7.1....

---

**Linkage disequilibrium and sequence diversity in a 500-kbp region around the adh1 locus in elite maize germplasm**

Mark Jung, Ada Ching, Dinakar Bhattramakki, Maureen Dolan, Scott Tinge, Michele Morgante, and Antoni Rafalski

(1) DuPont Crop Genetics, Experimental Station, P.O. Box 80353, Wilmington, DE 19880-0353, USA
(2) Pioneer Hi-Bred International, Inc., 7300 NW 62nd Avenue, P.O. Box 1004, Johnston, IA 50131-1004, USA
(3) Dipartimento di Produzione Vegetale e Tecnologie Agrarie, Universita di Udine, Via delle Scienze 208, 33100 Udine, Italy

Received: 26 January 2004 Accepted: 2 April 2004 Published online: 6 August 2004

**Mark Jung** Email: mark.t.jung@cgr.dupont.com

... 1). Gene locations were defined by several methods. Annotations provided in Tikhonov et al. (1999) were first used, then FGENESH gene-finding software ...

---

**Isolation, Characterization and Expression Analysis of a Leaf-specific Phosphoenolpyruvate Carboxylase Gene in Oryza sativa**

Chang-Fa Lin, Chun Wei, Li-Zhi Jiang, Ke-Gui Li, Xiao-Yin Qian, Kotb Attia, Jin-Shui Yang

A1 State Key Laboratory of Genetic Engineering Institute of Genetics, School of Life Sciences, Fudan University Shanghai 200433 P.R.China
A2 National Key Laboratory for Soil Erosion and Dry Land Farming on Loess Plateau Northwest Sci-Tech University of Agriculture and Forest Shannxi 712100 P.R.China

... tools of GeneMark (http://opal.biology.gatech.edu/geneMark/) and Softberry (http://www.softberry.com). For the isolation of putative ...

---

**Genome-Wide Analysis of the GRAS Gene Family in Rice and Arabidopsis**

Chaoguang Tian, Ping Wan, Shouhong Sun, Jiayang Li, and Mingsheng Chen

(1) Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Datun Road, Chaoyang District, Beijing, 100101, China

*Mingsheng Chen Email: mschen@genetics.ac.cn*
Combining transcriptome data with genomic and cDNA sequence alignments to make confident functional assignments for *Aspergillus nidulans* genes

Andrew H. SIMS  
Manda E. GENT  
Geoffrey D. ROBSON  
Nigel S. DUNN-COLEMAN  
Stephen G. OLIVER

**School of Biological Sciences, University of Manchester, The Michael Smith Building, Oxford Road, Manchester M13 9PT, UK. E-mail: steve.oliver@man.ac.uk**

**Genencor International Inc., 925 Page Mill Road, Palo Alto, CA 94304, USA.**

Combining transcriptome data with genomic and cDNA sequence alignments to make confident functional assignments for *Aspergillus nidulans* genes.

...Kingdom. Page 2. Genewise, *FgeneSH*, *FgeneSH+* consisting of 9541 putative open reading frames (ORFs) was released in June 2003. We...

TAG Theoretical and Applied Genetics

**DOI:** 10.1007/s00122-004-1603-2  
**Issue:** Volume 109, Number 1  
**Date:** June 2004  
**Pages:** 129 – 139

Gene content and density in banana (*Musa acuminata*) as revealed by genomic sequencing of BAC clones

R. Aert  
L. Sági  
G. Volckaert

...gsc.riken.go.jp), *fgenesh* version 1.1 (Salamov and Solovyev 2000; http://www.softberry.com), genemark.hmm version 2.2a (Lukashin and Borodovsky 1998; http ...

**Genome Research** 14:2503-2509, 2004

**Resources**

EAnnot: A genome annotation tool using experimental evidence

Li Ding, Aniko Sabo, Nicolas Berkowicz, Rekha R. Meyer, Yoram Shotland, Mark R. Johnson, Kymberlie H. Pepin, Richard K. Wilson and John Spieth

**Genome Sequencing Center, Washington University School of Medicine, St. Louis, Missouri 63110, USA**

Article and publication are at http://www.genome.org/cgi/doi/10.1101/gr.3152604.

**Corresponding author.**

**E-mail** lding@watson.wustl.edu; fax (314) 286-1810.

...Some ab initio programs, such as Genscan (Burge and Karlin 1997) and *Fgenesh* (Salamov and Solovyev 2000) are based on intrinsic characteristics of coding ...

...Some ab initio programs, such as Genscan (Burge and Karlin 1997) and Fgenesh (Salamov and Solovyev 2000) are based on intrinsic characteristics of coding sequence (e.g., codon usage, consensus splice sites, etc.) and require training on known genes from the organism…
…To further evaluate the performance of EAnnot, we compared EAnnot predictions with Ensembl, Genscan, and Fgenesh predictions using manual annotation as a standard. While Genscan and Fgenesh are ab initio programs, Ensembl takes into account experimental data, a feature shared with EAnnot. Ensembl predicted 1037 known genes with 1798 transcripts and 1457 EST genes with 2308 transcripts for chromosome 6 (build 31), while Fgenesh and Genscan predicted 6230 and 6225 genes, respectively. We evaluated the performance of each program with respect to splice sites, transcripts, and genes across all of chromosome 6.…
We have found that the standard gene-discovery programs FGENESH, GeneMark and GENSCAN annotate segments of most retrotransposons and many inverted-repeat transposable elements as genes. Using FGENESH to annotate maize BAC clones, for instance, 70–100% of the predicted genes are actually from transposable elements...


**Spotted leaf11, a Negative Regulator of Plant Cell Death and Defense, Encodes a U-Box/Armadillo Repeat Protein Endowed with E3 Ubiquitin Ligase Activity**

Li-Rong Zeng\(^a\), Shaohong Qu\(^a\), Alicia Bordeos\(^b\), Chengwei Yang\(^c\), Marietta Baraoidan\(^b\), Hongyan Yan\(^a\), Qi Xié\(^a\), Baek Hie Nahm\(^d\), Hei Leung\(^b\) and Guo-Liang Wang\(^a\,\(^1\)

\(^a\) Department of Plant Pathology, Ohio State University, Columbus, Ohio 43210
\(^b\) International Rice Research Institute, Metro Manila, Philippines
\(^c\) State Key Lab for Biocontrol, Sun Yat-sen (Zhongshan) University, Guangzhou, China 510275
\(^d\) Department of Biological Science, Myongji University, Kyonggido, Korea 449728

\(^1\) To whom correspondence should be addressed. E-mail wang.620@osu.edu; fax 614-292-4455

... in spl11. Exons predicted in G3 by the programs GENSCAN and Fgenesh using different matrixes are displayed in dark gray. (D) RFLP ...

---

**Source:** Human Genomics, Volume 1, Number 2, January 2004, pp. 146-149(4)

**Publisher:** Henry Stewart Publications

The truth about mouse, human, worms and yeast

**Authors:** David R. Nelson\(^1\); Daniel W. Nebert\(^2\)

\(^1\): Department of Molecular Sciences and The UT Center of Excellence in Genomics and Bioinformatics, University of Tennessee, Memphis, Tennessee 38163, USA  
\(^2\): Department of Environmental Health and Center for Environmental Genetics (CEG), University of Cincinnati Medical Center, Cincinnati, Ohio 45267-0056, USA

... unpublished data, 2003; see also Ref. [7], FGENESH, 21 TWINSCAN 22 and the Ensembl annotation pipeline. 23 The output of the four ...

---


A comprehensive transcript index of the human genome generated using microarrays and computational approaches

Eric E Schadt\(^1\), Stephen W Edwards\(^*\,\(^1\), Debraj GuhaThakurta\(^1\), Dan Holder\(^2\), Lisa Ying\(^2\), Vladimir Svetnik\(^2\), Amy Leonardson\(^1\), Kyle W Hart\(^3\), Archie Russell\(^1\), Guoya Li\(^1\), Guy Cavet\(^1\), John Castle\(^1\), Paul McDonagh\(^4\), Zhengyan Kan\(^1\), Ronghua Chen\(^1\), Andrew Kasarskis\(^1\), Mihai Margarint\(^1\), Ramon M Caceres\(^1\), Jason M Johnson\(^1\), Christopher D Armour\(^1\), Philip W Garrett-Engele\(^1\), Nicholas F Tsinoremas\(^5\) and Daniel D Shoemaker\(^1\)
…GrailEXP 4.0 [47], GENSCAN 1.0 [48], FGENESH [49], and FGENESH+ [49] *ab initio* gene-prediction algorithms were run independently across the entire genome assembly to augment alignment-based gene identification methods. GrailEXP 4.0, GENSCAN 1.0, and FGENESH version 1.c were run with default parameters for human sequence. GrailEXP used expressed sequence evidence from RefSeq, UniGene and DoubleTwist HGI to refine gene predictions. FGENESH+ was run with protein sequences from BLASTX with E-score lower than 10^{-5}. When multiple protein alignments overlapped, all overlapping protein sequences were clustered with BLASTClust [50] and the lowest E-score hit was used by FGENESH+…

…Additionally, 3' sim4 and 3' or 5' FGENESH+ predicted exons that were short and/or distant from internal predicted exons were removed…

