

Referências

- Agarwala, R. et al. NCBI's Genome Annotation project – current status. Human Genome Meeting 2001, Genome Informatics. Disponível em: <<http://hgm2001.hgu.mrc.ac.uk/Abstracts/Publish/Workshops/Workshop09/hgm0074.htm>>. Acesso em: 06 maio 2004.
- Altman, R. et al. RiboWeb: An Ontology-Based System for Collaborative Molecular Biology. IEEE Intelligent Systems, v. 14, n.5, p. 68-76, 1999.
- Altschul, S.F. Fundamentals of Database Searching, Trends Guide to Bioinformatics, Elsevier Science, p. 7-9, 1998.
- Altschul, S. F. et al. A basic local alignment search tool, Journal of Molecular Biology, v. 215, p. 403-410, 1990.
- Amzi! inc., Amzi! Prolog. Disponível em: <<http://www.amzi.com/>>. Acesso em: 06 julho 2004.
- APBI - Asia Pacific BioGrid Initiative. Disponível em: <<http://www.apbionet.org/grid/apbiobox/>>. Acesso em 06 maio 2004.
- Attwood, T.K. et al. PRINTS and its automatic supplement, prePRINTS, Nucleic Acids Research, v. 31, n. 1, p. 400-402, 2003.
- Azevedo, V., Pires, P.F., Mattoso, M., Handling Dissimilarities of Autonomous and Equivalent Web Services, Workshop on Web Services, e-Business, and the Semantic Web (WES): Foundations, Models, Architecture, Engineering and Applications, CAiSE'03, The 15th International Conference on Advanced Information Systems Engineering, 2003.
- Bateman, A. et al. The Pfam Protein Families Database, Nucleic Acids Research, v. 30, n.1, p. 276-280, 2002.
- Baker, P.G. et al. An Ontology for Bioinformatics Applications in Bioinformatics, v. 15, n. 6, p. 510-520, 1999.
- Baker, P.G. et al. TAMBIS: Transparent Access to Multiple Bioinformatics Information Sources. An Overview. In Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology (ISMB'98), p. 25-34, Menlow Park, California, 1998. AAAI Press.
- Barton, G. J. Protein Sequence Alignment and Database Scanning, Protein Structure Prediction - a Practical Approach. Editado por M. J. E. Sternberg, IRL Press - Oxford University Press, 1996.
- Benson, D. A. et al. GenBank, Nucleic Acids Research, v. 31, p. 23-27, 2003.
- Bendtsen, J.D., et al. Improved prediction of signal peptides: SignalP 3.0., Journal of Molecular Biology, v. 340, p. 783-795, 2004.
- Boeckmann, B. et al. The Swiss-Prot protein knowledgebase and its supplement TrEMBL in 2003, Nucleic Acids Research, v. 31, p. 365-370, 2003.

Brejová, B. et al. C.Patten. Project Report for CS798g, University of Waterloo, 2000. Disponível em: <<http://citeseer.nj.nec.com/brejova00finding.html>>. Acesso em: 06 maio 2004.

Casanova, M.A., Lemos, M. Optimized Buffer Management for Sequence Comparison in Molecular Biology Databases, Ed. C.J.P.Lucena, Monografia da Ciência da Computação nº 01/01, Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, 2001.

Casey, D.K. Human Genome Program, U.S. Department of Energy, Genomics and Its Impact on Science and Society: A 2003 Primer, November 2003. Disponível em: <http://www.ornl.gov/sci/techresources/Human_Genome/publicat/primer/index.shtml>. Acesso em: 06 maio 2004.

Cannataro, M. et al. Proteus, a Grid based Problem Solving Environment for Bioinformatics: Architecture and Experiments, IEEE Computational Intelligence Bulletin, v.3, n.1, p. 7-18, 2004.

Cavalcanti, M.C. et al. Managing Structural Genomic Workflows using Web Services. Data & knowledge engineering. Elsevier, v.53, n.1, p.45 - 74, 2005. Disponível em <http://authors.elsevier.com/sd/article/S0169023X04001120>. Acesso em: 03 agosto 2004.

CBRG - Computational Biochemistry Research Group, MultAlign: Multiple Sequence Alignment Tools, Disponível em: <<http://cbrg.inf.ethz.ch/Server/MultAlign.html>>. Acesso em: 06 julho 2004.

CGR - Center for Genome Research, Whitehead Institute, Community Annotation Project. Disponível em: <<http://www-genome.wi.mit.edu/annotation/microbes/methanosarcina/sarcinaCAP/>>. Acesso em: 06 maio 2004.

Clamp, M. et al. Ensembl 2002: accommodating comparative genomics, Nucleic Acids Research, v. 31, n.1, p. 38-42, 2003.

Corpet, F. Multiple sequence alignment with hierarchical clustering, Nucleic Acids Research, v. 16, n. 22, p. 10881-10890, 1988.

Crookes, D. Introduction to programming in PROLOG, New York : Prentice Hall, 1988.

DAML Services Coalition - The DARPA Agent Markup Language – OWL-S. Disponível em: <<http://www.daml.org/services/owl-s/>>. Acesso em 10 agosto 2004.

Davidson S.B. et al. K2/Kleisli and GUS: Experiments in Integrated Access to Genomic Data Sources, IBM Systems Journal, v. 4, n.2, p. 512-531, 2001.

Dayhoff, M. O. Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, v. 5, n.3, 1978.

DBM - Departamento de Bioquímica Médica, Universidade Federal do Rio de Janeiro. Disponível em: <<http://www.bioqmed.ufrj.br>>. Acesso em: 06 maio 2004.

DDBJ - DNA Data Bank of Japan - Ssearch Help, Disponível em: <http://helix.genes.nig.ac.jp/homology/ssearch-e_help.html>. Acesso em: 06 agosto 2004.

Delcher, A.L. et al. Improved microbial gene identification with GLIMMER, *Nucleic Acids Research*, v. 27, n. 23, p. 4636-4641, 1999.

Dublin Core Metadata Initiative. Dublin Core Metadata Element Set, Version 1.1: Reference Description. Disponível em: <<http://dublincore.org/documents/dces/>>. Acesso em: 06 maio 2004.

Eddy, S. HMMPFam - Search sequences against an HMM database. Disponível em: <<http://bioweb.pasteur.fr/seqanal/interfaces/hmmpfam.html>>. Acesso em: 06 julho 2004.

Elmasri, R. Navathe, S.B. *Fundamentals of Database Systems*, Addison-Wesley, 2000.

EMBL-EBI, European Bioinformatics Institute, Swiss-Prot and TrEMBL. Disponível em: <<http://us.expasy.org/sprot/>>. Acesso em: 03 agosto 2004.

Ermolaeva, M.D. et al. Prediction of Transcription Terminators in Bacterial Genomes, *Journal of Molecular Biology*, v. 301, p. 27-33, 2000.

EUROGRID Project, BioGrid, Disponível em: <<http://biogrid.icm.edu.pl/>>. Acesso em 06 julho 2004.

Ewing, B. et al. Base-Calling of Automated Sequencer Traces using Phred. I. Accuracy Assessment, *Genome Research*, v. 8, p.175-185, 1998.

Ewing, B., Green, P. Base-Calling of Automated Sequencer Traces using Phred. II. Error Probabilities, *Genome Research*, v. 8, p. 186-194, 1998.

Expasy, Swiss-Prot and TrEMBL. Disponível em: <<http://www.ebi.ac.uk/swissprot/>>. Acesso em: 03 agosto 2004.

Falquet, L. et al. The PROSITE database, its status in 2002, *Nucleic Acids Research*, v. 30, p. 235-238, 2002.

Frishman, D. et al. Functional and structural genomics using PEDANT, *Bioinformatics*, v. 17, p. 44-57, 2001.

Foster, I., Kesselman, C., Tuecke, S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations, *International J. Supercomputer Applications*, v. 15, n. 3, 2001. Disponível em: <<http://www.globus.org/research/papers/anatomy.pdf>>. Acesso em: 06 maio 2004.

Gaasterland, T., Sensen, C.W. MAGPIE: Automated Genome Interpretation, *Trends in Genetics*, v. 12, p. 76-78, 1996.

Garey, M., Johnson, D. *Computers and Intractability - A Guide to the Theory of NP-Completeness*, W.H. Freeman and Co., San Francisco, 1978.

GO - Gene Ontology Consortium. Disponível em: <<http://www.geneontology.org>>. Acesso em: 17 de maio 2004.

Generic Model Organism Project, GBROWSER. Disponível em: <<http://gmod.sourceforge.net/>>. Acesso em: 06 maio 2004.

