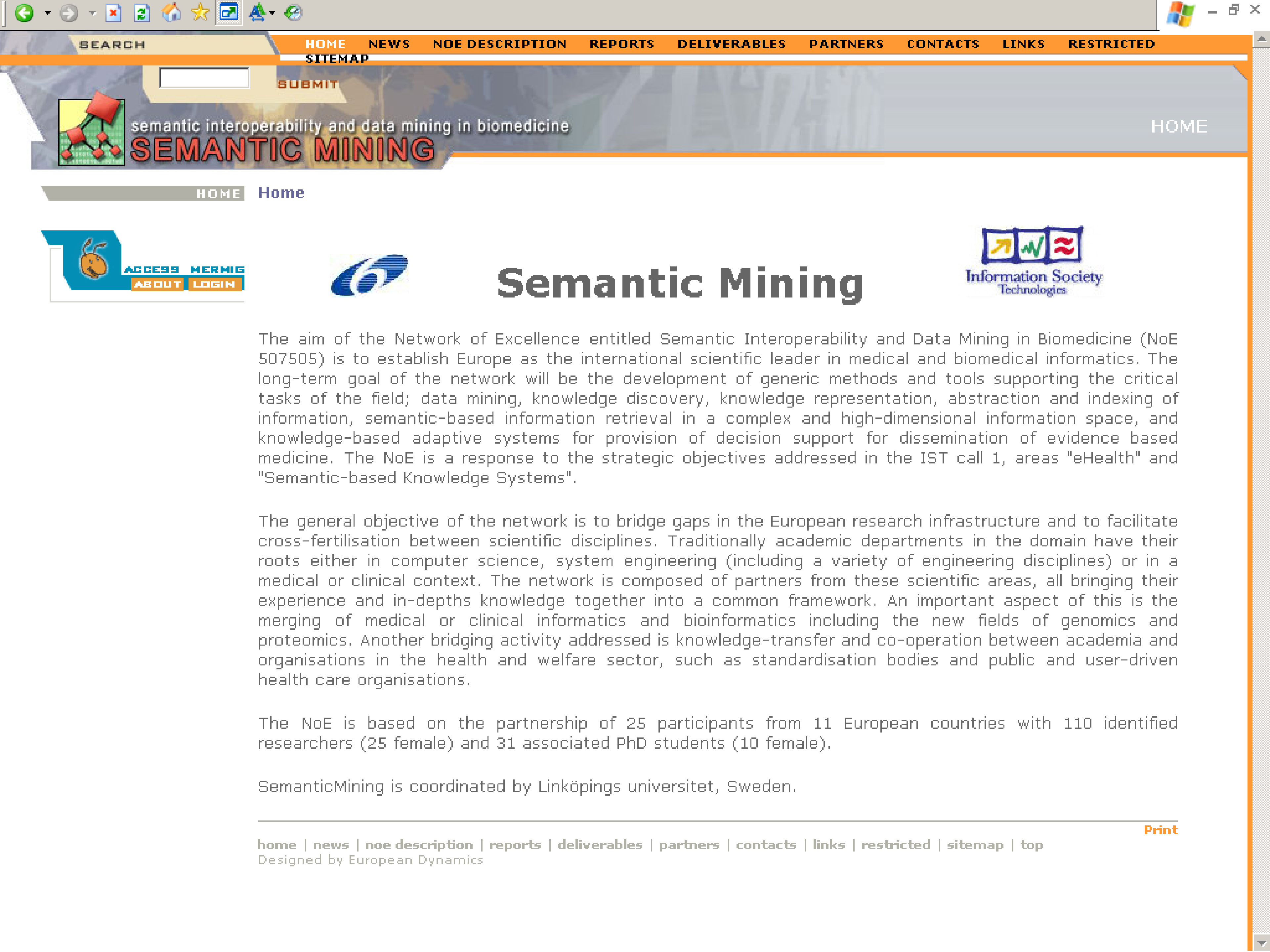


Home

Semantic Mining

The aim of the Network of Excellence entitled Semantic Interoperability and Data Mining in Biomedicine (NoE 507505) is to establish Europe as the international scientific leader in medical and biomedical informatics. The long-term goal of the network will be the development of generic methods and tools supporting the critical tasks of the field; data mining, knowledge discovery, knowledge representation, abstraction and indexing of information, semantic-based information retrieval in a complex and high-dimensional information space, and knowledge-based adaptive systems for provision of decision support for dissemination of evidence based medicine. The NoE is a response to the strategic objectives addressed in the IST call 1, areas "eHealth" and "Semantic-based Knowledge Systems".

The general objective of the network is to bridge gaps in the European research infrastructure and to facilitate cross-fertilisation between scientific disciplines. Traditionally academic departments in the domain have their roots either in computer science, system engineering (including a variety of engineering disciplines) or in a medical or clinical context. The network is composed of partners from these scientific areas, all bringing their experience and in-depths knowledge together into a common framework. An important aspect of this is the merging of medical or clinical informatics and bioinformatics including the new fields of genomics and proteomics. Another bridging activity addressed is knowledge-transfer and co-operation between academia and organisations in the health and welfare sector, such as standardisation bodies and public and user-driven health care organisations.

The NoE is based on the partnership of 25 participants from 11 European countries with 110 identified researchers (25 female) and 31 associated PhD students (10 female).

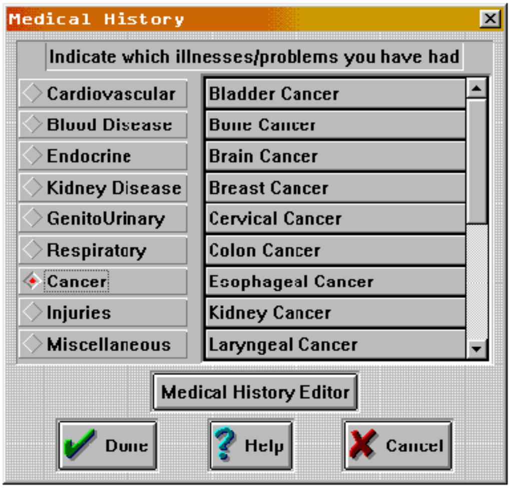
SemanticMining is coordinated by Linköpings universitet, Sweden.