*Genome Research* 14:988-995, 2004

**ENSEMBL Special**

*GeneWise and GeneWise*

Ewan Birney1,3, Michele Clamp2 and Richard Durbin2
1 *The European Bioinformatics Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK;* 2 *The Wellcome Trust Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK*

…There has been a long history of successful *ab initio* programs which do not use any additional evidence to predict genes on genomic DNA, of which Genscan (Burge and Karlin 1997•) and Fgenes (Solovyev and Salamov 1997•) are two of the most successful cases.…

…Another class of evidence-based gene prediction programs are ones which use external evidence to influence the scoring of potential exons, including SGP-2 (Parra et al. 2003•), Genie (Kulp et al. 1996•), Genomescan (Yeh et al. 2001•), HMMGene (Krogh 2000•), and Fgenesh++ (Solovyev and Salamov 1997•)…

Published online before print February 5, 2004, 10.1073/pnas.0308430100

*PNAS* | February 17, 2004 | vol. 101 | no. 7 | 1910-1915

Type I MADS-box genes have experienced faster birth-and-death evolution than type II MADS-box genes in angiosperms

Jongmin Nam††, Joonyul Kim‡‡, Shinyoung Lee‡, Gynheung An§, Hong Ma† and Masatoshi Nei†

†Institute of Molecular Evolutionary Genetics and Department of Biology, Pennsylvania State University, University Park, PA 16802; and §National Research Laboratory of Plant Functional Genomics, Division of Molecular and Life Sciences, Pohang University of Science and Technology, Pohang 790-784, Korea

Contributed by Masatoshi Nei, December 22, 2003
Because annotation of rice gene was still in progress at the time of this study, we ourselves conducted gene annotation by using the computer program **FGENESH** (www.softberry.com) from the genome sequences obtained from TIGR and the Rice Genome Database (China) (25).

---

**Sequence analysis of the long arm of rice chromosome 11 for rice–wheat synteny**

Nagendra K. Singh¹, Saurabh Raghuvarsri², Subodh K. Srivastava¹, Anupama Gaur², Ajit K. Pal¹, Vivek Dalal¹, Archana Singh¹, Irfan A. Ghazi¹, Ashutosh Bhargav¹, Mahavir Yadav¹, Anupam Dixit¹, Kamlesh Batra¹, Kishor Gaikwad¹, Tilak R. Sharma¹, Amitabh Mohanty², Arvind K. Bharti³, Anita Kapur³, Vikrant Gupta³, Dibyendu Kumar³, Shubha Vij², Ravi Vyedianathan², Parul Khurana², Sulabha Sharma², W. Richard McCombie³, Joachim Messing⁴, Rod Wing⁵, Takuji Sasaki⁶, Paramjit Khurana², Trilochan Mohapatra¹, Jitendra P. Khurana² and Akhilesh K. Tyagi²

1. Indian Initiative for Rice Genome Sequencing, National Research Centre on Plant Biotechnology, Indian Agricultural Research Institute, New Delhi, 110012, India
2. Indian Initiative for Rice Genome Sequencing, Department of Plant Molecular Biology, University of Delhi South Campus, New Delhi, 110021, India
3. Genome Research Centre, Cold Spring Harbor Laboratory, 1 Bungtown Road, New York, USA
4. The Plant Genome Initiative at Rutgers–Waksman Institute, Rutgers University, 190 Frelinghuysen Road, Piscataway, NJ 08873, USA
5. Department of Plant Sciences, 303 Forbes Building, Arizona Genomics Institute, The University of Arizona, Tucson, AZ 85721, USA
6. Department of Genome Research, National Institute of Agrobiological Sciences, 1-2, Kannondai 2-chome, Tsukuba, Ibaraki, 305-8602, Japan

**Akhilesh K. Tyagi Email:** akhilesh@genomeindia.org

... Wherever RiceGAAS data were not available, the genes were predicted by **FGENESH** trained for monocot plant species (http://www.softberry.com/berry.phtml). ...

---

**Annotation of a 95-kb *Populus deltoides* genomic sequence reveals a disease resistance gene cluster and novel class I and class II transposable elements**

M. Lescot¹, 4, S. Rombauts¹, J. Zhang¹, S. Aubourg¹, 5, C. Mathé¹, 6, S. Jansson², P. Rouzé¹, 3, and W. Boerjan¹

1. Department of Plant Systems Biology, Flanders Interuniversity Institute for Biotechnology, Ghent University, Technologypark 927, 9052 Gent, Belgium
2. Department of Plant Physiology, University of Umeå, 901 87 Umeå, Sweden
3. Laboratoire Associé de l’Institut National de la Recherche Agronomique (France), Ghent University, 9052 Gent, Belgium
4. **Present address:** CIRAD-Biotrop, TA40/03, 34398 Montpellier Cedex 5, France
5. **Present address:** Unité de Recherche en Génomique Végétale, INRA, 91057 Evry Cedex, France
6. **Present address:** Laboratoire de Biologie Vasculaire, Institut de Pharmacologie et Biologie Structurale, 205 route de Narbonne, 31077 Toulouse Cedex, France
P. Rouzé Email: pierre.rouze@psb.ugent.be

1999; http://www.tigr.org/tdb/glimmerm/glmr_form.html), and FgenesH for dicots or monocots (Salamov and Solovyev 2000; http://www.softberry.com/).

BIOINFORMATICS, 2004 vol.20 N.9 p.1416-1427

J Yuan, B Bush, A Elbrecht, Y Liu, T Zhang, W Zhao … -

... suchasGRAIL (Lopezetal., 1994; Roberts, 1991; Uberbacher et al., 1996), GENESCOPE (Murakami and Takagi, 1998), fgenesh (Salamov and Solovyev, 2000), GeneMark ...

Molecular Plant Pathology Volume 5 Issue 6 Page 515 - November 2004

Pathogen profile
Heading for disaster: Fusarium graminearum on cereal crops

RUBELLA S. GOSWAMI AND H. CORBY KISTLER*
*Correspondence: E-mail: HCKIST@umn.edu

... This pipeline uses a combination of the programs FGENESH and FGENESH+ (Salamov and Solovyev, 2000) modified by Softberry ( http://www.softberry.com ) with ...


MIPS: analysis and annotation of proteins from whole genomes

H. W. Mewes*1,2, C. Amid1, R. Arnold1, D. Frishman2, U. Güldener1, G. Mannhaupt2, M. Münsterkötter1, P. Pagel1, N. Strack2, V. Stümpflen1, J. Warfsmann1 and A. Ruepp1
1 Institute for Bioinformatics (MIPS), GSF National Research Center for Environment and Health, Ingolstaedter Landstrasse 1, D-85764 Neuherberg, Germany and 2 Technische Universität München, Chair of Genome Oriented Bioinformatics, Center of Life and Food Science, D-85350 Freising-Weißenstephan, Germany
*To whom correspondence should be addressed at Institute for Bioinformatics (MIPS), GSF National Research Center for Environment and Health, Ingolstädter Landstraße 1, D-85764 Neuherberg, Germany. Tel: +49 89 3187 3580; Fax: +49 89 3187 3585; Email: w.mewes@gsf.de

... The genome of 40 Mb encodes 10 000 proteins automatically predicted by the program FGENESH (http://softberry. com), specifically trained for Neurospora. ...

Annual Review of Genomics and Human Genetics
Vol. 5: 15-56 (Volume publication date September 2004)

COMPARATIVE GENOMICS

Webb Miller, Kateryna D. Makova, Anton Nekrutenko, and Ross C. Hardison
The Center for Comparative Genomics and Bioinformatics, The Huck Institutes of Life Sciences, and the Departments of Biology, Computer Science and Engineering, and Biochemistry and Molecular Biology, Pennsylvania State University, University Park, Pennsylvania; email: webb@bx.psu.edu, kdm16@psu.edu, anton@bx.psu.edu, rch8@psu.edu

... These algorithms include Genscan, the most popular gene prediction tool (24), GenMark (117), FgenesH (155), GeneID (144), and others (for an excellent overview ...

DNA and Cell Biology May 2004, Vol. 23, No. 5: 311-324

Harbinger Transposons and an Ancient HARBI1 Gene Derived from a Transposase
We used FGENESH (Salamov and Solovyev, 2000) and GeneScan (Burge and Karlin, 1997) for the identification of exons and introns. The $dN/dS$ ratio,

\begin{align*}
\text{Proc Natl Acad Sci U S A.} & \quad 2004 \text{ February 17; 101}(7): 1910–1915. \\
\text{Published online 2004 February 5. doi: 10.1073/pnas.0308430100.}
\end{align*}

Type I MADS-box genes have experienced faster birth-and-death evolution than type II MADS-box genes in angiosperms

Jongmin Nam, Joonyul Kim, Shinyoung Lee, Gynheung An, Hong Ma, and Masatoshi Nei

\begin{align*}
\text{† Institute of Molecular Evolutionary Genetics and Department of Biology, Pennsylvania State University, University Park, PA 16802; and § National Research Laboratory of Plant Functional Genomics, Division of Molecular and Life Sciences, Pohang University of Science and Technology, Pohang 790-784, Korea}
\end{align*}

\begin{align*}
\text{† To whom correspondence should be addressed. E-mail: jyn101@psu.edu.}
\end{align*}

\begin{align*}
\text{§ Present address: Michigan State University–Department of Energy Plant Research Laboratory and Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824.}
\end{align*}

... of rice gene was still in progress at the time of this study, we ourselves conducted gene annotation by using the computer program fgenesh (www.softberry.com ...
Plant Molecular Biology
DOI: 10.1023/B:PLAN.0000028768.21587.dc
Issue: Volume 54, Number 1 Date: January 2004 Pages: 55 - 69

A total of 54,397 putative genes could be predicted for the rice genome from this data set (Table 3) using the FGENESH program with the default setting for ...