- Gish, W. Publicação eletrônica [mensagem pessoal]. Mensagens recebidas por <melissa@inf.puc-rio.br> em junho, julho e agosto de 2004.
- Goble, C.A. Supporting Web-based Biology with Ontologies, Proceedings of the Third IEEE ITAB00 Arlington, p.384–390, 2000.
- Goble, C.A. et al. Transparent access to multiple bioinformatics information sources, IBM Systems Journal, v. 4, n.2, p.532-551, 2001.
- Green, P. Documentation for Phrap. Disponível em: <<http://bozeman.mbt.washington.edu/phrap.docs/phrap.html>>. Acesso em: 06 maio 2004.
- Gruber, T. R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. Int., Journal of Human-Computer Studies, v. 43, 1995.
- Guarino, N. Formal Ontology and Information Systems, Formal Ontology in Information Systems, Ed. Amsterdam, Netherlands: IOS Press, 1998.
- Harger, C., et al. The Genome Sequence DataBase (GSDB): improving data quality and data access, Nucleic Acids Research, v. 26, p.21-26, 1998.
- Henikoff, S. et al. Automated construction and graphical presentation of protein blocks from unaligned sequences, Gene-COMBIS, Gene 163, p. 17-26, 1995.
- Henikoff, J.G. et al. Increased coverage of protein families with the blocks database servers, Nucleic Acids Research, v. 28, p. 228-230, 2000.
- Hoersch, S. et al. The GeneQuiz Web server: protein functional analysis through the Web, Trends in Biochemical Sciences, v. 25, p. 33-35, 2000.
- Höhl, M., Kurtz, S., Ohlebusch, E. Efficient Multiple Genome Alignment, Proceedings of the Tenth International Conference on Intelligent Systems for Molecular Biology, Bioinformatics, v. 18, p. 312s-320s, 2002.
- Hoon S, et al. Biopipe: a flexible framework for protocol-based bioinformatics analysis, Genome Research, v. 13, n.8, p.1904-1915, 2003.
- Horowitz, E. et al., Computer Algorithms: C++, Ed. W. H. Freeman Company, 1996.
- Horrocks, I. DAML+OIL: a reason-able Web ontology language, International Conference on Extending Database Technology (EDBT), 2002.
- Huang, X., Madan, A. CAP3: A DNA sequence assembly program, Genome Research, v. 9, p. 868-877, 1999.
- Human Genome Project. Bioinformatics: Human Genome Research in Progress. Disponível em: <http://www.ornl.gov/sci/techresources/Human_Genome/research/informatics.shtml>. Acesso em: 30 junho 2004 (a).
- Human Genome Project. About the Human Genome Project. Disponível em: <http://www.ornl.gov/sci/techresources/Human_Genome/project/about.shtml>. Acesso em: 30 junho 2004 (b).
- IBM Corp., The Era of Grid Computing: A new standard for successful IT strategies. Disponível em: <http://www-1.ibm.com/grid/pdf/it_exec_brief.pdf>. Acesso em: 07 maio 2004.

IGH - Institut de Génétique Humaine, Montpellier France, Bioinformatics Unit, GeneStream. Disponível em: http://xylian.igh.cnrs.fr/getseq/genbank_sequence_finder.html. Acesso em: 06 maio 2004.

Jonassen, I. Efficient Discovery of Conserved Patterns using a Pattern Graph. Technical Report 118, Department of Informatics, University of Bergen, Norway, 1996.

Jones, D.T. Threader - Protein Fold Recognition by Optimal Protein Sequence Threading. Disponível em: <http://www.hgmp.mrc.ac.uk/Registered/Option/threader.html>. Acesso em 06 julho 2004.

Kanehisa, M. Databases of Biological Information, Trends Guide to Bioinformatics, Elsevier Science, p. 24-26, 1998.

Karlin, S., Altschul, S.F., Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes, Proc. Natl. Acad. Sci. USA, vol. 87, p. 2264-2268, 1990.

Karp, P. et al. The EcoCyc Database, Nucleic Acids Research, v. 30, n. 1, p. 56-58, 2002.

Kim, J. Computers are from Mars, Organisms are from Venus, Computer, v. 35, n. 7, p. 25-32, 2002.

Klatte, D.H., Abiview Disponível em: <http://bioinformatics.weizmann.ac.il/software/abiview/abiview.html>. Acesso em: 04 de agosto 2004.

Kulikova, T. et al. The EMBL Nucleotide Sequence Database, Nucleic Acids Research, v. 32, p. D27-D30, 2004.

LBI – Laboratório de Bioinformática, Instituto de Computação – Universidade de Campinas, Cancer Annotation Project. Disponível em: <http://cancer.lbi.ic.unicamp.br/>. Acesso em: 06 maio 2004.

Lee, C., Irizarry, K. The GeneMine System for genome/proteome annotation and collaborative data mining, IBM Systems Journal, v. 40, n.2, p. 592-603, 2001.

Lewis, S.E. et al. Apollo: a sequence annotation editor, Genome Biology, v. 3, n.12, 2002.

Lowe, T.M., Eddy, S.R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence, Nucleic Acids Research, v. 25, 955-964, 1997.

Lemos, M. Gerenciamento de Memória para Comparação de Biossequências, Dissertação de Mestrado. Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, 2000 (a).

Lemos, M., Casanova, M.A. Algoritmos para Análises de Sequências. Ed. C.J.P.Lucena, Monografia da Ciência da Computação nº 05/00, Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, 2000 (b).

Lemos, M., Basílio, A., Casanova, M.A. Um Estudo de Montagem de Fragmentos de Sequências. Ed. C.J.P.Lucena, Monografia da Ciência da Computação nº

05/03, Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, 2003 (a).

Lemos, M., Lifschitz, S. Memory Management for BLAST Processing. 1st International Workshop on Biological Data Management, In conjunction with DEXA 2003, Prague, Czech Republic, p.5-9, 2003 (b).

Lemos, M., Poggi, M., Casanova, M.A. Padrões em Biossequências. Ed. C.J.P.Lucena, Monografia da Ciência da Computação nº 17/03, Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, 2003 (c).

Lemos, M., Seibel, L.F.B., Casanova, M.A. Sistema de Anotações em Biossequências. Ed. C.J.P.Lucena, Monografia da Ciência da Computação nº 04/03, Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, 2003 (d).

Lemos, M., Seibel, L.F.B., Casanova, M.A. BioNotes: A System for Biosequence Annotation. 1st International Workshop on Biological Data Management, In conjunction with DEXA 2003, Prague, Czech Republic, p.16-20, 2003 (e).

Lemos, M., Casanova, M.A., Seibel, L.F.B., Macedo, J.A.F., Miranda, A.B., Ontology-Driven Workflow Management for Biosequence Processing Systems, a ser publicado em Database and Expert Systems Applications (DEXA), 15th International Conference, Zaragoza, Spain, 2004 (a).

Lemos, M., Seibel, L.F.B., Casanova, M.A. Functional Requirements of Biosequence Annotation Systems. Ed. C.J.P.Lucena, Monografia da Ciência da Computação nº 03/04, Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, 2004 (b).

Letovsky, S.I. et al. GDB: the Human Genome Database, Nucleic Acids Research, v. 26, p. 94-99, 1998.

Marchler-Bauer, A. et al. CDD: a curated Entrez database of conserved domain alignments, Nucleic Acids Research, v. 31, p. 383-387, 2003.

Mattson, M. Object-oriented Frameworks: A survey of methodological issues. M.Sc. thesis. University College of Karlshrona, Ronneby, Sweden, 1996.

Medigue, C. et al. Imagene: an integrated computer environment for sequence annotation and analysis, Bioinformatics, v. 15, p. 2-15, 1999.

Meyer, F. GenDB—an open source genome annotation system for prokaryote genomes, Nucleic Acids Research, v. 31, n. 8, p. 2187-2195, 2003.

Modrek, B. Lee, C. Alternative Splicing Annotation Project. Disponível em: <<http://www.bioinformatics.ucla.edu/HASDB/generic.php3>>. Acesso em: 07 maio 2004.

Moller, S. et al. EDITtoTrEMBL: a distributed approach to high-quality automated protein sequence annotation, Bioinformatics, vol. 15, p. 219-227, 1999.

MRC – Medical Resource Council Rosalind Franklin Centre for Genomics Research (RFCGR) - GCG - Sequence analysis package. Disponível em: <<http://www.hgmp.mrc.ac.uk/Registered/Option/gcg.html>>. Acesso em: 07 agosto 2004.

NCBI - National Center of Biotechnology Information, Homepage. Disponível em: <<http://www.ncbi.nlm.nih.gov/>>. Acesso em: 07 maio 2004 (a).

NCBI - National Center of Biotechnology Information, BLAST – XML. Disponível em: <<ftp://ftp.ncbi.nlm.nih.gov/blast/documents/xml/>>. Acesso em: 26 junho 2004 (b).

NCBI - National Center of Biotechnology Information, The BLAST Databases. Disponível em: <<ftp://ftp.ncbi.nih.gov/blast/db/>>. Acesso em: 07 maio 2004 (c).

NCBI - National Center of Biotechnology Information, FASTA Format. Disponível em: <<http://www.ncbi.nlm.nih.gov/BLAST/fasta.shtml>>. Acesso em: 26 junho 2004 (d).

NCBI - National Center of Biotechnology Information, Genbank Flat File Format. Disponível em: <<http://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html>>. Acesso em: 26 junho 2004 (e).

NCBI - National Center for Biotechnology Information. Open Reading Frame Finder. Disponível em: <<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>>. Acesso em: 07 maio 2004 (f).

NCBI - National Center of Biotechnology Information, XML at NCBI. Disponível em: <<http://www.ncbi.nih.gov/IEB/ToolBox/XML/>>. Acesso em: 26 junho 2004 (g).

NCBI - National Center of Biotechnology Information, BLAST. Disponível em: <<http://www.ncbi.nlm.nih.gov/BLAST/>>. Acesso em: 06 julho 2004 (h).

Needleman, S.B., Wunsch, C.D. A general method applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins, *Journal of Molecular Biology*, v.48, p.443-453, 1970.