Designed by European Dynamics

© Wegener Falkplan BV for Trotty

Application Areas... ...Research Areas

Health Statistics



Medical History

Indicate which illnesses/problems you have had

○ Cardiovascular	Bladder Cancer
○ Blood Disease	Bone Cancer
○ Endocrine	Brain Cancer
○ Kidney Disease	Breast Cancer
○ GenitoUrinary	Cervical Cancer
○ Respiratory	Colon Cancer
● Cancer	Esophageal Cancer
○ Injuries	Kidney Cancer
○ Miscellaneous	Laryngeal Cancer

Health Care

Bioinformatics

- Knowledge engineering
    - Ontology engineering
    - Coding, indexing and information retrieval
    - Data mining, knowledge extraction and representation
    - Natural Language Processing
    - The Semantic Web
- …to support application areas
    - Information and decision support
    - Infrastructure for health care information systems

Integration

- … to bridge gaps in the European research infrastructure and to facilitate cross-fertilisation between disciplines …
  - Computer science (engineers, logicians, linguists) [6 partners]
  - Bioinformatics and medical informatics [11 partners]
  - Health care organisations, standardisation bodies [6 partners]
  - Philosophy [2 partners]
  - SMEs [2 partners]

Partners

- Biomedical Engineering, Medical Informatics, Linköping University, Sweden
- Computer Science, Linköping University, Sweden
- Committee Nomenclature, Properties and Units in Lab Medicine, Linköping University, Sweden
- Karolinska Institutet, Stockholm, Sweden
- Sahlgrenska University Hospital,Göteborg, Sweden
- Dept of Swedish, Göteborg University, Sweden
- Dept of Medical Informatics, Universitätsklinikum Freiburg, Germany
- Jena University Language and Information Engineering (JULIE),Friedrich-Schiller-Universität, Jena, Germany"
- IFOMIS, Saarland, Germany
- Institute of Informatics and Applied Mathematics,Christian-Albrechts-University of Kiel, Germany
- Division of Medical Informatics, Geneve University Hospital, Switzerland
- Dept of Computer Science, University of Manchester, UK
- Centre for Health Informatics and Multiprofessional Education, University College London, UK
- The Information Technology Research Institute, University of Brighton, UK
- Public Health and Medical Informatics Laboratory, Broussais University Hospital,Paris, France
- Institute of Cognitive Science, Laboratory for Applied Ontology , Italy
- European Bioinformatics Institute, UK
- National Institute for Strategic Health Research, Budapest, Hungary
- WHO Collaborating Centre for Classification of Diseases in the Nordic countries, Uppsala University, Sweden
- The National Board of Health and Welfare, Sweden
- National Research and Development Centre for Welfare and Health, Finland
- KITH AS, Norway
- National Board of Health, Denmark
- Merrall-Ross International Ltd, UK
- European Dynamics S.A., Greece

Semantic Mining Board

- Hans Åhlfeldt, coordinator, Linköping University, Sweden
- Gunnar Klein, Karolinska Institutet, Sweden
- Jeremy Rogers, University of Manchester, UK
- Patrick Ruch, University Hospital Geneva, Switzerland
- Stefan Schulz, University Hospital, Freiburg Germany
- Arne Kverneland, National Board of Health. Denmark

Scientific Advisory Committee

- Alan Rector, Manchester, UK
- Robert Baud, Geneva, Switzerland
- Cornelius Rosse, Seattle, USA
- Chris Chute, Rochester, USA
- Anita Burgun, Rennes, France
- Jean-Marie Rodrigues, Saint Etienne, France

Research Areas

- Principles in ontology engineering
  - examples: FMA, GO, SNOMED CT
- Evaluation of SNOMED CT
  - strategies and experiences from evaluation and translation
- Concept systems in laboratory medicine
  - communication between bioinformatics, laboratory medicine and the EHR
- Multi-lingual medical dictionaries
  - English, German, French, Portuguese, Spanish, Swedish …
- Data/text mining in bioinformatics
  - NLP, IR applied in biomedicine (at EBI)
- The semantic-based electronic health record
  - contribution to standards, information models and concept systems
- What can ontologies do for health statistics?
  - information quality versus aggregation level
  - use of SNOMED CT as aggregation system

Ontology Engineering: Objectives, Activities

- Share understanding across 3 communities
  - Philosophy, Logicians, Engineers
- Coordinate future research efforts
- Coordinate input to standardisation activities
  - ISO, CEN, IEEE and HL7.
- Argue case for ontology-based biomedical vocabularies and coding systems
- Develop migration pathways
- Contribute to a consensus on a biomedical "upper ontology".
- Contribute to the convergence of biomedical ontologies
- Saarbrucken workshop: SNOMED CT

2006

2005

2004

WP20: Multilingual Lexicon

Three lines of work:

- MorphoSaurus subword lexicon: Links minimal, semantically atomic lexical units in 6 languages (approx. 80,000 entries, 27,000 equivalence classes). Purpose: Cross-language text retrieval, semantic interface between medical dictionaries
- Semi automated lexical acquisition: generating Spanish subwords out of Portuguese subwords, and Swedish out of German and English ones.
- Common Lexicon Interchange Format

  Based on the (EU-funded) MULTEXT morpho-syntactic description. Facilitates the re-use of lexical resources