Current Genetics
DOI: 10.1007/s00294-003-0451-y
Issue: Volume 44, Number 6 Date: January 2004 Pages: 329 - 338

Chromosome rearrangements in isolates that escape from het-c heterokaryon incompatibility in Neurospora crassa

Qijun Xiang and N. Louise Glass
Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720-3102, USA

... Hypothetical proteins are predicted from FGENESH calls with overlapping Blastx hits (but not with trusted homology), while Predicted ...

Molecular Genetics and Genomics
DOI: 10.1007/s00438-004-0990-z
Issue: Volume 271, Number 4 Date: May 2004 Pages: 402 - 415

Genome-wide identification of NBS genes in japonica rice reveals significant expansion of divergent non-TIR NBS-LRR genes

1. State Key Laboratory of Pharmaceutical Biotechnology, Department of Biology, Nanjing University, 210093 Nanjing, China
... to 5000–10,000 bp from both ends of the hits, and then the expanded nucleotide fragments were reannotated using the gene-finding programs FGENESH (http://www ...)

**Genetic control of branching in foxtail millet**

Andrew N. Doust,* Katrien M. Devos,‡† Michael D. Gadberry,*‡ Mike D. Gale,‡ and Elizabeth A. Kellogg*

*Department of Biology, University of Missouri, 8001 Natural Bridge Road, St Louis, MO 63121; and †John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, United Kingdom

‡ Present address: Departments of Crop and Soil Sciences and Plant Biology, Miller Plant Sciences Building, University of Georgia, Athens, GA 30602.

Communicated by Peter H. Raven, Missouri Botanical Garden, St. Louis, MO, April 23, 2004

... Each of these contigs was scanned by using **fgenesh** (28), and identified ORFs were translated and compared with ORFs from other contigs from the same QTL region ...

**Mol. Biol. Evol.** 21(9):1769-1780. 2004

**Merlin, a New Superfamily of DNA Transposons Identified in Diverse Animal Genomes and Related to Bacterial IS1016 Insertion Sequences**

Cédric Feschotte†

Departments of Plant Biology and Genetics, The University of Georgia, Athens

Correspondence: E-mail: cedric@plantbio.uga.edu.

... coding sequences were assembled by removing introns predicted with more than 85% confidence by NetGene2 (http://www.cbs.dtu.dk) and/or **FGENESH** (http://genomic ...)

**Genome Research** 14:1924–1931 ©2004

**Gene Loss and Movement in the Maize Genome**

Jinsheng Lai,1 Jianxin Ma,2,3 Zuzana Swigonová,4 Wusirika Ramakrishna,2,4 Eric Linton,1,5 Victor Llacac,1,6 Bahattin Tanyolac,1,7 Yong-Jin Park,2,8 O-Young Jeong,2,9 Jeffrey L. Bennetzen,2,3 and Joachim Messing1,10

1Waksman Institute of Microbiology, Rutgers University, Piscataway, New Jersey 08854-8020, USA; 2Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907-1392, USA

Present addresses: 3Department of Genetics, University of Georgia, Athens, GA 30602-7223, USA; 4Department of Biological Sciences, 740 DOW, Michigan Tech University, Houghton, MI 49931, USA; 5Plant Biology Labs, Michigan State University, East Lansing, MI 48824, USA; 6Analytical and Genomic Technologies, Crop Genetics R&D, DuPont Agriculture & Nutrition, Wilmington, DE 19880-0353, USA; 7Department of Bioengineering, Ege University, Izmir, 35100 Turkey; 8National Institute of Agricultural Biotechnology, Suwon 441-707, Republic of Korea; 9National Institute of Crop Science, Suwon 441-857, Republic of Korea.

10Corresponding author. E-MAIL mess@waksman.rutgers.edu; FAX (732) 445-0072.

... The **FGENESH** program predicted four, of which three (gene 1d in the maize orp1 region; gene 5a, 5b in the rice r1 region) would produce truncated proteins ...
Systematic genome-wide approach to positional candidate cloning for identification of novel human disease genes

H. Kiyosawa 1, T. Kawashima 2, D. Silva 3, N. Petrovsky 4, Y. Hasegawa 2,5, K. Sakai 6 and Y. Hayashizaki 2,5,7

Affiliations: 1Technology and Development team for Mammalian Cellular Dynamics, Bioresource Center, RIKEN Tsukuba Institute, Tsukuba, Ibaraki, 2Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center, RIKEN Yokohama Institute and 3John Curtin School of Medical Research, Australian National University and 4Medical Informatics Centre, University of Canberra, Canberra, Australian Capital Territory, Australia 5Division of Genomic Information Resource Exploration, Science of Biological Supramolecular Systems, Yokohama City University, Graduate School of Integrated Science, Tsurumi-ku, Yokohama, Kanagawa, 6Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, Mishima, Shizuoka, and 7Genome Science Laboratory, RIKEN, Wako, Saitama, Japan,

Correspondence: Yoshihide Hayashizaki, Laboratory for Exploration Research Group, RIKEN Genomic Sciences Center, RIKEN Yokohama Institute, Suehiro-cho 1-7-22, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. Email: rgscerg@gsc.riken.go.jp

The data described in the present paper can be accessed on the web at http://genome.gsc.riken.go.jp/m/pcc/

Cryptococcus neoformans Kin1 protein kinase homologue, identified through a Caenorhabditis elegans screen, promotes virulence in mammals

Eleftherios Mylonakis 1, Alexander Idnurm 2, Roberto Moreno 1, Joseph El Khoury 1,3,4, James B. Rottman 2, Frederick M. Ausubel 6,7, Joseph Heitman 2,8,9,10 and Stephen B. Calderwood 1,11

1Division of Infectious Diseases, Massachusetts General Hospital, Boston, MA 02114, USA.
2Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA.
3Center for Immunology and Inflammatory Diseases, and
4Division of Rheumatology, Allergy, and Immunology, Massachusetts General Hospital, Boston, MA 02114, USA.
5Archemix Corporation, Cambridge, MA 02139, USA.
6Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA.
7Department of Genetics, Harvard Medical School, Boston, MA 02115, USA.
8Division of Infectious Diseases,
9Department of Medicine, and
10Howard Hughes Medical Institute, Duke University Medical Center, Durham, NC 27710, USA.
11Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA 02115, USA.

*E-mail scalderwood@partners.org
Sequences were compared with the H99 genome database at Duke University, and genes predicted in these regions by FGENESH software (http://www.softberry.com).

TAG Theoretical and Applied Genetics
DOI: 10.1007/s00122-003-1499-2
Issue: Volume 108, Number 5 Date: March 2004 Pages: 903 - 913

Original Paper
Characterization of soybean genomic features by analysis of its expressed sequence tags

Ai-Guo Tian¹, Jun Wang², Peng Cui², Yu-Jun Han², Hao Xu², Li-Juan Cong², Xian-Gang Huang², Xiao-Ling Wang², Yong-Zhi Jiao², Bang-Jun Wang¹, Yong-Jun Wang¹, Jin-Song Zhang¹ and Shou-Yi Chen¹

1. Plant Biotechnology Laboratory, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Datun Road, 100101 Beijing, China
2. Beijing Genomics Institute, Chinese Academy of Sciences, 101300 Beijing, China
Jin-Song Zhang Email: jszhang@genetics.ac.cn
Shou-Yi Chen Email: sychen@genetics.ac.cn

... six BAC-contig sequences of M. truncatula were analyzed and the results based on the gene prediction program FGENESH (Arabidopsis match/FGENESH prediction) (http ...

Current Proteomics, Volume 1, Number 1, January 2004, pp. 41-48(8)

Annotation of the Human Genome by High-Throughput Sequence Analysis of Naturally Occurring Proteins

Authors: McGowan S.J.¹; Terrett J.¹; Brown C.G.¹; Adam P.J.¹; Aldridge L.¹; Allen J.C.¹; Amess B.¹; Andrews K.A.¹; Barnes M.¹; Barnwell D.E.¹

... The polymorphic ‘hypothetical transcriptome’ was cons- tructed from transcripts predicted by FGENES, FGENESH (Softberry Inc, Mount Kisco, NY, USA), GENSCAN ...

Chinese Science Bulletin 2004 Vol. 49 No. 4 355-362

The VER2 promoter contains repeated sequences and requires vernalization for its activity in winter wheat (Triticum aestivum L.)