NIH - National Institutes of Health. Disponível em: <<http://www.nih.gov/>>. Acesso em: 30 junho 2004.

The North Carolina Genomics and Bioinformatics Consortium, The North Carolina BioGrid Project. Disponível em: <<http://www.ncbiogrid.org/>>. Acesso em: 06 julho 2004.

Oracle Corporation. Disponível em: <<http://www.oracle.com/>>. Acesso em: 06 julho 2004.

Özsu, M.T., Valduriez, P., *Principles of Distributed Database Systems*, Second Edition, Prentice Hall, 1999.

Pearson, S. Distributed Annotation System, Disponível em: <<http://www.biodas.org/>>. Acesso em: 07 maio 2004.

Pearson, W. R., Lipman, D. J. Improved Tools for Biological Sequence Comparison, *PNAS*, v. 85, p. 2444- 2448, 1988.

Felsenstein, J. Department of Genetics, University of Washington, PHYLIP: the *PHY*Logeny *I*nference *P*ackage. Disponível em: <<http://evolution.genetics.washington.edu/phylip.html>>. Acesso em: 05 de julho de 2004.

Pop, M., Salzberg, S.L., Shumway, M. Genome Sequence Assembly: Algorithms and Issues, *Computer*, v. 35, n.7, p. 47-54, 2002.

Rational Genomics, Visual Genome. Disponível em: <<http://www.rationalgenomics.com/visualgenome.html>>. Acesso em: 07 maio 2004.

Reese, M.G. et al. Genome Annotation Assesment in Drosophila melanogaster, *Genome Research*, v. 10, n. 4, p. 483-501, 2000.

Rice P., Longden I., Bleasby A., EMBOSS: the European Molecular Biology Open Software Suite. *Trends in Genetics*, v.16, n.6, p. 276-277, 2000.

Rigoutsos, I. Floratos, A. Combinatorial Pattern Discovery in Biological Sequences: The TEIRESIAS Algorithm. *Bioinformatics*, v. 14, n.1, p.55-67, 1998. Errata apareceu em *Bioinformatics*, v.14, n.2, p.229, 1998.

Rutherford, K. et. al. Artemis: sequence visualisation and annotation, *Bioinformatics*, v. 16, n. 10, p. 944-945, 2000.

Sali, A. MODELLER - A Program for Protein Structure Modeling. Disponível em: <<http://salilab.org/modeller/manual/manual.html>>. Acesso em: 06 julho 2004.

The Sanger Institute: Informatics Analysis Software: Alfresco, November 2003. Disponível em: <<http://www.sanger.ac.uk/Software/Alfresco/>>. Acesso em: 07 maio 2004.

Seibel, L. BioAXS: Uma arquitetura de dados e aplicações da Biologia Molecular, Tese de Doutorado, Departamento de Informática, PUC-Rio, 2002.

Seibel, L. F. B., Lifschitz, S. A Genome Database Framework, Database and Expert Systems Applications, 12th International Conference, DEXA 2001, p. 319-329, 2001.

Seibel, L.F.B., Lemos, M., Lifschitz, S. Implementation Issues of Bio-AXS: an Object-oriented Framework for Integrating Biological Data and Applications. Seventh International Database Engineering and Applications Symposium (IDEAS'03), p. 409, 2003.

Servant, F. et al. ProDom: Automated clustering of homologous domains, *Briefings in Bioinformatics*, v. 3 n. 3, p. 246-251, 2002.

Schulze-Kremer, S. Ontologies for Molecular Biology. In *Proceedings of the Third Pacific Symposium on Biocomputing*, p. 693-704, AAAI Press, 1998.

Silberschatz, A., Korth, H.F., Sudarshan, S. *Sistema de Banco de Dados*, Makron Books do Brasil Editora Ltda, 1999.

Smith, T.F., Waterman, M.S. Identification of Common Molecular Subsequences, *Journal of Molecular Biology*, v. 147, p. 195-197, 1981.

Sonnhammer, E. L. L., Durbin, R. A workbench for Large Scale Sequence Homology Analysis, *Comput. Applic. Biosci.*, v. 10, p. 301-307, 1994.

Sousa, M.V., et al. *Gestão da Vida – Genoma e Pós-Genoma*. Bluhm, Brasília, DF: Ed. UnB, 2001.

SRS – Sequence Retrieval System, Expsy Proteomics Server. Disponível em: <<http://us.expsy.org/srs5/>>. Acesso em: 06 julho 2004.

Stanford Medical Informatics at the Stanford University School of Medicine, The Protégé Ontology Editor and Knowledge Acquisition System. Disponível em: <<http://protege.stanford.edu/>>. Acesso em: 06 julho 2004.

Stein, L., Rozen, S., Goodman, N. Managing laboratory workflow with LabBase. In Proceedings of the 1994 Conference on Computers in Medicine (CompMed94). World Scientific Publishing Company, 1995.

Stevens, R. D., Robinson, A.J., Goble, C.A. myGrid: personalised bioinformatics on the information grid, *Bioinformatics*, v. 19, p. 302-304, 2003.

Stevens, R.D., et al. A Classification of Tasks in Bioinformatics. Disponível em: <<http://imgproj.cs.man.ac.uk/tambis/questionnaire/bio-queries.html>>. Acesso em: 04 agosto 2004.

Stoffel, K., Taylor, M., Hendler, J., Efficient Management of Very Large Ontologies, Proc. 14th Nat'l Conf. AI, MIT-AAAI Press, Menlo Park, Calif., 1997.

Sumpter, R., Whitepaper on Data Management, Lawrence Livermore National Laboratory. The IEEE Metadata Workshop, 1994.

Swofford, D. PAUP: Phylogenetic Analysis Using Parsimony. Disponível em: <<http://paup.csit.fsu.edu/>>. Acesso em: 07 maio 2004.

SYBYL® 6.7.1 Tripos Inc., TRIPOS Online, Molecular Modeling and Visualization. Disponível em: <<http://www.tripos.com/sciTech/inSilicoDisc/moleculeModeling/index.html>>. Acesso em: 07 maio 2004.

Technelysium Pty Ltd., Chromas. Disponível em: <<http://www.technelysium.com.au/chromas.html>>. Acesso em: 04 de agosto 2004.

TIGR - The Institute for Genomic Research, RBSFinder. Disponível em: <<http://www.tigr.org/software/>>. Acesso em: 07 maio 2004 (a).

TIGR - The Institute for Genomic Research, Bioinformatics Department, Manatee. Disponível em: <<http://manatee.sourceforge.net/>>. Acesso em: 06 maio 2004 (b).

Thompson, J.D., Higgins, D.G., Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice, *Nucleic Acids Research*, v. 22, p. 4673-4680, 1994.

U.S. Department of Energy. Disponível em: <<http://www.doe.gov/>>. Acesso em: 30 junho 2004.

Weerawarana, S., Curbera, F., Business Process with BPEL4WS: Understanding BPEL4WS, Part 1, 2002. Disponível em: <<http://www-106.ibm.com/developerworks/webservices/library/ws-bpelcoll1/>>. Acesso em: 10 agosto 2004.

Westbrook, J. et al., The Protein Data Bank: unifying the archive, *Nucleic Acids Research*, v. 30, n. 1, p. 245-248, 2002.

WfMC - Workflow Management Coalition (WfMC), The Workflow Reference Model, Document Number TC00-1003. Document Status - Issue 1.1, 1995. Disponível em: <<http://www.wfmc.org/standards/docs/tc003v11.pdf>>. Acesso em: 07 maio 2004.

WfMC - Workflow Management Coalition, Workflow Management Coalition Terminology and Glossar, Technical Report WfMC-TC-1011 3.0. Brussels, 1999.

WfMC - Workflow Management Coalition, Workflow Process Definition Interface -- XML Process Definition Language (XPDL), 2002. Disponível em: <<http://www.wfmc.org/standards/docs.htm>>. Acesso em: 11 agosto 2004.

WfMC - Workflow Management Coalition, The Workflow Handbook 2004, Fischer, L. (ed.). Disponível em <<http://www.wfmc.org/information/handbook04.htm>>. Acesso em 07 maio 2004.

Wroe, C. et al. A suite of DAML+OIL Ontologies to Describe Bioinformatics Web Services and Data, International Journal of Cooperative Information Systems, v. 12, n. 2, p.197-224, 2003.

WU-BLAST - Washington University BLAST Archives – BLAST Manual. Disponível em: <<http://blast.wustl.edu/doc/blast1.pdf>>. Acesso em: 06 julho 2004(a).

WU-BLAST - Washington University BLAST Archives – BLAST Memory Requirements. Disponível em: <<http://blast.wustl.edu/blast/Memory.html>>. Acesso em: 06 julho 2004 (b).

Wu, C. H. et al. The Protein Information Resource: an integrated public resource of functional annotation of proteins, Nucleic Acids Research, v. 30, p. 35-37, 2002.

W3C, Annotated DAML+OIL Ontology Markup, W3C Note, 2001. Disponível em: <<http://www.w3.org/TR/daml+oil-walkthru>>. Acesso em: 17 de maio 2004(a).

W3C, DAML+OIL Reference Description, W3C Note, 2001. Disponível em: <<http://www.w3.org/TR/daml+oil-reference>>. Acesso em: 17 de maio 2004(b).