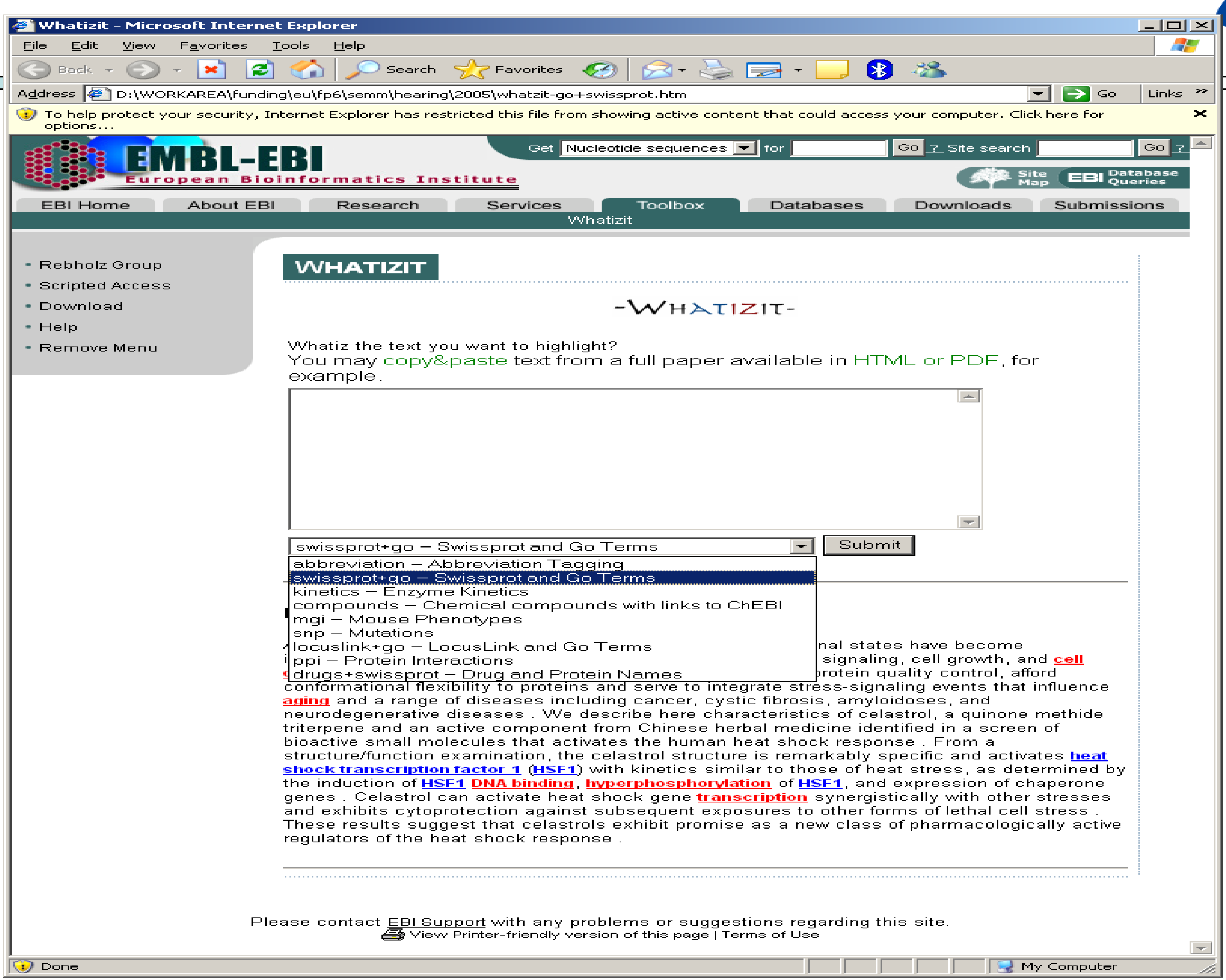
Health Statistics WP23

- ~8 participants
  - (Finland, Hungary, Sweden, Denmark)
- Documenting problems with European Health Statistics
  - Kick-off July
    - Hungary
  - 2 Workshops October
    - Sweden
    - Iceland
- Ontologies for health indicators
- Reliability of health indicators

WP24: Information Retrieval and Data Mining

- Semantic Interoperability
  - Normalized vocabulary (Gene Ontology, MeSH…)
  - Online integration tool: http://www.ebi.ac.uk/Rebholz-srv/whatizit/form.jsp
- Information Retrieval and Extraction
  - Gene and Proteins, Drugs…
  - Protein Functions: apoptosis-induction…
  - Cellular Components: membrane, mitochondria..
  - Biological Processes: digestion, reproduction…
- Knowledge coupling
  - Uni-Prot (EU), MGI, LocusLink (US)
    → via Sequence Retrieval System
    → Need new Tools for Images and Full-text articles !

Entity Types



Address: D:\WORKAREA\funding\eu\fp6\semm\hearing\2005\whatzit-go+swissprot.htm

To help protect your security, Internet Explorer has restricted this file from showing active content that could access your computer. Click here for options...

Whatiz the text you want to highlight?
You may copy&paste text from a full paper available in HTML or PDF, for example.