XU Wenzhong¹, WANG Xin¹, FENG Qi², ZHANG Lei², LIU Yaoguang³, HAN Bin², CHONG Kang¹, XU Zhihong & TAN Kehui
1. Research Center for Molecular Developmental Biology, Key Lab of Photosynthesis and Environmental Molecular Physiology, Institute of Botany, Chinese Academy of Sciences (CAS), Beijing 100093, China;
2. National Center for Gene Research, CAS, Shanghai 200233, China;
3. Genetic Engineering Laboratory, College of Life Sciences, South China Agricultural University, Guangzhou 510642, China
Correspondence should be addressed to Chong Kang (e-mail: mailto:chongk@ns.ibcas.ac.cn)

... Sequence analyses were finished using biological softwares on Internet, such as FGENESH 1.0 (Prediction of potential genes in Plant (Dct) genomic DNA). ...
Sequence variations of simple sequence repeats on chromosome-4 in two subspecies of the Asian cultivated rice

Can Li¹, Yu Zhang¹, Kai Ying¹, Xiaolei Liang¹ and Bin Han¹
¹National Center for Gene Research, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 500 Caobao Road, Shanghai 200233, China

... To characterize the possible relationship between SSRs and genes predicted by using FGENESH, we investigated the distribution of SSRs in the rice chromosome-4 ...

Long-range patterns of diversity and linkage disequilibrium surrounding the maize Y1 gene are indicative of an asymmetric selective sweep

Kelly Palaisa,* Michele Morgante,†‡ Scott Tingey,†§ and Antoni Rafalski*†§¶
¹DuPont Crop Genetics, Molecular Genetics Group, 1 Innovation Way, Newark, DE 19711; ²Department of Plant and Soil Sciences and Delaware Biotechnology Institute, University of Delaware, Newark, DE 19716; and ³Dipartimento di Produzione Vegetale e Tecnologie Agrarie, Università degli Studi di Udine, Via della Scienze 208, 33100 Udine, Italy
* To whom correspondence should be addressed. E-mail: j-antoni.rafalski@cgr.dupont.com.
† Present address: DuPont Crop Genetics Research, DuPont Experimental Station Building E353, Wilmington, DE 19880.

... The Y1 gene comprises a small uninterrupted gene island consisting of the Y1, an MLO homolog lying immediately downstream of Y1, and an fgenesh-predicted gene ...

Positional cloning of the rice Rf-1 gene, a restorer of BT-type cytoplasmic male sterility that encodes a mitochondria-targeting PPR protein

H. Akagi¹,², A. Nakamura², Y. Yokozeki-Misono², A. Inagaki²,³, H. Takahashi¹, K. Mori¹ and T. Fujimura³
¹Laboratory of Plant Breeding and Genetics, Department of Biological Production, Faculty of Bioresource Sciences, Akita Prefectural University, Kaidoubata-Nishi 241-7, Shimoshinjyo-Nakano, 010-0195 Akita, Japan
²Biochemical Technology Section, Life Science Laboratory, Performance Materials R&D Center, Mitsui Chemicals, Togo 1144, 297-0017 Mobara, Japan
³Institute of Agricultural and Forest Engineering, University of Tsukuba, Tennoudai 1-1-1, Tsukuba, 305-8572 Ibaraki, Japan

4. Present address: Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Yoshida-Honmachi, Sakyo-ku, 606-8501 Kyoto, Japan

H. Akagi Email: akagi@akita-pu.ac.jp

... Software Develop- ment, Tokyo). Genomic sequences were also analyzed using gene prediction programs, genescan and fgenesh. Table 1 DNA ...

Genome Research 14:942-950, 2004
Commonly, we use Genscan for ab initio prediction in human, mouse, and rat, but the system is equally applicable to other methods such as FgenesH (Solovyev et al.).

---

**Transcript abundance of rml1, encoding a putative GT1-like factor in rice, is up-regulated by *Magnaporthe grisea* and down-regulated by light**

Rong Wang a,b, Guofan Hong a,b, Bin Han a,*
National Center for Gene Research, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 500 Caobao Road, Shanghai 200233, China
Shanghai Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yueyang Road, Shanghai 200031, China

... 1A and 2A). The structure of rml1 a is as same as that predicted by FGENESH software. The 3V UTR of rml1 a is confirmed with the length of 596 bp. ...

---

**Identification of conserved gene structures and carboxy-terminal motifs in the Myb gene family of *Arabidopsis* and *Oryza sativa* L. ssp. *Indica***

Cizhong Jiang1, Xun Gu1, 2 and Thomas Peterson1
1Department of Genetics, Development and Cell Biology, and Department of Agronomy, Iowa State University, Ames, IA 50011, USA
2LHB Center for Bioinformatics and Biological Statistics, Iowa State University, Ames, IA 50011, USA

... FGeneSH has been used successfully to predict genes in rice [9], and GenScan was used together with it to predict genes by taking rice genomic sequences as ...

---

**The sirodesmin biosynthetic gene cluster of the plant pathogenic fungus *Leptosphaeria maculans***

Donald M. Gardiner1, Anton J. Cozijnsen1, Leanne M. Wilson2, M. Soledade C. Pedras2 and Barbara J. Howlett2
Analysis of the Molecular Evolutionary History of the Ascorbate Peroxidase Gene Family: Inferences from the Rice Genome

Felipe Karam Teixeira, Larissa Menezes-Benavente, Rogério Margis, Márcia Margis-Pinheiro

(1) Laboratório de Genética Molecular Vegetal, Departamento de Genética, UFRJ, 21944-970 Rio de Janeiro, Brasil
(2) Departamento de Bioquímica, Instituto de Química, UFRJ, 21944-970 Rio de Janeiro, Brasil

... Genomic sequences were also analyzed in the FGENESH gene structure prediction program (http://www.softberry.com/) and GeneMark program (http...
Inaugural Articles

Rapid recent growth and divergence of rice nuclear genomes

Jianxin Ma and Jeffrey L. Bennetzen*
Department of Genetics, University of Georgia, Athens, GA 30602
*To whom correspondence should be addressed. E-mail: maize@uga.edu.
Contributed by Jeffrey L. Bennetzen, May 25, 2004

... Almost all LTR-retrotransposons, including solo LTRs, identified in our studies were predicted as genes by the gene-finding program fgenesh (data not shown). ...

The Plant Journal
Volume 37 Issue 4 Page 517 -527 - February 2004
doi:10.1046/j.1365-313X.2003.01976.x

Xa26, a gene conferring resistance to Xanthomonas oryzae pv. oryzae in rice, encodes an LRR receptor kinase-like protein

Xinli Sun, Yinglong Cao, Zhifen Yang, Caiguo Xu, Xianghua Li, Shiping Wang; and Qifa Zhang
National Key Laboratory of Crop Genetic Improvement, National Center of Crop Molecular Breeding, Huazhong Agricultural University, Wuhan 430070, China
*For correspondence (fax +86 27 87287092; e-mail swang@mail.hzau.edu.cn)

... al., 1997). Gene prediction programs used were genscan (Burge and Karlin, 1997) and fgenesh (http://www.softberry.com). Promoter ...

Genome Research 14:1916–1923 ©2004 by Cold Spring Harbor Laboratory Press ISSN 1088-9051/04

Close Split of Sorghum and Maize Genome Progenitors

Zuzana Swigonová,1,6 Jinsheng Lai,1,6 Jianxin Ma,2,3 Wusirika Ramakrishna,2,4
Victor Llaca,1,5 Jeffrey L. Bennetzen,2,3 and Joachim Messing1,7
1Waksman Institute of Microbiology, Rutgers University, Piscataway, New Jersey 08854, USA;
2Department of Biological Sciences and Genetics Program, West Lafayette, Indiana 47907, USA
Present addresses: 3Department of Genetics, University of Georgia, Athens, GA 30602, USA;
4Department of Biological Sciences, Michigan Tech University, MI 49931, USA; 5Analytical and Genomic Technologies, Crop Genetics R&D, DuPont Agriculture & Nutrition, Wilmington, DE 19880, USA.
6These authors contributed equally to this work.
7Corresponding author.
E-MAIL messing@waksman.rutgers.edu; FAX (732) 445-0072.

Page 1. Close Split of Sorghum and Maize Genome Progenitors Zuzana Swigonová, 1,6 Jinsheng Lai, 1,6 Jianxin Ma, 2,3 Wusirika Ramakrishna, 2,4 ...
Physical mapping and putative candidate gene identification of a quantitative trait locus Ctb1 for cold tolerance at the booting stage of rice

K. Saito1, Y. Hayano-Saito1, W. Maruyama-Funatsuki1, Y. Sato1 and A. Kato1
(1) National Agricultural Research Center for Hokkaido Region, Hitsujigaoka 1, Toyohira, Sapporo, Hokkaido, 062-8555, Japan
K. Saito Email: kjsaito@affrc.go.jp

... GENSCAN, RICEHMM, FGENESH, MZEF), a splice prediction program (SPICEPREDICTOR), homology search analysis programs (BLAST, HMMER, ...

TAG Theoretical and Applied Genetics
DOI: 10.1007/s00122-004-1697-6
Issue: Volume 109, Number 4 Date: August 2004 Pages: 690 - 699

The anthracnose resistance locus Co-4 of common bean is located on chromosome 3 and contains putative disease resistance-related genes

M. Melotto1,4, M. F. Coelho1, A. Pedrosa-Harand2, J. D. Kelly3 and L. E. A. Camargo1
1. Departamento de Fitopatologia, Laboratório de Genética Molecular, ESALQ, Universidade de São Paulo, Piracicaba, SP, C.P. 9, 13418-900, Brazil
2. Department of Cell Biology and Genetics, Institute of Botany, University of Vienna, Rennweg 14, Vienna, 1030, Austria
3. Department of Crop and Soil Sciences, Michigan State University, East Lansing, MI 48824, USA
4. Present address: MSU-DOE Plant Research Laboratory, Michigan State University, 206 Plant Biology Building, East Lansing, MI 48824, USA
M. Melotto Email: melottom@msu.edu

... and Karlin 1997; http://genes.mit.edu/GENSCAN.html) and FGENESH (http://www.softberry.com/)—using Arabidopsis as the model organism. ...