W3C, OWL Web Ontology Language Overview, 2003. Disponível em: <<http://www.w3.org/TR/owl-features/>>. Acesso em: 17 de maio 2004(a).

W3C, OWL Web Ontology Language Reference, W3C Working Draft, 2003. Disponível em: <<http://www.w3.org/TR/owl-ref/>>. Acesso em: 17 de maio 2004(b).

W3C, Resource Description Framework (RDF). Disponível em: <<http://www.w3.org/RDF>>. Acesso em: 17 de maio 2004 (a).

W3C, XML Schema. Disponível em: <<http://www.w3.org/XML/Schema>>. Acesso em: 07 agosto 2004 (b).

W3C, Extensible Markup Language – XML. Disponível em: <<http://www.w3.org/XML/>>. Acesso em: 07 de agosto 2004 (c).

W3C, XML Pipeline Definition Language, Disponível em: <http://www.w3.org/TR/xml-pipeline/>. Acesso em: 08 agosto 2004 (d).

Wroe, C. et al. A Suite of DAM+OIL Ontologies to Describe Bioinformatics Web services and Data, International Journal of Cooperative Information Systems, v. 12, n. 2, 2003.

Anexo 1 – Definição da Ontologia

Este anexo apresenta duas ontologias de processos de Bioinformática. A primeira, escrita em OWL, modela os processos, contêineres, conexões e projetos como instâncias em OWL, e os programas de análise, tipos de dados de entrada e saída e tipos de recursos como classes em OWL. A segunda, escrita em Amzi-Prolog, modela basicamente uma taxonomia de tipos de programas de análise, dados de entrada e saída e recursos, acrescida de propriedades e relacionamentos entre estes objetos.

Ontologia em OWL

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rss="http://purl.org/rss/1.0/"
  xmlns:jms="http://jena.hpl.hp.com/2003/08/jms#"
  xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns="http://www.puc-rio.br/melissa/Bio#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
```

```
xmlns:vcard="http://www.w3.org/2001/vcard-rdf/3.0#"
xmlns:daml="http://www.daml.org/2001/03/daml+oil#"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xml:base="http://www.puc-rio.br/melissa/Bio">
<owl:Ontology rdf:about="">
<owl:Class rdf:ID="Project">
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    Modela um projeto de Bioinformatica.
</rdfs:comment>
</owl:Class>

<owl:Class rdf:ID="Complete">
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    Modela um projeto para sequenciamento completo do genoma de um organismo.
</rdfs:comment>
<rdfs:subClassOf rdf:resource="#Project"/>
</owl:Class>

<owl:Class rdf:ID="EST">
<rdfs:subClassOf>
    <owl:Class rdf:about="#Project"/>
</rdfs:subClassOf>
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Modela um projeto cujo objetivo e obter somente as sequencias codificantes.

```
</rdfs:comment>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Workflow">
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Modela uma composicao de processos, modelando chamadas a programas de Bioinformatica que analisam biossequencias e que ajudam um pesquisador a interpreta-las.

```
</rdfs:comment>
```

```
</owl:Class>
```

```
<owl:ObjectProperty rdf:ID="Process_Used">
```

```
<rdfs:range rdf:resource="#Process"/>
```

```
<rdfs:domain rdf:resource="#Workflow"/>
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Relaciona um workflow a um processo.

```
</rdfs:comment>
```

```
</owl:ObjectProperty>
```

```
<owl:ObjectProperty rdf:ID="Container_Used">
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Relaciona um workflow a uma contêiner.

```
</rdfs:comment>
```

```
<rdfs:range rdf:resource="#Container"/>
<rdfs:domain rdf:resource="#Workflow"/>
</owl:ObjectProperty>
```

```
<owl:ObjectProperty rdf:ID="Connection_Used">
  <rdfs:range rdf:resource="#Connection"/>
  <rdfs:domain rdf:resource="#Workflow"/>
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    Relaciona um workflow a uma conexao.
  </rdfs:comment>
</owl:ObjectProperty>
```

```
<owl:Class rdf:ID="Process">
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    Modela uma chamada a um programa de Bioinformatica.
  </rdfs:comment>
</owl:Class>
```

```
<owl:Class rdf:ID="Filter">
  <rdfs:subClassOf>
    <owl:Class rdf:about="#Process"/>
  </rdfs:subClassOf>
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Um filtro analisa um conjunto de dados gerado por um processo e extrai partes dele para futuro processamento.

```
</rdfs:comment>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Constructive">
```

```
<rdfs:subClassOf>
```

```
<owl:Class rdf:about="#Process"/>
```

```
</rdfs:subClassOf>
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Um processo construtivo cria novos conjuntos de dados, pertinentes ao domínio de análise de Bioinformática.

```
</rdfs:comment>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Gene_Prediction">
```

```
<rdfs:subClassOf rdf:resource="#Constructive"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Genome_Comparison">
```

```
<rdfs:subClassOf rdf:resource="#Constructive"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Phylogenetic Analysis">
```

```
<rdfs:subClassOf rdf:resource="#Constructive"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Base_Identification">  
  <rdfs:subClassOf rdf:resource="#Constructive"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="Pattern_Discovery">  
  <rdfs:subClassOf rdf:resource="#Constructive"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="Molecular_Prediction">  
  <rdfs:subClassOf rdf:resource="#Constructive"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="Sequence_Assembly">  
  <rdfs:subClassOf rdf:resource="#Constructive"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="Pattern_Recognition">  
  <rdfs:subClassOf rdf:resource="#Constructive"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="Sequence_Alignment">
```



```
<rdfs:subClassOf rdf:resource="#Constructive"/>
</owl:Class>
```

```
<owl:Class rdf:ID="Multiple_Alignment">
  <rdfs:subClassOf rdf:resource="#Sequence_Alignment"/>
</owl:Class>
```

```
<owl:Class rdf:ID="MultiAlign">
  <rdfs:subClassOf rdf:resource="#Multiple_Alignment"/>
</owl:Class>
```

```
<owl:Class rdf:ID="CLUSTAL_W">
  <rdfs:subClassOf rdf:resource="#Multiple_Alignment"/>
</owl:Class>
```

```
<owl:Class rdf:ID="Pairwise_Alignment">
  <rdfs:subClassOf rdf:resource="#Sequence_Alignment"/>
</owl:Class>
```

```
<owl:Class rdf:ID="Global_Alignment">
  <rdfs:subClassOf rdf:resource="#Pairwise_Alignment"/>
</owl:Class>
```

```
<owl:Class rdf:ID="Local_Alignment">  
  <rdfs:subClassOf rdf:resource="#Pairwise_Alignment"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="Smith_Waterman">  
  <rdfs:subClassOf rdf:resource="#Lobal_Alignment"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="ssearch">  
  <rdfs:subClassOf rdf:resource="#Smith_Waterman"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="BLAST">  
  <rdfs:subClassOf rdf:resource="#Lobal_Alignment"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="BLASTN">  
  <rdfs:subClassOf rdf:resource="#BLAST"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="TBLASTN">  
  <rdfs:subClassOf rdf:resource="#BLAST"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="BLASTP">  
  <rdfs:subClassOf rdf:resource="#BLAST"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="BLASTX">  
  <rdfs:subClassOf rdf:resource="#BLAST"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="TBLASTX">  
  <rdfs:subClassOf rdf:resource="#BLAST"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="FAST">  
  <rdfs:subClassOf rdf:resource="#Local_Alignment"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="FASTA">  
  <rdfs:subClassOf rdf:resource="#FAST"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="TFASTA3">  
  <rdfs:subClassOf rdf:resource="#FAST"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="FASTX3">
```

```
<rdfs:subClassOf rdf:resource="#FAST"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="TFASTX3">
```

```
<rdfs:subClassOf rdf:resource="#FAST"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="FASTY3">
```

```
<rdfs:subClassOf rdf:resource="#FAST"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="TFASTY3">
```

```
<rdfs:subClassOf rdf:resource="#FAST"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="External_Control">
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Um processo de controle externo ajuda o pesquisador a gerenciar a execucao do workflow.

```
</rdfs:comment>
```

```
<rdfs:subClassOf rdf:resource="#Process"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Verification_Point">
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Usado sempre em conjunto com um processo P, indica uma condicao, avaliada sobre os contêineres de entrada de P, que deve ser satisfeita para que P seja executado.

```
</rdfs:comment>
```

```
<rdfs:subClassOf rdf:resource="#External_Control"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Stop_Point">
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Indica um ponto em que a execucao do workflow deve parar temporariamente para que o pesquisador analise os resultados intermediarios ja gerados.

```
</rdfs:comment>
```

```
<rdfs:subClassOf rdf:resource="#External_Control"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Exit_Point">
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Indica um ponto em que a execucao do workflow deve parar.

```
</rdfs:comment>
```

```
<rdfs:subClassOf>
```

```
<owl:Class rdf:about="#controle_externo"/>
```

```
</rdfs:subClassOf>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Internal_Control">
```

```
<rdfs:subClassOf rdf:resource="#Process"/>
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Um processo de controle interno e um processo acrescentado automaticamente pelo sistema de gerencia de workflow para que o funcionamento do workflow se torne coerente, viavel ou mais eficiente.

```
</rdfs:comment>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Inspection_Point">
```

```
<rdfs:subClassOf rdf:resource="#Internal_Control"/>
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Um processo de inspecao verifica se os dados de entrada e o resultado da execucao de um processo estao corretos ou nao.