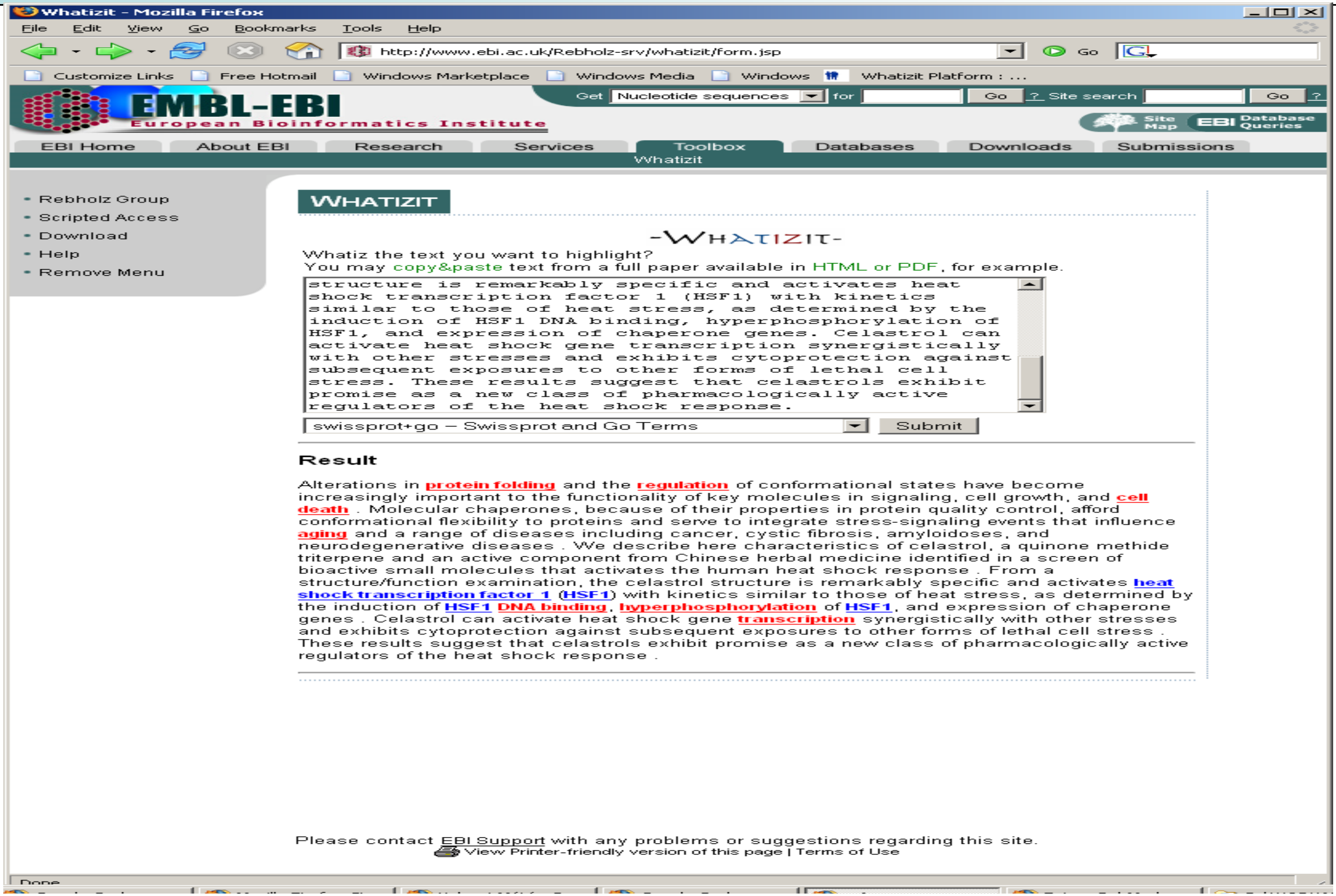
swissprot+go – Swissprot and Go Terms

- abbreviation – Abbreviation Tagging
- swissprot+go – Swissprot and Go Terms
- kinetics – Enzyme Kinetics
- compounds – Chemical compounds with links to ChEBI
- mgi – Mouse Phenotypes
- snp – Mutations
- locuslink+go – LocusLink and Go Terms
- ppi – Protein Interactions
- drugs+swissprot – Drug and Protein Names

nal states have become
signaling, cell growth, and **cell**
rotein quality control, afford
conformational flexibility to proteins and serve to integrate stress-signaling events that influence **aging** and a range of diseases including cancer, cystic fibrosis, amyloidoses, and neurodegenerative diseases . We describe here characteristics of celastrol, a quinone methide triterpene and an active component from Chinese herbal medicine identified in a screen of bioactive small molecules that activates the human heat shock response . From a structure/function examination, the celastrol structure is remarkably specific and activates **heat shock transcription factor 1** (**HSF1**) with kinetics similar to those of heat stress, as determined by the induction of **HSF1** **DNA binding**, **hyperphosphorylation** of **HSF1**, and expression of chaperone genes . Celastrol can activate heat shock gene **transcription** synergistically with other stresses and exhibits cytoprotection against subsequent exposures to other forms of lethal cell stress . These results suggest that celastrols exhibit promise as a new class of pharmacologically active regulators of the heat shock response .

Please contact EBI Support with any problems or suggestions regarding this site.
View Printer-friendly version of this page | Terms of Use

Whatizit !



Whatiz the text you want to highlight?
You may copy&paste text from a full paper available in HTML or PDF, for example.

```
structure is remarkably specific and activates heat
shock transcription factor 1 (HSF1) with kinetics
similar to those of heat stress, as determined by the
induction of HSF1 DNA binding, hyperphosphorylation of
HSF1, and expression of chaperone genes. Celastrol can
activate heat shock gene transcription synergistically
with other stresses and exhibits cytoprotection against
subsequent exposures to other forms of lethal cell
stress. These results suggest that celastrols exhibit
promise as a new class of pharmacologically active
regulators of the heat shock response.
```

swissprot+go – Swissprot and Go Terms

Result

Alterations in **protein folding** and the **regulation** of conformational states have become increasingly important to the functionality of key molecules in signaling, cell growth, and **cell death** . Molecular chaperones, because of their properties in protein quality control, afford conformational flexibility to proteins and serve to integrate stress-signaling events that influence **aging** and a range of diseases including cancer, cystic fibrosis, amyloidoses, and neurodegenerative diseases . We describe here characteristics of celastrol, a quinone methide triterpene and an active component from Chinese herbal medicine identified in a screen of bioactive small molecules that activates the human heat shock response . From a structure/function examination, the celastrol structure is remarkably specific and activates **heat shock transcription factor 1** (**HSF1**) with kinetics similar to those of heat stress, as determined by the induction of **HSF1** **DNA binding**, **hyperphosphorylation** of **HSF1**, and expression of chaperone genes . Celastrol can activate heat shock gene **transcription** synergistically with other stresses and exhibits cytoprotection against subsequent exposures to other forms of lethal cell stress . These results suggest that celastrols exhibit promise as a new class of pharmacologically active regulators of the heat shock response .