Structural and functional analysis of rice genome

Department of Plant Molecular Biology, University of Delhi South Campus, Benito Juarez Road, New Delhi 110 021, India

... It integrates results from several gene prediction software such as GENSCAN (Burge and Karlin 1997), FGENESH (Sala- mov and Solovyev 2000), RiceHMM (Sakata ...


Comparative Analysis of the Receptor-Like Kinase Family in Arabidopsis and Rice

Shin-Han Shiua, Wojciech M. Karlowskiib, Runsun Pand, Yun-Huei Tzengac, Klaus F. X. Mayerb and Wen-Hsiung Li1,1
a Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637
b Munich Information Center for Protein Sequences/Institute of Bioinformatics, GSF, National Research Center for Environment and Health, Neuherberg 85764, Germany
c Department of Mathematics, National Tsing Hua University, Hsinchu, Taiwan 300
d Institute of Information Science, Academia Sinica, Taiwan 115
To whom correspondence should be addressed. E-mail whli@uchicago.edu; fax 773-702-9740.

... a permissive E value cutoff of 1. The rice genes from the indica subspecies was predicted using the whole genome shotgun assembly with FGENESH (Solovyev, 2002).

---


---

**Incongruent Patterns of Local and Global Genome Size Evolution in Cotton**

Corrinne E. Grover,¹ HyeRan Kim,² Rod A. Wing,² Andrew H. Paterson,³ and Jonathan F. Wendel,⁴¹

¹Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, Iowa 50011, USA; ²Arizona Genomics Institute, University of Arizona, Tucson, Arizona 85721, USA; ³Plant Genome Mapping Laboratory, University of Georgia, Athens, Georgia 30602, USA

... Potential genes were predicted by three independent programs: FGENESH (http://www.softberry.com/), ...
Angelike Stathopoulos¹, Bergin Tam¹, Matthew Ronshaugen¹, Manfred Frasch² and Michael Levine¹,³
¹Department of Molecular and Cell Biology, Division of Genetics & Development, University
of California, Berkeley, California 94720-3204, USA; ²Brookdale Department of Molecular
Cell and Developmental Biology, Mount Sinai School of Medicine, New York, New York
10029, USA

... FGF protein sequences used in alignment and phylogenetic reconstruction were gathered from
GenBank or inferred from genomic sequence using GENESCAN (Burge and Karlin 1997) and
FGENESH...

---

**Genome Research** 14:1888-1901, 2004

**Organization and Evolution of a Gene-Rich Region of the Mouse Genome: A 12.7-Mb
Region Deleted in the Del(13)Svea36H Mouse**

Ann-Marie Mallon¹,⁴, Laurens Wilming²,³, Joseph Weekes¹, James G.R. Gilbert², Jennifer
Ashurst¹, Sandrine Peyrefitte², Lucy Matthews², Matthew Cadman¹, Richard McKeone¹, Chris
A. Sellick¹, Ruth Arkell¹, Marc R.M. Botcherby³, Mark A. Strivens¹, R. Duncan Campbell³,
Simon Gregory²,⁵, Paul Denny¹, John M. Hancock²,⁶, Jane Rogers² and Steve D.M. Brown¹
¹Medical Research Council Mammalian Genetics Unit, Harwell, Oxfordshire, United Kingdom;
²Wellcome Trust Sanger Institute, Hinxton Genome Campus, United Kingdom; ³Medical
Research Council Rosalind Franklin Centre for Genomics Research, Hinxton Genome Campus,
United Kingdom

... Ab initio gene structures were predicted using FGENESH (Salamov and Solovyev 2000) and
GENSCAN...

---


**Annotation of the Human Genome by High-Throughput Sequence Analysis of Naturally
Occurring Proteins**

**Authors:** McGowan S.J.¹; Terrett J.¹; Brown C.G.¹; Adam P.J.¹; Aldridge L.¹; Allen J.C.¹;
Amess B.¹; Andrews K.A.¹; Barnes M.¹; Barnwell D.E.¹

**Affiliations:** ¹: Oxford GlycoSciences plc, The Forum, 86 Milton Park, Abingdon, OX14 4RY, UK.

... The polymorphic ‘hypothetical transcriptome’ was cons- tructed from transcripts predicted by
FGENES, FGENESH (Softberry Inc, Mount Kisco, NY, USA), GENSCAN...


**Long-range patterns of diversity and linkage disequilibrium surrounding the maize YI
gene are indicative of an asymmetric selective sweep**

Kelly Palaisa, * Michele Morgante, †‡ Scott Tingey, †‡ and Antoni Rafalski *†‡¶
*DuPont Crop Genetics, Molecular Genetics Group, 1 Innovation Way, Newark, DE 19711; "Department of Plant
and Soil Sciences and Delaware Biotechnology Institute, University of Delaware, Newark, DE 19716; and
‡Dipartimento di Produzione Vegetale e Tecnologie Agrarie, Università degli Studi di Udine, Via della Scienze 208,
33100 Udine, Italy

* To whom correspondence should be addressed. E-mail: j-antoni.rafalski@cgr.dupont.com.
The Y1 gene comprises a small uninterrupted gene island consisting of the Y1, an MLO homolog lying immediately downstream of Y1, and an fgenesh-predicted gene ...

Nucleic Acids Research, 2003, Vol. 31, No. 4 1148-1155

Characterization of Arabidopsis thaliana ortholog of the human breast cancer susceptibility gene 1: AtBRCA1, strongly induced by gamma rays

S. Lafarge and M.-H. Montané*
CEA Cadarache, DSV-DEVM, Laboratoire de Radiobiologie Végétale, Bat 185, F-13108 St Paul Lez Durance Cedex, France
*To whom correspondence should be addressed. Tel: +33 4 42 25 35 56; Fax: + 33 4 42 25 26 25; Email: marie-helene.montane@cea.fr
Received November 18, 2002; Accepted December 5, 2002
DDBJ/EMBL/GenBank accession no. AF515728.

…Gene structure prediction was done on software implemented on the Softberry web page (http://www.softberry.com/), analysis of protein domains using the SMART…
…The gene structure of At4g21070 was determined with three gene structure prediction software packages (Softberry, GenScan, Grail). …. To resolve this ambiguity in intron–exon prediction, we postulated the presence of two genes given by Softberry prediction software and performed northern blotting and 5’ RACE to characterize the structural organization of the At4g21070 locus…

Gene expression of a gene family in maize based on noncollinear haplotypes

Rentao Song and Joachim Messing*
Waksman Institute, Rutgers, The State University of New Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854-8020
Communicated by Brian A. Larkins, University of Arizona, Tucson, AZ, May 19, 2003, (received for review 2002 April 10)
* To whom correspondence should be addressed. E-mail: messing@mbcl.rutgers.edu.

.. The FGENESH program (Softberry, Mount Kisco, NY) was used for gene prediction analysis.

Gene discovery in the hamster: a comparative genomics approach for gene annotation by sequencing of hamster testis cDNAs

Sreedhar Oduru,1 Janee L Campbell,1 SriTulasi Karri,1 William J Hendry,3 Shafiq A Khan,1 and Simon C Williams2,1
1Department of Cell Biology & Biochemistry, Texas Tech University Health Sciences Center, Lubbock, Texas. USA
2Southwest Cancer Center at UMC, Lubbock, Texas, USA
3Department of Biological Sciences, Wichita State University, Wichita, Kansas, USA
... Two gene prediction programs were used **FGENESH** [http://www.softberry.com/berry.phtml](http://www.softberry.com/berry.phtml) and **GENEMARK** [http://opal.biology.gatech.edu/GeneMark/eukhmm.cgi?org=H...](http://opal.biology.gatech.edu/GeneMark/eukhmm.cgi?org=H)...
...We used FGENESH (ref. 18; www.softberry.com) for identifying genes encoded by TEs.

Iron homeostasis related genes in rice

Jeferson Gross¹, II; Ricardo José Stein II; Arthur Germano Fett-Neto I, II; Janette Palma Fett I, II
¹Universidade Federal do Rio Grande do Sul, Centro de Biotecnologia, Porto Alegre, RS, Brazil
IIUniversidade Federal do Rio Grande do Sul, Departamento de Botânica, Porto Alegre, RS, Brazil
³Botanical Institute of Ludwig-Maximilians-Universität, München, Germany

The prediction algorithms were GenScan (Burge and Karlin, 1997; http://genes.mit.edu/GENSCAN.html), GenomeScan (Burge and Karlin, 1997; http://genes.mit.edu/genomescan.html), FGENESH (Salamov and Solovyev, 2000; http://www.softberry.com/berry.phtml?topic=gfind), GeneMark.hmm (Borodovsky and Lukashin, unpublished; http://opal.biology.gatech.edu/GeneMark/eukhmm.cgi) and GrailEXP (Xu and Uberbacher, 1997; http://compbio.ornl.gov/grailexp/).