```
</rdfs:comment>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Format_Transformation_Process">
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Um processo de transformacao de formato, como o nome indica, aplica uma transformacao de formato em um conjunto de dados.

```
</rdfs:comment>
```

```
<rdfs:subClassOf rdf:resource="#Internal_Control"/>
```

```
</owl:Class>
```

```
<owl:DatatypeProperty rdf:ID="Popularity">
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

```
    Medida da porcentagem de pesquisadores que conhecem e utilizam os processos de uma classe.
```

```
</rdfs:comment>
```

```
<rdfs:domain rdf:resource="#Process"/>
```

```
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#float"/>
```

```
</owl:DatatypeProperty>
```

```
<owl:DatatypeProperty rdf:ID="Performance">
```

```
<rdfs:domain rdf:resource="#Process"/>
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

```
    Medida da quantidade de recursos computacionais (como tempo de CPU, acesso a disco, etc...) consumidos pelos processos de uma classe.
```

```
</rdfs:comment>
```

```
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#float"/>
```

```
</owl:DatatypeProperty>
```

```
<owl:DatatypeProperty rdf:ID="Fidelity">
```

```
<rdfs:domain rdf:resource="#Process"/>
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

```
    Medida de quao proximo do otimo estao, normalmente, os resultados gerados pelos processos de uma classe.
```

```
</rdfs:comment>
```

```
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#float"/>
</owl:DatatypeProperty>
```

```
<owl:DatatypeProperty rdf:ID="Cost">
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    Medida do custo financeiro para execucao dos processos de uma classe.
  </rdfs:comment>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#float"/>
  <rdfs:domain rdf:resource="#Process"/>
</owl:DatatypeProperty>
```

```
<owl:DatatypeProperty rdf:ID="Default">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#float"/>
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    Medida que indica o quanto os processos de uma classe sao indicados como opcao padrao para o tipo de tarefa a que se propoem.
  </rdfs:comment>
  <rdfs:domain rdf:resource="#Process"/>
</owl:DatatypeProperty>
```

```
<owl:DatatypeProperty rdf:ID="Adequacy ">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#float"/>
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    Medida que indica o quanto os processos de uma classe sao indicados como opcao para um tipo de projeto.
```



```
</rdfs:comment>
```

```
<rdfs:domain rdf:resource="#Process"/>
```

```
</owl:DatatypeProperty>
```

```
<owl:Class rdf:ID="Container">
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Modela uma estruturas de dados responsavel por armazenar e gerenciar um conjunto de dados compartilhado no workflow. Os contêineres que nao sao destino de alguma conexao e que modelam bancos de dados de biosequencias sao tambem chamados de recursos (nao sendo modelados como uma classe separada).

```
</rdfs:comment>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Chromatogram_Set">
```

```
<rdfs:subClassOf rdf:resource="#Container"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Pattern_Set">
```

```
<rdfs:subClassOf rdf:resource="#Container"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Sequence_Set">
```

```
<rdfs:subClassOf rdf:resource="#Container"/>
```

```
</owl:Class>
```

```
<owl:Class rdf:ID="Aminoacid_Sequence_Set">  
  <rdfs:subClassOf rdf:resource="#Sequence_Set"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="Nucleotide_Sequence_Set">  
  <rdfs:subClassOf rdf:resource="#Sequence_Set"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="Contig_Nucleotide_Sequence_Set">  
  <rdfs:subClassOf rdf:resource="#Nucleotide_Sequence_Set"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="Read_Nucleotide_Sequence_Set">  
  <rdfs:subClassOf rdf:resource="#Nucleotide_Sequence_Set"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="Genome_Nucleotide_Sequence_Set">  
  <rdfs:subClassOf rdf:resource="#Nucleotide_Sequence_Set"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="ORF_Nucleotide_Sequence_Set">
```

```
<rdfs:subClassOf rdf:resource="#Nucleotide_Sequence_Set"/>
</owl:Class>

<owl:DatatypeProperty rdf:ID="Container_Type">
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    Propriedadde derivada do tipo da conexao.
  </rdfs:comment>
  <rdfs:domain rdf:resource="#Container"/>
  <rdfs:range>
    <owl:DataRange>
      <owl:oneOf rdf:parseType="Resource">
        <rdf:first rdf:datatype="http://www.w3.org/2001/XMLSchema#string">gradative</rdf:first>
        <rdf:rest rdf:parseType="Resource">
          <rdf:first rdf:datatype="http://www.w3.org/2001/XMLSchema#string">not-gradative</rdf:first>
          <rdf:rest rdf:parseType="Resource">
            <rdf:rest rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#nil"/>
            <rdf:first rdf:datatype="http://www.w3.org/2001/XMLSchema#string">mix</rdf:first>
          </rdf:rest>
        </rdf:rest>
      </owl:oneOf>
    </owl:DataRange>
  </rdfs:range>
</owl:DatatypeProperty>
```

```
<owl:DatatypeProperty rdf:ID="Access_Type">
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    Indica se o contêiner e publico ou privado.
  </rdfs:comment>
  <rdfs:domain rdf:resource="#Container"/>
  <rdfs:range>
    <owl:DataRange>
      <owl:oneOf rdf:parseType="Resource">
        <rdf:first rdf:datatype="http://www.w3.org/2001/XMLSchema#string">public</rdf:first>
        <rdf:rest rdf:parseType="Resource">
          <rdf:first rdf:datatype="http://www.w3.org/2001/XMLSchema#string">private</rdf:first>
          <rdf:rest rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#nil"/>
        </rdf:rest>
      </owl:oneOf>
    </owl:DataRange>
  </rdfs:range>
</owl:DatatypeProperty>
```

```
<owl:DatatypeProperty rdf:ID="Estimated_Max_Size">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#float"/>
  <rdfs:domain rdf:resource="#Container"/>
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Captura, como o nome indica, o tamanhos maximo estimado para o contêiner.

```
</rdfs:comment>
```

```
</owl:DatatypeProperty>
```

```
<owl:DatatypeProperty rdf:ID="Estimated_Min_Size">
```

```
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#float"/>
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Captura, como o nome indica, o tamanhos minimo estimado para o contêiner.

```
</rdfs:comment>
```

```
<rdfs:domain rdf:resource="#Container"/>
```

```
</owl:DatatypeProperty>
```

```
<owl:DatatypeProperty rdf:ID="Format">
```

```
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
```

```
<rdfs:domain rdf:resource="#Container"/>
```

```
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Indica o formato dos dados armazenados no contêiner.

```
</rdfs:comment>
```

```
</owl:DatatypeProperty>
```

```
<owl:DatatypeProperty rdf:ID="Container_Quality">
```

```
<rdfs:range>
```

```
<owl:DataRange>
```

```
<owl:oneOf rdf:parseType="Resource">
  <rdf:rest rdf:parseType="Resource">
    <rdf:first rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Medium</rdf:first>
    <rdf:rest rdf:parseType="Resource">
      <rdf:rest rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#nil"/>
      <rdf:first rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Low</rdf:first>
    </rdf:rest>
  </rdf:rest>
  <rdf:first rdf:datatype="http://www.w3.org/2001/XMLSchema#string">High</rdf:first>
</owl:oneOf>
</owl:DataRange>
</rdfs:range>
<rdfs:domain rdf:resource="#Container"/>
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
  Caracteriza a qualidade dos dados em um contêiner, tipicamente um recurso.
</rdfs:comment>
</owl:DatatypeProperty>

<owl:Class rdf:ID="Connection">
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    Modela uma conexao ligando processo a contêiner (ou vice-versa).
  </rdfs:comment>
</owl:Class>
```

```

<owl:DatatypeProperty rdf:ID="Connection_Type">
  <rdfs:range>
    <owl:DataRange>
      <owl:oneOf rdf:parseType="Resource">
        <rdf:rest rdf:parseType="Resource">
          <rdf:rest rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#nil"/>
          <rdf:first rdf:datatype="http://www.w3.org/2001/XMLSchema#string">not-gradative</rdf:first>
        </rdf:rest>
        <rdf:first rdf:datatype="http://www.w3.org/2001/XMLSchema#string">gradative</rdf:first>
      </owl:oneOf>
    </owl:DataRange>
  </rdfs:range>
  <rdfs:domain rdf:resource="#Connection"/>
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    Indica como o processo escreve os itens de dados no contêiner, ou le itens de dados do contêiner.
  </rdfs:comment>
</owl:DatatypeProperty>

<owl:ObjectProperty rdf:ID="Source">
  <protege:allowedParent rdf:resource="http://www.w3.org/2002/07/owl#Thing"/>
  <protege:allowedParent rdf:resource="#Process"/>
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">

```

Indica um contêiner ou processo que e a origem da conexao.

```
</rdfs:comment>
<rdfs:range rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
<rdfs:domain rdf:resource="#Connection"/>
<protege:allowedParent rdf:resource="#Container"/>
</owl:ObjectProperty>
```

```
<owl:ObjectProperty rdf:ID="Target">
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
```

Indica um contêiner ou processo que e o destino da conexao.