Please contact EBI Support with any problems or suggestions regarding this site.
View Printer-friendly version of this page | Terms of Use

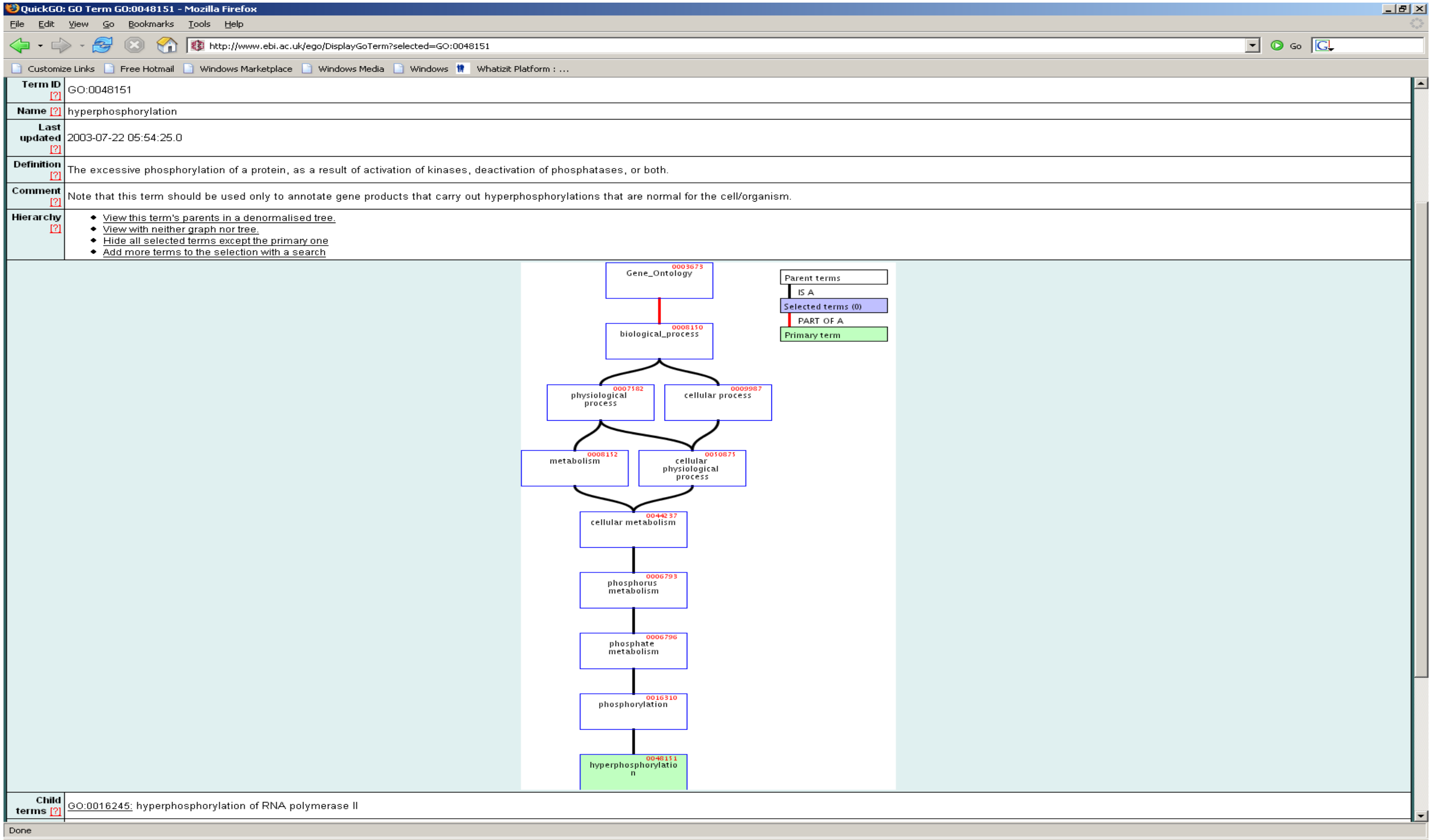
Biomedical Text (MEDLINE Abstract)

Alterations in protein folding and the regulation of conformational states have become increasingly important to the functionality of key molecules in signaling, cell growth, and cell death. Molecular chaperones, because of their properties in protein quality control, afford conformational flexibility to proteins and serve to integrate stress-signaling events that influence aging and a range of diseases including cancer, cystic fibrosis, amyloidoses, and neurodegenerative diseases. We describe here characteristics of celastrol, a quinone methide triterpene and an active component from Chinese herbal medicine identified in a screen of bioactive small molecules that activates the human heat shock response. From a structure/function examination, the celastrol structure is remarkably specific and activates heat shock transcription factor 1 (HSF1) with kinetics similar to those of heat stress, as determined by the induction of HSF1 DNA binding, hyperphosphorylation of HSF1, and expression of chaperone genes. Celastrol can activate heat shock gene transcription synergistically with other stresses and exhibits cytoprotection against subsequent exposures to other forms of lethal cell stress. These results suggest that celastrols exhibit promise as a new class of pharmacologically active regulators of the heat shock response.

Ontology-driven Knowledge Coupling (GO)

Alterations in **protein folding** and the **regulation** of conformational states have become increasingly important to the functionality of key molecules in signaling, cell growth, and **cell death** . Molecular chaperones, because of their properties in protein quality control, afford conformational flexibility to proteins and serve to integrate stress-signaling events that influence **aging** and a range of diseases including cancer, cystic fibrosis, amyloidoses, and neurodegenerative diseases . We describe here characteristics of celastrol, a quinone methide triterpene and an active component from Chinese herbal medicine identified in a screen of bioactive small molecules that activates the human heat shock response . From a structure/function examination, the celastrol structure is remarkably specific and activates heat shock transcription factor 1 (HSF1) with kinetics similar to those of heat stress, as determined by the induction of HSF1 **DNA binding**, **hyperphosphorylation** of HSF1, and expression of chaperone genes . Celastrol can activate heat shock gene **transcription** synergistically with other stresses and exhibits cytoprotection against subsequent exposures to other forms of lethal cell stress . These results suggest that celastrols exhibit promise as a new class of pharmacologically active regulators of the heat shock response .