Characterization of the Maize Endosperm Transcriptome and Its Comparison to the Rice Genome

J Lai, N Dey, CS Kim, AK Bharti, S Rudd, KFX Mayer …

... for the rice genome from this data set (Table 3) using the FGENESH program with the default setting for monocotyledonous genes (http://www.softberry.com). ...
The ORF was identified by using FGENESH http://www.softberry.com/berry.phtml/ and GENSCAN http://genes.mit.edu/GENSCAN.html.

**Australasian Plant Pathology**
Volume 32 Number 4 2003 pp. 511-519

Small scale functional genomics of the blackleg fungus, *Leptosphaeria maculans*: analysis of a 38 kb region

Alexander Idnurm, Janet L. Taylor, M. Soledade C. Pedras and Barbara J. Howlett

... vertebrate and Arabidopsis settings; Burge and Karlin 1997) and FGENESH on Neurospora crassa and Schizosaccharomyces pombe settings (www.softberry.com), as ...

**Barley Genetics Newsletter** Volume 32
Hard-copy edition pages 34 - 37

**MAPPING AND SEQUENCING OF THE BARLEY PUTATIVE HYPERSENSITIVE INDUCED REACTION GENES**

Nils Rostoks\(^1\), David Kudrna\(^1\) and Andris Kleinhofs\(^1,2\)

\(^1\) Department of Crop and Soil Sciences, Washington State University, Pullman, WA 99164
\(^2\) School of Molecular Biosciences, Washington State University, Pullman, WA 99164

The full length coding sequence was reconstructed using a combination of FGENESH gene prediction program (http://www.softberry.com/) and alignment with cDNAs from the other barley HIR groups.

**TAG Theoretical and Applied Genetics**
DOI: 10.1007/s00294-003-0391-6
Issue: Volume 43, Number 5 Date: August 2003 Pages: 351 - 357

**Characterisation of the mating-type locus of the plant pathogenic ascomycete Leptosphaeria maculans**

Anton J. Cozijnsen\(^A1\) and Barbara J. Howlett \(^A1\)

\(^A1\) School of Botany The University of Melbourne 3010 Victoria Australia

... Genes, introns, exons and transcription initiation sites were predicted by analysis with FGENESH (www.softberry.com) on Neurospora crassa and...

Received May 23, 2003; Accepted September 23, 2003.

**Ds tagging of BRANCHED FLORETLESS 1 (BFL1) that mediates the transition from spikelet to floret meristem in rice (Oryza sativa L.)**

Qian-Hao Zhu\(^1,2\), Mohammad Shamsul Hoque\(^1,2\), Elizabeth S Dennis, \(^1,2\) and Narayana M Upadhyaya\(^2\)

\(^1\)CSIRO Plant Industry, GPO Box 1600, Canberra, ACT 2601, Australia
\(^2\)NSW Agricultural Genomics Centre, Wagga Wagga, Australia
...Analyses of 10-kb sequence of the ~30-kb contig 239 from the China Rice Genome database harboring the \textit{BFL1} locus by gene prediction programs \textbf{FGENESH} \url{http://www.softberry.com/berry.phtml/} and \textbf{GENSCAN} \url{http://genes.mit.edu/GENSCAN.html} identified a single-exon gene capable of encoding a protein with the DNA binding domain of the EREBP/AP2 family of plant transcription factors \cite{26,36}, 1515 bp downstream from the \textit{Ds} insertion.

.. The ORF was identified by using \textbf{FGENESH} \url{http://www.softberry.com/berry.phtml/} and \textbf{GENSCAN} \url{http://genes.mit.edu/GENSCAN.html}. Alignment of EREBP/AP2 domains was performed using programs of Genetics Computer Group Wisconsin software suit \cite{11}.

\section*{Map-Based Cloning of Leaf Rust Resistance Gene \textit{Lr21} From the Large and Polyploid Genome of Bread Wheat}

Li Huang\textsuperscript{a}, Steven A. Brooks\textsuperscript{a}, Wanlong Li\textsuperscript{a}, John P. Fellers\textsuperscript{b}, Harold N. Trick\textsuperscript{a}, and Bikram S. Gill\textsuperscript{a}

\textsuperscript{a} Wheat Genetics Resource Center, Department of Plant Pathology, Kansas State University, Manhattan, Kansas 66506-5502
\textsuperscript{b} USDA-ARS, Plant Science and Entomology Unit, Kansas State University, Manhattan, Kansas 66506-5502

Corresponding author: Bikram S. Gill, 4024 Throckmorton, Kansas State University, Manhattan, KS 66506-5502., \textsuperscript{a}bsg@ksu.edu (E-mail)

...In addition, \textbf{FGENESH 1.1} \url{http://www.softberry.com} was used for gene prediction (with monocot genomic DNA parameters).

\section*{Nucleic Acids Research, 2003, Vol. 31, No. 1 229-233}

The TIGR rice genome annotation resource: annotating the rice genome and creating resources for plant biologists

Qiaoping Yuan, Shu Ouyang, Jia Liu, Bernard Suh, Foo Cheung, Razvan Sultana, Dan Lee, John Quackenbush and C. Robin Buell

The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA

*To whom correspondence should be addressed. Tel: +1 301 8383558; Fax: +1 301 8380208; Email: rbuell@tigr.org

Received August 14, 2002; Revised and Accepted October 2, 2002

...The rice sequences were processed with multiple \textit{ab initio} gene finders including \textbf{FGENESH} \url{http://www.softberry.com},...
OsSET1, a novel SET-domain-containing gene from rice

Yun-Kuan Liang, Ying Wang, Yong Zhang, Song-Gang Li, Xiao-Chun Lu, Hong Li, Cheng Zou, Zhi-Hong Xu and Shu-Nong Bai*

PKU-Yale Joint Research Center of Agricultural and Plant Molecular Biology, National Key Laboratory of Protein Engineering and Plant Gene Engineering, College of Life Sciences, Peking University, 5 Yiheyuan Road, Beijing 100871, PR China

* To whom correspondence should be addressed. Fax: +86 10 6275 1526. E-mail: shunongb@pku.edu.cn

... It localizes at chromosome three in rice genome at the contig 1300

http://www.softberry.com/berry.phtml?topic=gfind&prg=FGENESH; GenBank accession number ...

---

**BMC Genomics** 2003, 4:22
Published: 3 June 2003
Received: 31 January 2003
Accepted: 3 June 2003
This article is available from: http://www.biomedcentral.com/1471-2164/4/22

---

**Gene discovery in the hamster: a comparative genomics approach for gene annotation by sequencing of hamster testis cDNAs**

Sreedhar Oduru1, Janee L Campbell1, SriTulasi Karri1, William J Hendry3, Shafiq A Khan1 and Simon C Williams*1,2

Address: 1Department of Cell Biology & Biochemistry, Texas Tech University Health Sciences Center, Lubbock, Texas, USA, 2Southwest Cancer Center at UMC, Lubbock, Texas, USA and 3Department of Biological Sciences, Wichita State University, Wichita, Kansas, USA

Email: Sreedhar Oduru - oduru.sreedhar@ttuhsc.edu; Janee L Campbell - janee.campbell@ttuhsc.edu; SriTulasi Karri - phrsk@ttuhsc.edu; William J Hendry - william.hendry@wichita.edu; Shafiq A Khan - shafiq.khan@ttuhsc.edu; Simon C Williams* - simon.williams@ttuhsc.edu

* Corresponding author

...Two gene prediction programs were used FGENESH http://www.softberry.com/berry.phtml and GENEMARK http://opal.biology.gatech.edu/GeneMark/eukhmm.cgi?org=H.sapiens

---

**TAG Theoretical and Applied Genetics**
DOI: 10.1007/s00122-003-1499-2
Issue: Volume 108, Number 5 Date: March 2004 Pages: 903 - 913

---

**Characterization of soybean genomic features by analysis of its expressed sequence tags**

Ai-Guo Tian1, Jun Wang2, Peng Cui2, Yu-Jun Han2, Hao Xu2, Li-Juan Cong2, Xian-Gang Huang2, Xiao-Ling Wang2, Yong-Zhi Jiao2, Bang-Jun Wang1, Yong-Jun Wang1, Jin-Song Zhang1 and Shou-Yi Chen1
Box 1. Useful human genome annotation and browser URLs

Human genome browsers
- UCSC Human Genome Browser: http://genome.cse.ucsc.edu/cgi-bin/hgGateway/

Ab initio gene prediction programs. *Ab initio* gene predictors rely on the statistical qualities of exons rather than on homologies. Examples of such prediction programs include Genscan [14] (used by Celera, ORNL and Ensembl), fgenesh [15] (used in the Softberry and UCSC browsers as well as Celera’s pipeline) and GrailEXP [16] (ORNL).
For each *Drosophila* GPCR, prediction of gene structure was made in **FGENESH** ([www.softberry.com](http://www.softberry.com); ref. 21) by using about 20 kb of genomic sequence surrounding highly conserved regions, particularly for 5 prime and 3 prime ends of ORFs. Putative *Drosophila* GPCRs in the database were amplified by RT-PCR using primers based on gene predictions in the FGENESH gene finder ([www.softberry.com](http://www.softberry.com); ref. 21).