```
</rdfs:comment>
<protege:allowedParent rdf:resource="http://www.w3.org/2002/07/owl#Thing"/>
<rdfs:domain rdf:resource="#Connection"/>
<protege:allowedParent rdf:resource="#Process"/>
<rdfs:range rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
<protege:allowedParent rdf:resource="#Container"/>
</owl:ObjectProperty>
<Pattern_Set rdf:ID="PFam"/>
<Pattern_Set rdf:ID="Blocks"/>
<Pattern_Set rdf:ID="Prosites"/>
<Aminoacid_Sequence_Set rdf:ID="TR-EMBL"/>
<Aminoacid_Sequence_Set rdf:ID="Swiss-Prot"/>
<Aminoacid_Sequence_Set rdf:ID="PIR"/>
```



```
<Aminoacid_Sequence_Set rdf:ID="Genbank-NR"/>  
<Nucleotide_Sequence_Set rdf:ID="EMBL"/>  
<Nucleotide_Sequence_Set rdf:ID="Genbank-NT"/>  
</rdf:RDF>
```

Ontologia em Prolog

```
%% Project
```

```
is_project(est).
```

```
is_project(complete).
```

```
%% Container
```

```
is_container(chromatogram_set).
```

```
is_container(sequence_set).
```

```
is_container(pattern_set).
```

```
is_container(alignment_set).
```

```
is_container(S2) :- isa_container(S1,S2), is_container(S1).
```

```
isa_container(sequence_set, nucleotide_sequence_set).
```

```
isa_container(sequence_set, aminoacid_sequence_set).
```

```
isa_container(nucleotide_sequence_set, nucleotide_sequence_resource).
```

```
isa_container(nucleotide_sequence_resource, genbank_nt).
```

```
isa_container(nucleotide_sequence_resource, embl).
```

```
isa_container(aminoacid_sequence_set, aminoacid_sequence_resource).
```

```
isa_container(aminoacid_sequence_resource, genbank_nr).
```

```
isa_container(aminoacid_sequence_resource, pir).
```

isa_container(aminoacid_sequence_resource, swissprot).
isa_container(aminoacid_sequence_resource, trembl).
isa_container(pattern_set, pattern_resource).
isa_container(pattern_resource, prosite).
isa_container(pattern_resource, pfam).
isa_container(pattern_resource, blocks).
isa_container(nucleotide_sequence_set, read_nucleotide_sequence_set).
isa_container(nucleotide_sequence_set, contig_nucleotide_sequence_set).
isa_container(nucleotide_sequence_set, orf_nucleotide_sequence_set).
isa_container(nucleotide_sequence_set, genome_sequence_set).
isa_container(aminoacid_sequence_set, orf_aminoacid_sequence_set).
isa_container(aminoacid_sequence_set, contig_aminoacid_sequence_set).
isa_container(aminoacid_sequence_set, read_aminoacid_sequence_set).
isa_container(pattern_set, regular_expression_set).
isa_container(alignment_set, nucleotide_alignment_set).
isa_container(alignment_set, aminoacid_alignment_set).
is_container_type(gradative).
is_container_type(not_gradative).
is_container_type(mix).
is_container_accesstype(public).
is_container_accesstype(private).
is_container_dataformat(chromatogram_set_scf_format).
is_container_dataformat(chromatogram_set_abi_format).

```
is_container_dataformat(sequence_set_phd_format).
is_container_dataformat(sequence_set_ace_format).
is_container_dataformat(sequence_set_fasta_format).
is_container_dataformat(nucleotide_sequence_set_fasta_format).
is_container_dataformat(aminoacid_sequence_set_fasta_format).
is_container_dataformat(alignment_set_clustalw_format).
is_container_dataformat(alignment_set_multialign_format).
is_container_dataformat(alignment_set_ssearch_format).
is_container_dataformat(alignment_set_blast_format).
is_container_dataformat(alignment_set_fast_format).
is_container_dataformat(regular_expression_set_fasta_format).
```

```
%% Process
```

```
is_process(P) :- is_internal_control_process(P).
is_process(P) :- is_external_control_process(P).
is_process(P) :- is_filter_process(P).
is_process(P) :- is_internal_constructive_process(P).
```

```
%% Internal Control Process
```

```
is_internal_control_process(inspection).
is_internal_control_process(format_transformation).
is_internal_control_process(S2) :- isa_internal_control_process(S1,S2), is_internal_control_process(S1).
isa_constructive_process(format_transformation, phd2fasta).
```

```
isa_constructive_process(format_transformation, emboss_transeq).
```

```
%% External Control Process
```

```
is_external_control_process(if).
```

```
is_external_control_process(exit).
```

```
is_external_control_process(stop).
```

```
%% Filter Process
```

```
is_filter_process(bat).
```

```
is_filter_process(mspcrunch).
```

```
%% Constructive Process
```

```
is_constructive_process(base_identification).
```

```
is_constructive_process(sequence_assembly).
```

```
is_constructive_process(sequence_alignment).
```

```
is_constructive_process(gene_prediction).
```

```
is_constructive_process(pattern_discovery).
```

```
is_constructive_process(S2) :- isa_constructive_process(S1,S2), is_constructive_process(S1).
```