Gene Ontology Browser

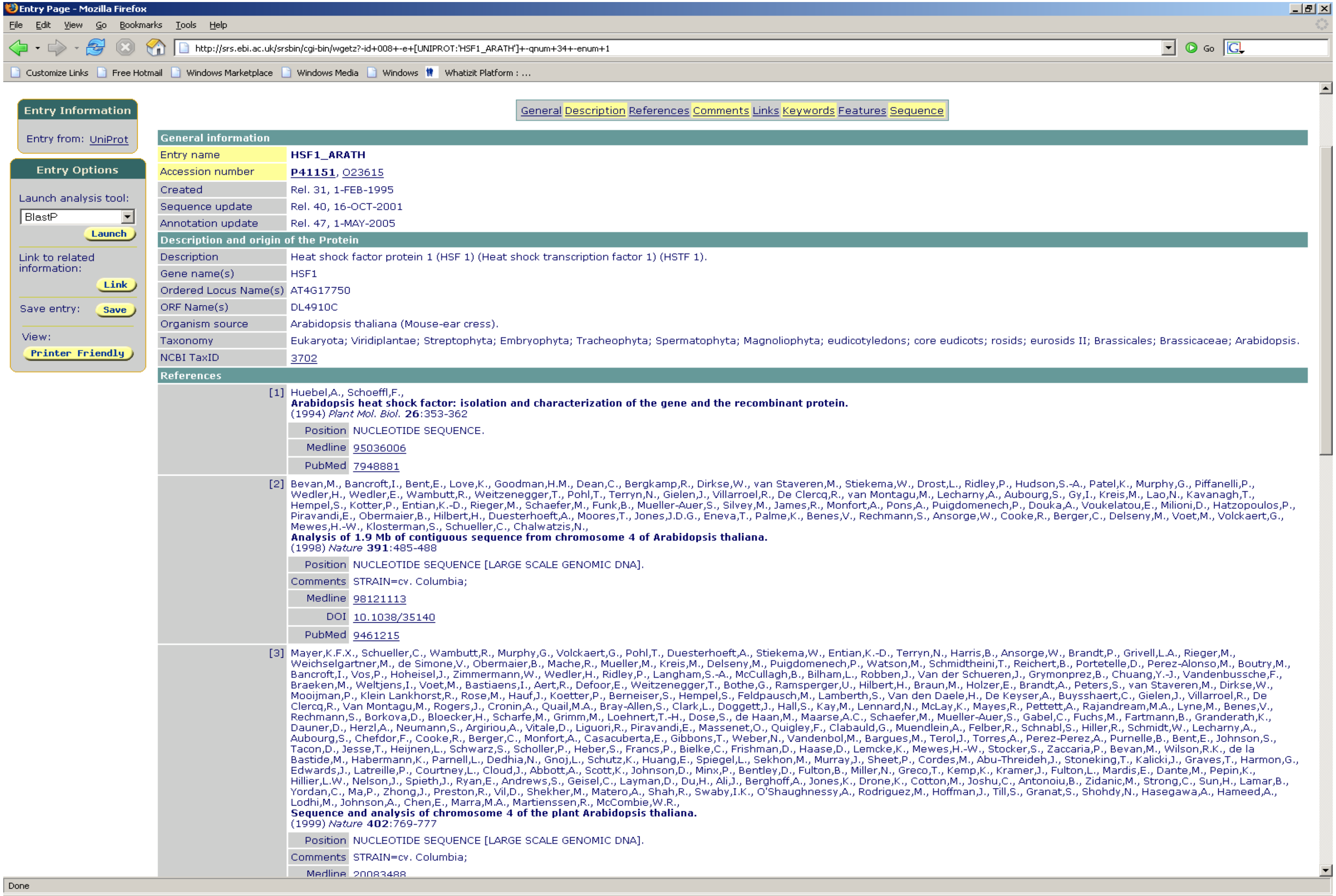


Term ID [?]	GO:0048151
Name [?]	hyperphosphorylation
Last updated [?]	2003-07-22 05:54:25.0
Definition [?]	The excessive phosphorylation of a protein, as a result of activation of kinases, deactivation of phosphatases, or both.
Comment [?]	Note that this term should be used only to annotate gene products that carry out hyperphosphorylations that are normal for the cell/organism.
Hierarchy [?]	View this term's parents in a denormalised tree. View with neither graph nor tree. Hide all selected terms except the primary one Add more terms to the selection with a search
Child terms [?]	GO:0016245: hyperphosphorylation of RNA polymerase II

Database-driven Knowledge Coupling (Swiss-Prot)

Alterations in protein folding and the regulation of conformational states have become increasingly important to the functionality of key molecules in signaling, cell growth, and cell death . Molecular chaperones, because of their properties in protein quality control, afford conformational flexibility to proteins and serve to integrate stress-signaling events that influence aging and a range of diseases including cancer, cystic fibrosis, amyloidoses, and neurodegenerative diseases . We describe here characteristics of celastrol, a quinone methide triterpene and an active component from Chinese herbal medicine identified in a screen of bioactive small molecules that activates the human heat shock response . From a structure/function examination, the celastrol structure is remarkably specific and activates **heat shock transcription factor 1** (**HSF1**) with kinetics similar to those of heat stress, as determined by the induction of **HSF1** DNA binding, hyperphosphorylation of **HSF1**, and expression of chaperone genes . Celastrol can activate heat shock gene transcription synergistically with other stresses and exhibits cytoprotection against subsequent exposures to other forms of lethal cell stress . These results suggest that celastrols exhibit promise as a new class of pharmacologically active regulators of the heat shock response .