Isocitrate Lyase Is Essential for Pathogenicity of the Fungus *Leptosphaeria maculans* to Canola (*Brassica napus*)

Alexander Idnurm and Barbara J. Howlett*
School of Botany, The University of Melbourne, Melbourne, Victoria 3010, Australia
Received 17 June 2002/ Accepted 29 July 2002

… The DNA sequence obtained was compared to those in the GenBank database by using **BLAST** (1), and genes were predicted by using **FGENESH** software ([http://www.softberry.com](http://www.softberry.com)) and **GENSCAN** ([www.bionavigator.com](http://www.bionavigator.com)).

Bio-Almanac

GLE ToxExpress, P Offering, CLG Annotated

... sequence data. The genes are identified with the **FGENESH11** gene modeling software exclusively licenced from **Softberry**, Inc. Automatic ...

**Analysis of 106 kb of contiguous DNA sequence from the D genome of wheat reveals high gene density** …

SA Brooks, L Huang, BS Gill, JP Fellers

... trix. In addition, **FGENESH 1.1** ([http://www.softberry.com](http://www.softberry.com)) was used for CDS prediction with monocot genomic DNA parameters. Both ...

**Molecular Genetics and Genomics**
DOI: 10.1007/s00438-002-0706-1
Issue: Volume 267, Number 6 Date: August 2002  Pages: 713 - 720

**Genome sequencing of a 239-kb region of rice chromosome 10L reveals a high frequency of gene duplication and a large chloroplast DNA insertion**

Q. Yuan, J. Hill, J. Hsiao, K. Moffat, S. Ouyang, Z. Cheng, J. Jiang, C. Buell
A1 The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA
A2 Department of Horticulture, University of Wisconsin, Madison, WI 53706, USA

... The sequences were analyzed with several gene prediction programs, including **FGENESH** ([http://www.softberry.com](http://www.softberry.com)), Genemark.hmm (rice matrix; http://opal.biology ...
Different Types and Rates of Genome Evolution Detected by Comparative Sequence Analysis of Orthologous Segments From Four Cereal Genomes

Wusirika Ramakrishna\textsuperscript{a}, Jorge Dubcovsky\textsuperscript{b}, Yong-Jin Park\textsuperscript{1,a}, Carlos Busso\textsuperscript{b}, John Emberton\textsuperscript{a}, Phillip SanMiguel\textsuperscript{c}, and Jeffrey L. Bennetzen\textsuperscript{a}

\textsuperscript{a} Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, \textsuperscript{b} Department of Agronomy and Range Science, University of California, Davis, California 95616 \textsuperscript{c} Purdue University Genomics Core, WSLR, Purdue University, West Lafayette, Indiana 47907

Corresponding author: Jeffrey L. Bennetzen, Hansen Bldg., Purdue University, West Lafayette, IN 47907., \texttt{maize@bilbo.bio.purdue.edu} (E-mail)

\textbf{FGENESH} (http://www.softberry.com/nucleo.html) with the maize training set was used for gene prediction in addition to GENSCAN (http://genes.mit.edu/GENSCAN.html) and GeneMark.hmm (http://genemark.biology.gatech.edu/Gene Mark/).

---

Genomic sequencing reveals gene content, genomic organization, and recombination relationships in barley

Nils Rostoks, Yong-Jin Park, Wusirika Ramakrishna, Jianxin Ma, Arnis Druka, Bryan A. Shiloff, Phillip J. SanMiguel, Zeyu Jiang, Robert Brueggeman, Devinder Sandhu, Kulvinder Gill, Jeffrey L. Bennetzen, Andris Kleinhofs

\textsuperscript{A1} Department of Crop and Soil Sciences, Washington State University, Pullman, WA 99164, USA
\textsuperscript{A2} Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA
\textsuperscript{A3} National Center for Genome Resources, 2935 Rodeo Park Drive East, Santa Fe, NM 87505, USA
\textsuperscript{A4} G302 Agronomy Hall, Iowa State University, Ames, IA 50011-1010, USA
\textsuperscript{A5} Department of Agronomy, University of Nebraska, Lincoln, NE 68583, USA
\textsuperscript{A6} School of Molecular Biosciences and Department of Crop and Soil Sciences, Washington State University, Pullman, WA 99164, USA

... version 1.0 with maize parameters. The \textbf{FGENESH} predictions were run at http://www.softberry.com/. BAC genomic regions were defined ...

---

**Structural organization of the barley D-hordein locus in comparison with its orthologous regions of**

YQ Gu, OD Anderson, CF Londeore, X Kong, RN ...

... et al. 1997) to search for additional genes. In addition, \textbf{FGENESH} (http://www.softberry.com/berry.phtml) and. GENESCAN (http://genes ...
The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases


* Department of Crop and Soil Sciences, Washington State University, Pullman, WA 99164-6420; † Department of Plant Pathology, 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108-6030; and § School of Molecular Biosciences, Washington State University, Pullman, WA 99164-4234

Communicated by Diter von Wettstein, Washington State University, Pullman, WA, May 13, 2002 (received for review March 25, 2002)

.. The gene prediction programs GENSCAN (http://genes.mit.edu/GENSCAN.html) and FGENESH (http://www.softberry.com), as well as NEURAL NETWORK PROMOTER PREDICTION (http://www.fruitfly.org/seq_tools/promoter.html) localized the putative transcription start site of the gene about 400 bp upstream of the translation start site.

---

**Plants Physiol.** 2002 December; 130(4): 1626–1635.


Received July 30, 2002; Accepted October 1, 2002.

**Contiguous Genomic DNA Sequence Comprising the 19-kD Zein Gene Family from Maize**

Rentao Song and Joachim Messing*

Waksman Institute, Rutgers, The State University of New Jersey, 190 Frelinghuysen Road, Piscataway, New Jersey 08854–8020

* Corresponding author; e-mail messing@mbcl.rutgers.edu; fax 732–445–0072.

.. Draft sequences generated from high-throughput DNA sequencing (phase II) were subjected to gene prediction programs with FGENESH (Softberry, Inc., Mount Kisco, NY).

**The Plant Cell**, Vol. 14, 3213-3223, December 2002, Copyright © 2002,

Received July 22, 2002; accepted September 26, 2002.

**Structural Analysis of the Maize Rp1 Complex Reveals Numerous Sites and Unexpected Mechanisms of Local Rearrangement**

Wusirika Ramakrishna, John Emberton, Matthew Ogden, Phillip SanMiguel and Jeffrey L. Bennetzen

* Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907

b Purdue University Genomics Core, Purdue University, West Lafayette, Indiana 47907

1 To whom correspondence should be addressed. E-mail maize@bilbo.bio.purdue.edu; fax 765-496-1496

... FGENESH (http://www.softberry.com/berry.phtml) with the monocot training set was used for gene prediction, in addition to GENSCAN (http://genes.mit.edu/GENSCAN ...

**Comparative Sequence Analysis of the Sorghum *Rph* Region and the Maize *Rp1* Resistance Gene Complex**

Wusirika Ramakrishna, John Emberton, Phillip SanMiguel, Matthew Ogden, Victor Llaca, Joachim Messing, and Jeffrey L. Bennetzen*

Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907 (W.R., J.E., M.O., J.L.B.); Purdue University Genomics Core, Purdue University, West Lafayette, Indiana 47907 (P.S.M.); and Waksman Institute, Rutgers University, Piscataway, New Jersey 08854 (V.L., J.M.)

* Corresponding author; e-mail maize@bilbo.bio.purdue.edu; fax 765–496–1496.

Received September 19, 2002; Accepted October 8, 2002.

…Annotation and sequence analysis were performed as described earlier (Dubcovsky et al., 2001; Song et al., 2001; Ramakrishna et al., 2002a). **FGENESH** ([http://www.softberry.com/berry.phtml](http://www.softberry.com/berry.phtml)) with the monocot training set was used for gene prediction in addition to GENSCAN ([http://genes.mit.edu/GENSCAN.html](http://genes.mit.edu/GENSCAN.html)) and GeneMark.hmm ([http://opal.biology.gatech.edu/GeneMark/eukhmm.cgi](http://opal.biology.gatech.edu/GeneMark/eukhmm.cgi)).

---

Rapid Genome Evolution Revealed by Comparative Sequence Analysis of Orthologous Regions from Four …

YQ Gu, D Coleman-Derr, X Kong, OD Anderson

... **FGENESH** ([http://www.softberry.com/nucleo.html](http://www.softberry.com/nucleo.html)) and GENESCAN ([http://genemark.mit.edu/GENESCAN.htm](http://genemark.mit.edu/GENESCAN.htm)) were used for gene prediction. ...

---

A Genome-Wide Screen Identifies Genes Required for Centromeric Cohesion

JJ Doyle, J Denarie, F Debelle, JC Prome, BB Amor, …

17. HK Choi et al., Genetics, in press. 18. **FGENESH**, see [www.softberry.com/berry.phtml](http://www.softberry.com/berry.phtml).

19. E. Bornberg-Bauer, E. Rivals, M. Vingron, Nucleic Acids Res. ...
OsPPR1, a pentatricopeptide repeat protein of rice is essential for the chloroplast biogenesis

Kodiveri M. Gothandam¹, Eun-Sook Kim¹, Hongjoo Cho¹ and Yong-Yoon Chung¹
(1) School of Life Sciences and Biotechnology, Korea University, Sungbuk-ku, 136-701, Seoul, Anam-Dong, Korea

... nucleotide and amino acid sequences were analyzed by the Basic Local Alignment Search Tool (BLAST) and the Softberry program (http://www.softberry.com/). ...