```
isa_constructive_process(base_identification, phred).
```

```
isa_constructive_process(base_identification, abiview).
```

```
isa_constructive_process(base_identification, chromas).
```

```
isa_constructive_process(sequence_assembly, cap3).
```

```
isa_constructive_process(sequence_assembly, phrap).
```

isa_constructive_process(sequence_assembly, tigr_assembler).
isa_constructive_process(sequence_alignment, multiple_alignment).
isa_constructive_process(sequence_alignment, pairwise_alignment).
isa_constructive_process(multiple_alignment, clustalW).
isa_constructive_process(multiple_alignment, multiAlign).
isa_constructive_process(pairwise_alignment, global_alignment).
isa_constructive_process(pairwise_alignment, local_alignment).
isa_constructive_process(global_alignment, ssearch).
isa_constructive_process(local_alignment, blast).
isa_constructive_process(local_alignment, fast).
isa_constructive_process(blast, blastp).
isa_constructive_process(blast, blastn).
isa_constructive_process(blast, tblastx).
isa_constructive_process(blast, tblastn).
isa_constructive_process(blast, blastx).
isa_constructive_process(fast, fasta3).
isa_constructive_process(fast, fasty3).
isa_constructive_process(fast, fastx3).
isa_constructive_process(fast, tfastx3).
isa_constructive_process(fast, tfasty3).
isa_constructive_process(fast, tfasta3).
isa_constructive_process(gene_prediction, glimmer).
isa_constructive_process(pattern_discovery, teireisias).

%% Constructive Process- Properties

%% Process - Description

is_process_description(base_identification,'Bases identification: writes the base calls of DNA sequences.').

is_process_description(sequence_assembly,'Sequence assembly programs.').

is_process_description(sequence_alignment,'Sequence alignment programs.').

is_process_description(gene_prediction,'Gene prediction programs.').

is_process_description(pattern_discovery,'Pattern discovery programs.').

is_process_description(phred,'Phred reads DNA sequencer trace data, calls bases, assigns quality values to the bases, and writes the base calls and quality values to output files.').

is_process_description(abiview,'Abiview is a free sequence assembly program.').

is_process_description(chromas,'Chromas is a sequence assembly program which requires Windows 95/NT4.0 or higher and is shareware').

is_process_description(cap3,'Cap3 is a DNA sequence assembly program.').

is_process_description(phrap,'Phrap is a program for assembling shotgun DNA sequence data').

is_process_description(tigr_assembler,'The TIGR Assembler is the classic assembly tool developed by TIGR to build a consensus sequence from smaller sequence fragments.').

is_process_description(multiple_alignment,'Alignment is calculated between multiple sequences.').

is_process_description(pairwise_alignment,'Alignment is calculated between two sequences.').

is_process_description(clustalW,'ClustalW is a multiple alignment program.').

is_process_description(multiAlign,'MultiAlign is a multiple alignment program').

is_process_description(global_alignment,'Global alignment is calculated taking into consideration the total length of the two sequences being compared.').

is_process_description(local_alignment,'Local alignment is calculated taking into consideration alignments between substrings of the sequences.').

is_process_description(sssearch,'Search compares a protein or DNA sequence to a sequence database using the Smith-Waterman algorithm.').

is_process_description(blast,'Blast is a sequence comparison program family.').

is_process_description(fast,'Fast is a sequence comparison program family.').

is_process_description(blastp,'Blastp compares an amino acid query sequence against a protein sequence database').

is_process_description(blastn,'Blastn compares a nucleotide query sequence against a nucleotide sequence database').

is_process_description(tblastx,'Tblastn compares a protein query sequence against a nucleotide sequence database dynamically translated in all six reading frames (both strands).').

is_process_description(tblastn,'Tblastx compares the six-frame translations of a nucleotide query sequence against the six-frame translations of a nucleotide sequence database.').

is_process_description(blastx,'Blastx compares the six-frame conceptual translation products of a nucleotide query sequence (both strands) against a protein sequence database').

is_process_description(fasta3,'Fasta3 scans a protein or DNA sequence library for similar sequences').

is_process_description(fasty3,'Fasty3 compares a DNA sequence to a protein sequence database, comparing the translated DNA sequence in forward and reverse frames.').

is_process_description(fastx3,'Fastx3 compares a DNA sequence to a protein sequence database, comparing the translated DNA sequence in forward and reverse frames.').

is_process_description(tfastx3,'Tfastx3 compares a protein sequence to a DNA sequence database, calculating similarities with frameshifts to the forward and reverse orientations.').

is_process_description(tfasty3,'Tfasty3 compares a protein sequence to a DNA sequence database, calculating similarities with frameshifts to the forward and reverse orientations.').

is_process_description(tfasta3,'Tfasta3 compares a protein sequence to a DNA sequence library, translating the DNA sequence library on-the-fly').

is_process_description(phd2fasta,'Phd2fasta is a program to make the fasta format sequence file from the .phd files generated by phred').

%% Process - Parameter

%% Phred - Disable phred base calling and set the current sequence to the ABI base calls that are read from the input file.

%% By default, the current sequence is set to the phred base calls.

is_constructive_process_parameter(phred,nocall,false).

%% Phred - Perform sequence trimming on the current sequence (to permit trimming off low quality segments of reads that are not destined for assembly).

is_constructive_process_parameter(phred,trim,false).

%% Phrap - forcelevel relaxes stringency to varying degree during final contig merge pass. Allowed values are integers from 0 (most stringent) to 10 (least stringent), inclusive.

is_constructive_process_parameter(phrap,forcelevel,0).

%% Phrap - maxgap is the maximum permitted size of an unmatched region in merging contigs, during first (most stringent) merging pass.

is_constructive_process_parameter(phrap,maxgap,30).

%% BLAST - word length.

is_constructive_process_parameter(blastp,w,3).

is_constructive_process_parameter(blastx,w,3).

is_constructive_process_parameter(tblastn,w,3).

is_constructive_process_parameter(tblastx,w,3).

is_constructive_process_parameter(blastn,w,11).

%% BLAST - number of matches one expects to observe by chance alone during the database search.

is_constructive_process_parameter(blastp,e,10).

is_constructive_process_parameter(blastx,e,10).

is_constructive_process_parameter(tblastn,e,10).

is_constructive_process_parameter(tblastx,e,10).

is_constructive_process_parameter(blastn,e,10).

%% BLAST - Using the -dbrecmax option, the record number of the last database sequence to search can be specified.

%% By default the BLAST programs search the entire database.

is_constructive_process_parameter(blastp,dbrecmax,entire_database).

is_constructive_process_parameter(blastx,dbrecmax,entire_database).

is_constructive_process_parameter(tblastn,dbrecmax,entire_database).

is_constructive_process_parameter(tblastx,dbrecmax,entire_database).

is_constructive_process_parameter(blastn,dbrecmax,entire_database).

%% BLAST - Using the -dbrecmin option, the record number of the first database sequence to search can be specified.

%% By default the BLAST programs search the entire database.

is_constructive_process_parameter(blastp,dbrecmin,entire_database).

is_constructive_process_parameter(blastx,dbrecmin,entire_database).

is_constructive_process_parameter(tblastn,dbrecmin,entire_database).

is_constructive_process_parameter(tblastx,dbrecmin,entire_database).

is_constructive_process_parameter(blastn,dbrecmin,entire_database).

%% BLAST - Governing Output

%% Parameter hspmax can be used to limit the number of HSPs reported per database sequence.

%% The default limit is 1000, which is ample leeway for most searches.

%% Notable exceptions are when long query sequences are used (e.g., an entire cosmid) and numerous repetitive or low-complexity (lowentropy) regions exist in the query and database sequences.

is_constructive_process_parameter(blastp,hspmax,1000).

is_constructive_process_parameter(blastx,hspmax,1000).

is_constructive_process_parameter(tblastn,hspmax,1000).

is_constructive_process_parameter(tblastx,hspmax,1000).

is_constructive_process_parameter(blastn,hspmax,1000).

%% BLAST - Governing Output

%% Parameter V is the maximum number of database sequences for which one-line descriptions will be reported.

%% The default value for V is 500.

is_constructive_process_parameter(blastp,b,10).

is_constructive_process_parameter(blastx,b,10).

is_constructive_process_parameter(tblastn,b,10).

is_constructive_process_parameter(tblastx,b,10).

is_constructive_process_parameter(blastn,b,10).

%% BLAST - Governing Output

%% Parameter B regulates the display of the high-scoring segment pairs (alignments). For positive values, B is the maximum number of database sequences for which high-scoring segment pairs will be reported.

%% This may be much smaller than the actual number of high-scoring segment pairs reported, since any given database sequence may yield several HSPs.

%% The default value for B is 250.

is_constructive_process_parameter(blastp,b,10).

is_constructive_process_parameter(blastx,b,10).

is_constructive_process_parameter(tblastn,b,10).

is_constructive_process_parameter(tblastx,b,10).

is_constructive_process_parameter(blastn,b,10).

%% ssearch - Governing Output

%% scores specify how many homologous sequences are reported in list of homology scores.

%% The default value is 100.

is_constructive_process_parameter(ssearch,scores,100).

%% ssearch - Governing Output

%% alignments specify how many alignments with homologous sequences are reported.

%% The default value is 100.

is_constructive_process_parameter(ssearch,alignments,100).

%% Process - Quality

is_constructive_process_quality(phred,cost,10).

is_constructive_process_quality(abiview,cost,0).

is_constructive_process_quality(chromas,cost,0).

is_constructive_process_quality(phred,popularity,10).

is_constructive_process_quality(abiview,popularity,5).
is_constructive_process_quality(chromas,popularity,5).
is_constructive_process_quality(phrap,popularity,10).
is_constructive_process_quality(cap3,popularity,10).
is_constructive_process_quality(tigr_assembler,popularity,5).
is_constructive_process_quality(phrap,adequacy, est, 5).
is_constructive_process_quality(phrap,adequacy, complete,10).
is_constructive_process_quality(cap3,adequacy,est,10).
is_constructive_process_quality(cap3,adequacy,complete,5).
is_constructive_process_quality(tigr_assembler,adequacy,est,10).
is_constructive_process_quality(tigr_assembler,adequacy,complete,5).
is_constructive_process_quality(blast,performance,8).
is_constructive_process_quality(fast,performance,5).
is_constructive_process_quality(blast,fidelity,5).
is_constructive_process_quality(fast,fidelity,8).
is_constructive_process_quality(blast,default,10).
is_constructive_process_quality(fast,default,5).
is_constructive_process_quality(blast,popularity,10).
is_constructive_process_quality(fast,popularity,5).
is_constructive_process_quality(fastx3,performance,10).
is_constructive_process_quality(fasty3,performance,5).
is_constructive_process_quality(fastx3,fidelity,5).
is_constructive_process_quality(fasty3,fidelity,10).

```
is_constructive_process_quality(tfasta3,performance,10).
is_constructive_process_quality(tfastx3,performance,7).
is_constructive_process_quality(tfasty3,performance,5).
is_constructive_process_quality(tfasta3,fidelity,5).
is_constructive_process_quality(tfastx3,fidelity,7).
is_constructive_process_quality(tfasty3,fidelity,10).
is_constructive_process_quality(tfastx3,default,10).
is_constructive_process_quality(tfasty3,default,10).
is_constructive_process_quality(tfasta3,default,5).
is_constructive_process_quality(blastp,performance,10).
is_constructive_process_quality(blastx,performance,8).
is_constructive_process_quality(tblastn,performance,6).
is_constructive_process_quality(tblastx,performance,5).
is_constructive_process_quality(blastp,fidelity,5).
is_constructive_process_quality(blastx,fidelity,6).
is_constructive_process_quality(tblastn,fidelity,8).
is_constructive_process_quality(tblastx,fidelity,10).
is_constructive_process_quality(blastp,default,10).
is_constructive_process_quality(blastx,default,8).
is_constructive_process_quality(tblastn,default,6).
is_constructive_process_quality(tblastx,default,5).
```

```
%% RELATIONSHIPS
```

%% Constructive Process - Filter Programs

is_constructive_filter_process(blastp,mspcrunch).

is_constructive_filter_process(blastp,bat).

%% Format Transformation Programs

is_internal_format_transformation_process(emboss_transeq, nucleotide_sequence_set_fasta_format,aminoacid_sequence_set_fasta_format).

%% Process and its input data, output data, resource

bioprocess_restriction(phred,[chromatogram_set],[read_nucleotide_sequence_set],[[]]).

bioprocess_restriction(abiview,[chromatogram_set],[read_nucleotide_sequence_set],[[]]).

bioprocess_restriction(chromas,[chromatogram_set],[read_nucleotide_sequence_set],[[]]).

bioprocess_restriction(cap3,[read_nucleotide_sequence_set],[contig_nucleotide_sequence_set],[[]]).

bioprocess_restriction(phrap,[read_nucleotide_sequence_set],[contig_nucleotide_sequence_set],[[]]).

bioprocess_restriction(clustalW,[sequence_set],[alignment_set],[[]]).

bioprocess_restriction(multiAlign,[sequence_set],[alignment_set],[[]]).

bioprocess_restriction(ssearch,[aminoacid_sequence_set],[aminoacid_alignment_set],[aminoacid_sequence_resource]).

bioprocess_restriction(ssearch,[nucleotide_sequence_set],[nucleotide_alignment_set],[nucleotide_sequence_resource]).

bioprocess_restriction(blastp,[aminoacid_sequence_set],[aminoacid_alignment_set],[aminoacid_sequence_resource]).

bioprocess_restriction(blastn,[nucleotide_sequence_set],[nucleotide_alignment_set],[nucleotide_sequence_resource]).

bioprocess_restriction(tblastx,[nucleotide_sequence_set],[aminoacid_alignment_set],[nucleotide_sequence_resource]).

bioprocess_restriction(blastx,[nucleotide_sequence_set],[aminoacid_alignment_set],[aminoacid_sequence_resource]).

bioprocess_restriction(tblastn,[aminoacid_sequence_set],[aminoacid_alignment_set],[nucleotide_sequence_resource]).