Swiss-Prot Records



http://srs.ebi.ac.uk/srsbin/cgi-bin/wgetz?-id+008+-e+[UNIPROT:'HSF1_ARATH']+-qnum+34+-enum+1

General information

Entry name	**HSF1_ARATH**
Accession number	**P41151**, O23615
Created	Rel. 31, 1-FEB-1995
Sequence update	Rel. 40, 16-OCT-2001
Annotation update	Rel. 47, 1-MAY-2005

Description and origin of the Protein

Description	Heat shock factor protein 1 (HSF 1) (Heat shock transcription factor 1) (HSTF 1).
Gene name(s)	HSF1
Ordered Locus Name(s)	AT4G17750
ORF Name(s)	DL4910C
Organism source	Arabidopsis thaliana (Mouse-ear cress).
Taxonomy	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI TaxID	3702

References

[1] Huebel,A., Schoeffl,F.,
Arabidopsis heat shock factor: isolation and characterization of the gene and the recombinant protein.
(1994) *Plant Mol. Biol.* **26**:353-362

Position	NUCLEOTIDE SEQUENCE.
Medline	95036006
PubMed	7948881

[2] Bevan,M., Bancroft,I., Bent,E., Love,K., Goodman,H.M., Dean,C., Bergkamp,R., Dirkse,W., van Staveren,M., Stiekema,W., Drost,L., Ridley,P., Hudson,S.-A., Patel,K., Murphy,G., Piffanelli,P., Wedler,H., Wedler,E., Wambutt,R., Weitzenegger,T., Pohl,T., Terryn,N., Gielen,J., Villarroel,R., De Clercq,R., van Montagu,M., Lecharny,A., Aubourg,S., Gy,I., Kreis,M., Lao,N., Kavanagh,T., Hempel,S., Kotter,P., Entian,K.-D., Rieger,M., Schaefer,M., Funk,B., Mueller-Auer,S., Silvey,M., James,R., Monfort,A., Pons,A., Puigdomenech,P., Douka,A., Voukelatou,E., Milioni,D., Hatzopoulos,P., Piravandi,E., Obermaier,B., Hilbert,H., Duesterhoeft,A., Moores,T., Jones,J.D.G., Eneva,T., Palme,K., Benes,V., Rechmann,S., Ansorge,W., Cooke,R., Berger,C., Delseny,M., Voet,M., Volckaert,G., Mewes,H.-W., Klosterman,S., Schueller,C., Chalwatzis,N.,
Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana.
(1998) *Nature* **391**:485-488

Position	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
Comments	STRAIN=cv. Columbia;
Medline	98121113
DOI	10.1038/35140
PubMed	9461215

[3] Mayer,K.F.X., Schueller,C., Wambutt,R., Murphy,G., Volckaert,G., Pohl,T., Duesterhoeft,A., Stiekema,W., Entian,K.-D., Terryn,N., Harris,B., Ansorge,W., Brandt,P., Grivell,L.A., Rieger,M., Weichselgartner,M., de Simone,V., Obermaier,B., Mache,R., Mueller,M., Kreis,M., Delseny,M., Puigdomenech,P., Watson,M., Schmidtheini,T., Reichert,B., Portetelle,D., Perez-Alonso,M., Boutry,M., Bancroft,I., Vos,P., Hoheisel,J., Zimmermann,W., Wedler,H., Ridley,P., Langham,S.-A., McCullagh,B., Bilham,L., Robben,J., Van der Schueren,J., Grymonprez,B., Chuang,Y.-J., Vandenbussche,F., Braeken,M., Weltjens,I., Voet,M., Bastiaens,I., Aert,R., Defoor,E., Weitzenegger,T., Bothe,G., Ramsperger,U., Hilbert,H., Braun,M., Holzer,E., Brandt,A., Peters,S., van Staveren,M., Dirkse,W., Mooijman,P., Klein Lankhorst,R., Rose,M., Hauf,J., Koetter,P., Berneiser,S., Hempel,S., Feldpausch,M., Lamberth,S., Van den Daele,H., De Keyser,A., Buysshaert,C., Gielen,J., Villarroel,R., De Clercq,R., Van Montagu,M., Rogers,J., Cronin,A., Quail,M.A., Bray-Allen,S., Clark,L., Doggett,J., Hall,S., Kay,M., Lennard,N., McLay,K., Mayes,R., Pettett,A., Rajandream,M.A., Lyne,M., Benes,V., Rechmann,S., Borkova,D., Bloecker,H., Scharfe,M., Grimm,M., Loehnert,T.-H., Dose,S., de Haan,M., Maarse,A.C., Schaefer,M., Mueller-Auer,S., Gabel,C., Fuchs,M., Fartmann,B., Granderath,K., Dauner,D., Herzl,A., Neumann,S., Argiriou,A., Vitale,D., Liguori,R., Piravandi,E., Massenet,O., Quigley,F., Clabauld,G., Muendlein,A., Felber,R., Schnabl,S., Hiller,R., Schmidt,W., Lecharny,A., Aubourg,S., Chefdor,F., Cooke,R., Berger,C., Monfort,A., Casacuberta,E., Gibbons,T., Weber,N., Vandenbol,M., Bargues,M., Terol,J., Torres,A., Perez-Perez,A., Purnelle,B., Bent,E., Johnson,S., Tacon,D., Jesse,T., Heijnen,L., Schwarz,S., Scholler,P., Heber,S., Francs,P., Bielke,C., Frishman,D., Haase,D., Lemcke,K., Mewes,H.-W., Stocker,S., Zaccaria,P., Bevan,M., Wilson,R.K., de la Bastide,M., Habermann,K., Parnell,L., Dedhia,N., Gnoj,L., Schutz,K., Huang,E., Spiegel,L., Sekhon,M., Murray,J., Sheet,P., Cordes,M., Abu-Threideh,J., Stoneking,T., Kalicki,J., Graves,T., Harmon,G., Edwards,J., Latreille,P., Courtney,L., Cloud,J., Abbott,A., Scott,K., Johnson,D., Minx,P., Bentley,D., Fulton,B., Miller,N., Greco,T., Kemp,K., Kramer,J., Fulton,L., Mardis,E., Dante,M., Pepin,K., Hillier,L.W., Nelson,J., Spieth,J., Ryan,E., Andrews,S., Geisel,C., Layman,D., Du,H., Ali,J., Berghoff,A., Jones,K., Drone,K., Cotton,M., Joshu,C., Antonoiu,B., Zidanic,M., Strong,C., Sun,H., Lamar,B., Yordan,C., Ma,P., Zhong,J., Preston,R., Vil,D., Shekher,M., Matero,A., Shah,R., Swaby,I.K., O'Shaughnessy,A., Rodriguez,M., Hoffman,J., Till,S., Granat,S., Shohdy,N., Hasegawa,A., Hameed,A., Lodhi,M., Johnson,A., Chen,E., Marra,M.A., Martienssen,R., McCombie,W.R.,
Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
(1999) *Nature* **402**:769-777