---

Cellulose Synthase-Like Genes of Rice¹

Samuel P. Hazen, John S. Scott-Craig, and Jonathan D. Walton*
Department of Energy-Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824
* Corresponding author; e-mail walton@msu.edu; fax 517-353-9168.

The Rice Genome Research Program cDNA clones were of high quality; all but one were viable and accurately annotated. The one exception, D22177, was chimeric, containing OsCSLA2 at one end and a predicted DNA-binding protein at the other. For all sequences, the corresponding proteins were deduced using gene prediction software from GeneMark (Atlanta; http://opal.biology.gatech.edu/GeneMark) and Softberry, Inc. (White Plains, NY; http://www.softberry.com), and by manual alignment with the Arabidopsis CsI proteins and with each other.

---

IVET experiments in Pseudomonas fluorescens reveal cryptic promoters at loci associated... Silby et al. Microbiology. 2004; 150: 518-520.

MW Silby, PB Rainey, SB Levy

... Using SoftBerry software (http://www.softberry.com/berry.phtml), -35 and -10 boxes and a transcriptional start site were predicted 84, 60 and 44 bp upstream of ...

---

Two Novel Fungal Virulence Genes Specifically Expressed in Appressoria of the Rice Blast Fungus

Received March 25, 2002; Accepted June 14, 2002.
Approximately 1.2- and 1.4-kb upstream sequences of GAS1 and GAS2 were sequenced and analyzed with several programs, including TRES (www.bioportal.bic.nus.edu.sg/tres), Expasy (www.expasy.org), and SoftBerry (www.softberry.com).


Characterizing CGI-94 (comparative gene identification-94) which is down-regulated in the hippocampus of early stage Alzheimer's disease brain


... Additionally, protein sequence analysis was performed using the following programs at ExPASy, http://www.expasy.ch; softberry, http://www/softberry.com/index ...
... 2003). Finally, there have been several commercial genome management products based on proprietary technology from Softberry, Celera, and Doubletwist.

Generation of T-DNA tagging lines with a bidirectional gene trap vector 2 and the establishment of...

G An

... annotated in the public databases, we undertook. 274. annotation with the **Softberry** program ([http://. 275. www.softberry.com/berry.phtml](http://www.softberry.com/berry.phtml)). Functional clas-. 276 ...

---

0021-9193/01/$04.00+0 DOI: 10.1128/JB.184.1.183-190.2002
Received 25 July 2001/ Accepted 11 October 2001

Regulation of the acuF Gene, Encoding Phosphoenolpyruvate Carboxykinase in the Filamentous Fungus *Aspergillus nidulans*

Michael J. Hynes, Oliver W. Draht, and Meryl A. Davis
Department of Genetics, University of Melbourne, Parkville, Victoria 3010, Australia

The Protein Sequence Analysis program ([http://www.softberry.com/protein.html](http://www.softberry.com/protein.html)) predicted a PEPCK (ATP) signature sequence between amino acids 275 and 290.

---

**Proteins: Structure, Function, and Genetics**
Volume 53, Issue S6, Pages 352 - 368
Supplement: Fifth Meeting on the Critical Assessment of Techniques for Protein Structure Prediction
Published Online: 15 Oct 2003
Received: 4 March 2003; Accepted: 23 June 2003
Digital Object Identifier (DOI) 10.1002/prot.10543

**Comp. Modeling: Assessment**

**Assessment of homology-based predictions in CASP5**

Anna Tramontano1 *, Veronica Morea2
1Department of Biochemical Sciences, “A. Rossi Fanelli,” University of Rome “La Sapienza,” Rome, Italy
2CNR Institute of Molecular Biology and Pathology, University of Rome “La Sapienza,” Rome, Italy
email: Anna Tramontano ([Anna.Tramontano@uniromal.it](mailto:Anna.Tramontano@uniromal.it))
*Correspondence to Anna Tramontano, Department of Biochemical Sciences “A. Rossi Fanelli,” University of Rome “La Sapienza,” P.le Aldo Moro, 5-00185 Rome, Italy

...Page 1. Assessment of Homology-Based Predictions in CASP5. Anna Tramontano 1 * and Veronica Morea 2 1 Department of Biochemical Sciences ...
Spectrum of \( \beta \)-thalassemia mutations and HbF levels in the heterozygous Moroccan population

Wafaa Lemsaddek \(^1\), Isabel Picanço \(^2\), Filomena Seuanes \(^2\), Lahoucine Mahmal \(^3\), Saâd Benchekroun \(^3\), Mohammed Khattab \(^4\), Paulo Nogueira \(^5\), Leonor Osório-Almeida \(^1\)*

\(^1\)Laboratório de Genética Molecular, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal
\(^2\)Laboratório de Hematologia, Instituto Nacional de Saúde Dr. Ricardo Jorge, Lisboa, Portugal
\(^3\)Service Hémato-Oncologie, Hôpital 20 Août CHU Ibn Rochd, Casablanca, Morocco
\(^4\)Service Hémato-Oncologie Pédiatrique, Hôpital d'Enfants CHU Rabat, Morocco
\(^5\)Observatório Nacional de Saúde, Instituto Nacional de Saúde Dr. Ricardo Jorge, Lisboa, Portugal

email: Leonor Osório-Almeida (mlo@fct.unl.pt)

*Correspondence to Leonor Osório-Almeida, Laboratório de Genética Molecular, Secção Autónoma de Biotecnologia, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

... Page 1. Spectrum of Thalassemia Mutations and HbF Levels in the Heterozygous Moroccan Population. Wafaa Lemsaddek, 1 Isabel Picanço ...

---

Pattern Discovery Allowing Wild-Cards, Substitution Matrices, and Multiple Score Functions

Alban Mancheron\(^1\) and Irena Rusu\(^1\)

(1) I.R.I.N., Université de Nantes, 2, Rue de la Houssinière, B.P. 92208, 44322 Nantes Cedex 3, France,

Page 1. Pattern Discovery Allowing Wild-Cards, Substitution Matrices, and Multiple Score Functions. Alban Mancheron and Irena Rusu.

---

Systematic genome-wide approach to positional candidate cloning for identification of novel human disease genes

Internal Medicine Journal

Volume 34 Issue 3 Page 79 - 90 - March 2004


Received 23 October 2003; accepted 15 December 2003
Affiliations: 1Technology and Development team for Mammalian Cellular Dynamics, Bioresource Center, RIKEN Tsukuba Institute, Tsukuba, Ibaraki, 2Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center, RIKEN Yokohama Institute and 3John Curtin School of Medical Research, Australian National University and 4Medical Informatics Centre, University of Canberra, Canberra, Australian Capital Territory, Australia 5Division of Genomic Information Resource Exploration, Science of Biological Supramolecular Systems, Yokohama City University, Graduate School of Integrated Science, Tsurumi-ku, Yokohama, Kanagawa, 6Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, Mishima, Shizuoka, and 7Genome Science Laboratory, RIKEN, Wako, Saitama, Japan,

Correspondence: Yoshihide Hayashizaki, Laboratory for Exploration Research Group, RIKEN Genomic Sciences Center, RIKEN Yokohama Institute, Suehiro-cho 1-7-22, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. Email: rgscerg@gsc.riken.go.jp

The data described in the present paper can be accessed on the web at http://genome.gsc.riken.go.jp/m/pcc/

Page 9-22

Long-term Efficacy of Cholinesterase Inhibitors

Serge Gauthier#
McGill Centre for Studies in Aging, Quebec, Canada
#Correspondence: Dr. Serge Gauthier, FRCPC McGill Centre for Studies in Aging
6825 LaSalle Boulevard, Verdun, Quebec, Canada H4H 1R3, Tel: +1 514 766 2010; Fax: +1 514 888 4050; Email: serge.gauthier@mcgill.ca


Plant Physiology, December 2003, Vol. 133, pp. 2040–2047,

Generation and Analysis of End Sequence Database for T-DNA Tagging Lines in Rice1

Suyoung An2, Sunhee Park2, Dong-Hoon Jeong, Dong-Yeon Lee, Hong-Gyu Kang, Jung-Hwa Yu, Jungh Hur, Sung-Ryul Kim, Young-Hea Kim, Miok Lee, Soo-Jin Han, Soo-Jin Kim, Jungwon Yang, Eunjoo Kim, Soo Jin Wi, Hoo Sun Chung, Jong-Pil Hong, Vitnary Choe, Hakyung Lee, Jung-Hee Choi, Jongmin Nam, Seong-Ryong Kim, Phun-Bum Park, Ky Young Park, Woo Taek Kim, Sunghwa Choe, Chin-Bum Lee, and Gynheung An*
…If a particular sequence had not yet been annotated in the public database, the sequence surrounding the insertion site was annotated using the Softberry program (http://www.softberry.com) and the GeneMark program (http://opal.biology.gatech.edu/GeneMark).

Biology, Yonsei University, Seoul 120–749, Korea (H.S.C., J.-P.H., W.T.K.); Department of Biology, Seoul National University, Seoul 151–747, Korea (V.C., S.C.); Department of Biology, Dong-eui University, Pusan 614–714, Korea. (H.-K.L., J.-H.C., C.-B.L.); and Institute of Molecular Evolutionary Genetics and Department of Biology, Pennsylvania State University, 208 Mueller Laboratory, University Park, Pennsylvania 16802 (J.N.)