```

bioprocess_restriction(fasta3,[aminoacid_sequence_set],[aminoacid_alignment_set],[aminoacid_sequence_resource]).
bioprocess_restriction(fasta3,[nucleotide_sequence_set],[nucleotide_alignment_set],[nucleotide_sequence_resource]).
bioprocess_restriction(fasty3,[nucleotide_sequence_set],[aminoacid_alignment_set],[aminoacid_sequence_resource]).
bioprocess_restriction(fastx3,[nucleotide_sequence_set],[aminoacid_alignment_set],[aminoacid_sequence_resource]).
bioprocess_restriction(tfastx3,[aminoacid_sequence_set],[aminoacid_alignment_set],[nucleotide_sequence_resource]).
bioprocess_restriction(tfasty3,[aminoacid_sequence_set],[aminoacid_alignment_set],[nucleotide_sequence_resource]).
bioprocess_restriction(tfasta3,[aminoacid_sequence_set],[aminoacid_alignment_set],[nucleotide_sequence_resource]).
bioprocess_restriction(glimmer,[genome_sequence_set],[orf_nucleotide_sequence_set],[[]]).
bioprocess_restriction(teireisias,[sequence_set],[regular_expression_set],[[]]).

```

```

bioprocess_restriction_input(P,R):- bioprocess_restriction(P,Le,Lo,Lr), in(R,Le).

```

```

in(R,[R|T]).                % R is first element of the list

```

```

in(R,[F|T]) :- in(R,T).     % R is the rest of the list

```

```

bioprocess_restriction_output(P,R):- bioprocess_restriction(P,Le,Lo,Lr), out(R,Lo).

```

```

out(R,[R|T]).

```

```

out(R,[F|T]) :- out(R,T).

```

```

bioprocess_restriction_resource(P,R):- bioprocess_restriction(P,Le,Lo,Lr), res(R,Lr).

```

```

res(R,[R|T]).

```

```

res(R,[F|T]) :- res(R,T).

```

```

%% Process and its input data, output data and resource format data

```


bioprocess_restriction_internal_format(phred,[chromatogram_set_scf_format],[sequence_set_phd_format],[]).
bioprocess_restriction_internal_format(phred,[chromatogram_set_abi_format],[sequence_set_phd_format],[]).
bioprocess_restriction_internal_format(phred,[chromatogram_set_scf_format],[sequence_set_fasta_format],[]).
bioprocess_restriction_internal_format(phred,[chromatogram_set_abi_format],[sequence_set_fasta_format],[]).
bioprocess_restriction_internal_format(abiview,[chromatogram_set_abi_format],[sequence_set_fasta_format],[]).
bioprocess_restriction_internal_format(chromas,[chromatogram_set_scf_format],[sequence_set_fasta_format],[]).
bioprocess_restriction_internal_format(chromas,[chromatogram_set_abi_format],[sequence_set_fasta_format],[]).
bioprocess_restriction_internal_format(cap3,[sequence_set_fasta_format],[sequence_set_fasta_format],[]).
bioprocess_restriction_internal_format(cap3,[sequence_set_fasta_format],[sequence_set_ace_format],[]).
bioprocess_restriction_internal_format(phrap,[sequence_set_fasta_format],[sequence_set_fasta_format],[]).
bioprocess_restriction_internal_format(phrap,[sequence_set_fasta_format],[sequence_set_ace_format],[]).
bioprocess_restriction_internal_format(clustalW,[sequence_set_fasta_format],[alignment_set_clustalw_format],[]).
bioprocess_restriction_internal_format(multiAlign,[sequence_set_fasta_format],[alignment_set_multialign_format],[]).
bioprocess_restriction_internal_format(ssearch,[sequence_set_fasta_format],[alignment_set_ssearch_format],[aminoacid_sequence_set_fasta_format]).
bioprocess_restriction_internal_format(blastp,[aminoacid_sequence_set_fasta_format],[alignment_set_blast_format],[aminoacid_sequence_set_fasta_format]).
bioprocess_restriction_internal_format(blastn,[nucleotide_sequence_set_fasta_format],[alignment_set_blast_format],[nucleotide_sequence_set_fasta_format]).
bioprocess_restriction_internal_format(tblastx,[nucleotide_sequence_set_fasta_format],[alignment_set_blast_format],[nucleotide_sequence_set_fasta_format]).
bioprocess_restriction_internal_format(blastx,[nucleotide_sequence_set_fasta_format],[alignment_set_blast_format],[aminoacid_sequence_set_fasta_format]).
bioprocess_restriction_internal_format(tblastn,[aminoacid_sequence_set_fasta_format],[alignment_set_blast_format],[aminoacid_sequence_set_fasta_format]).
bioprocess_restriction_internal_format(fasta3,[aminoacid_sequence_set_fasta_format],[alignment_set_fast_format],[aminoacid_sequence_set_fasta_format]).
bioprocess_restriction_internal_format(fasta3,[nucleotide_sequence_set_fasta_format],[alignment_set_fast_format],[aminoacid_sequence_set_fasta_format]).
bioprocess_restriction_internal_format(fasty3,[nucleotide_sequence_set_fasta_format],[alignment_set_fast_format],[aminoacid_sequence_set_fasta_format]).
bioprocess_restriction_internal_format(fastx3,[nucleotide_sequence_set_fasta_format],[alignment_set_fast_format],[aminoacid_sequence_set_fasta_format]).

```

bioprocess_restriction_internal_format(tfastx3,[aminoacid_sequence_set_fasta_format],[alignment_set_fast_format],[nucleotide_sequence_set_fasta_format]).
bioprocess_restriction_internal_format(tfasty3,[aminoacid_sequence_set_fasta_format],[alignment_set_fast_format],[nucleotide_sequence_set_fasta_format]).
bioprocess_restriction_internal_format(tfasta3,[aminoacid_sequence_set_fasta_format],[alignment_set_fast_format],[nucleotide_sequence_set_fasta_format]).
bioprocess_restriction_internal_format(glimmer,[sequence_set_fasta_format],[nucleotide_sequence_set_fasta_format],[ ]).
bioprocess_restriction_internal_format(teireisias,[sequence_set_fasta_format],[regular_expression_set_fasta_format],[ ]).
bioprocess_restriction_input_internal_format(P,R):- bioprocess_restriction_internal_format(P,Le,Lo,Lr), inIF(R,Le).
inIF(R,[R|T]).
inIF(R,[F|T]) :- inIF(R,T).

```

```

bioprocess_restriction_output_internal_format(P,R):- bioprocess_restriction_internal_format(P,Le,Lo,Lr), outIF(R,Lo).
outIF(R,[R|T]).
outIF(R,[F|T]) :- outIF(R,T).

```

```

bioprocess_restriction_resource_internal_format(P,R):- bioprocess_restriction_internal_format(P,Le,Lo,Lr), resIF(R,Lr).
resIF(R,[R|T]).
resIF(R,[F|T]) :- resIF(R,T).

```

%% Process and Containers - Connection Type

```

is_connection_gradative(phred,[yes],[yes],[ ]).
is_connection_gradative(phred,[yes],[yes],[ ]).
is_connection_gradative(abiview,[yes],[yes],[ ]).
is_connection_gradative(chromas,[yes],[yes],[ ]).

```

is_connection_gradative(cap3,[no],[yes],[]).
is_connection_gradative(phrap,[no],[yes],[]).
is_connection_gradative(clustalW,[no],[no],[]).
is_connection_gradative(multiAlign,[no],[no],[]).
is_connection_gradative(ssearch,[yes],[yes],[no]).
is_connection_gradative(blastp,[yes],[yes],[no]).
is_connection_gradative(blastn,[yes],[yes],[no]).
is_connection_gradative(tblastx,[yes],[yes],[no]).
is_connection_gradative(blastx,[yes],[yes],[no]).
is_connection_gradative(tblastn,[yes],[yes],[no]).
is_connection_gradative(fasta3,[yes],[yes],[no]).
is_connection_gradative(fasty3,[yes],[yes],[no]).
is_connection_gradative(fastx3,[yes],[yes],[no]).
is_connection_gradative(tfastx3,[yes],[yes],[no]).
is_connection_gradative(tfasty3,[yes],[yes],[no]).
is_connection_gradative(tfasta3,[yes],[yes],[no]).
is_connection_gradative(glimmer,[no],[yes],[]).
is_connection_gradative(teireisias,[yes],[yes],[]).

bioprocess_restriction_input_gradative(P,R):- is_connection_gradative(P,Le,Lo,Lr), inSW(R,Le).
inSW(R,[R|T]).
inSW(R,[F|T]) :- inSW(R,T).

```
bioprocess_restriction_output_gradative(P,R):- is_connection_gradative(P,Le,Lo,Lr), outSW(R,Lo).  
outSW(R,[R|T]).  
outSW(R,[F|T]) :- outSW(R,T).
```

```
bioprocess_restriction_resource_gradative(P,R):- is_connection_gradative(P,Le,Lo,Lr), resSW(R,Lr).  
resSW(R,[R|T]).  
resSW(R,[F|T]) :- resSW(R,T).
```