Position	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
Comments	STRAIN=cv. Columbia;
Medline	20083488

Evaluation

- Q2. Sharing of resources and use of research software tools

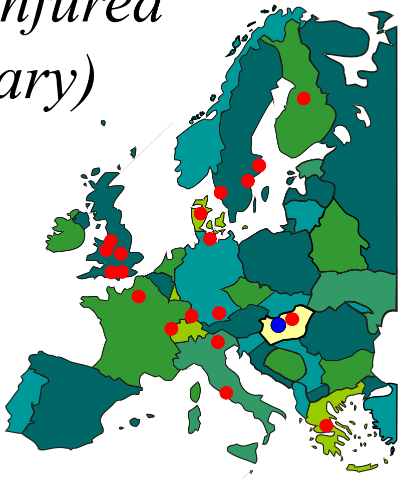
  Good
- Q6. Short and medium-term visits

  To be improved
- Q7. Co-authoring of research papers, PhD…

  To be improved

Summer School : July 4th-10th 2004

Balatonfured
(Hungary)



© Zentai László, 2000

Summer School : July 4th-10th 2004

2004 Summer School: A Summary

DELEGATES

- 80 participants
  - 18 out of 23 partner sites represented
  - 9 granted non-NoE PhD Students
- 29 Speakers
  - 1 invited (Cornelius Rosse)
  - 28 from NoE
- 19 Student Posters

PROGRAMME

- Ontology masterclass
- 1 day workshops
  - Ontology
  - Semantic Web
  - Health Statistics
- 2nd Assembly Meeting
- Social Programme

Mobility Program

- Objective: Exchange of PhD students
- Inventory of PhD-study programmes, procedures
- Launch of mobility program March 2005
  - One or two medium-term visits
  - 10-15 short-term visits, 1 week - 1 month

Knowledge sharing

- Workshop on the Gene Ontology, Leipzig, May 29
- Workshop on NLP for Biomedical Applications at the COLING conference, Geneva, August 23-27
- TERMINFO and Scientific Advisory Committee at MEDINFO2004
- WHO-FIC meeting on Classifications in Health Care, Reykjavik, Iceland, October 24-30
- Description Logics and SNOMED CT, Saarbrücken, Nov 22-23
- Workshop on EHR at Satellite Conference to EUROREC, Brussels, Nov 25-27
- Workshop on Mereotopolgy in Freiburg, Jan 23-24

WP13: Workshop on Natural Language Processing

- Goals:
  1. expand visibility of the semanticmining workshop;
  2. establish forum for outside/inside network cooperation;
  3. federate the NLP community in the biomedical domain;
  4. organize a shared task to stimulate research in the domain, following well established challenges such as the TREC Genomics (http://trec.nist.gov/) or BioCreative(http://www.pdg.cnb.uam.es/BioLINK/BioCreative.eval.html).

Workshop

- Audience
  - Satellite of COLING: computer scientists, linguists, logicians…
  - Natural Language Processing/Information Retrieval
  - Medical informatics and Bioinformatics
  - 60 registered participants
- Distribution
  - Table
- Paper selection
  - 7 regular papers out of 30 submissions
  - 5 posters
- Dissemination
  - Workshop printed proceedings
  - Website
  - Special issue under preparation (IJMI - Elsevier)

Shared Task I

- Background
  - Information access tools is increasing to support literature survey,
  - Online ‘portals’ where scientists can navigate
  - Genetics and disease databases
  - Ambiguous nomenclature: Gene/RNA/proteins
  - Scale up methods for processing full text articles etc.
- Task
  - Annotate Gene and Protein Names (GPNs)

    i.e. find beginning and end of GPNs

Shared task II

- MEDLINE Corpus

  Trained on 2000 abstracts / Tested on 200

- Evaluation

  IOB recall and precision-like metrics

- Participation
  - 12 participant team

Knowledge sharing

- standardisation activities performed in e.g. CEN TC251 and HL7
- developers of the Foundational Model of Anatomy (FMA)
- developers of the Gene Ontology (GO)
- developers of SNOMED CT
- developers of IUPAC and LOINC (in the area of laboratory medicine)

Upcoming Events

- Symposium on Semantic Mining in Biomedicine
  - EBI, April 10-13
- Ontology and Biomedical Informatics
  - Rome, April 29 - May 2, in cooperation with IMIA WG6,
- Workshop on SNOMED CT
  - Date and place to be fixed
- Workshop on Human issues in handling large scale ontologies
  - AIME/IJCAI, Aberdeen, July 24-27
- Workshops at Summer School, June 29 - July 4
  - The “Boundary problem” between Information and Terminology models
  - The Semantic Web
  - Concept systems in laboratory medicine
  - Text mining from EHRs
  - Gender issues in computer science
- Check this … www.semanticmining.org