Using a Database Wiki for Biological Database Curation

James Cheney
University of Edinburgh

with Heiko Mueller, Sam Lindley,
Jo Sharman, Tony Harmar, and Peter Buneman

Wikis vs databases

- Much easier to get started using wikis
  - easy to edit, versioning for free
  - don't have to think hard about structure
- But many projects (eventually) need:
  - structured data, fast querying, access control
  - *databases!*

Where we come in

- Our background:
  - Databases, programming languages
- Our research:
  - Archiving/versioning for structured data
  - Provenance for DB queries and updates
  - Advanced Web programming languages

Database Wikis

- Goal: Make databases usable by everyone.
  - Secondary: A general database curation platform
- Add tree-structured *data resources* to wikis
  - editable via Web browser; can import data as XML
  - support annotation, provenance, versioning
- Provide for *embedding* data into Wiki pages
  - path queries, "stylesheets", updatable "views"

Database Wikis

- Goal: Make databases usabl
  - Secondary: A general databa
- Add tree-structured *data re*
  - editable via Web browser; can import data as XML
  - support annotation, provenance, versioning
- Provide for *embedding* data into Wiki pages
  - path queries, "stylesheets", updatable "views"

... a lot like Semantic MediaWiki, but allows for **nested** data

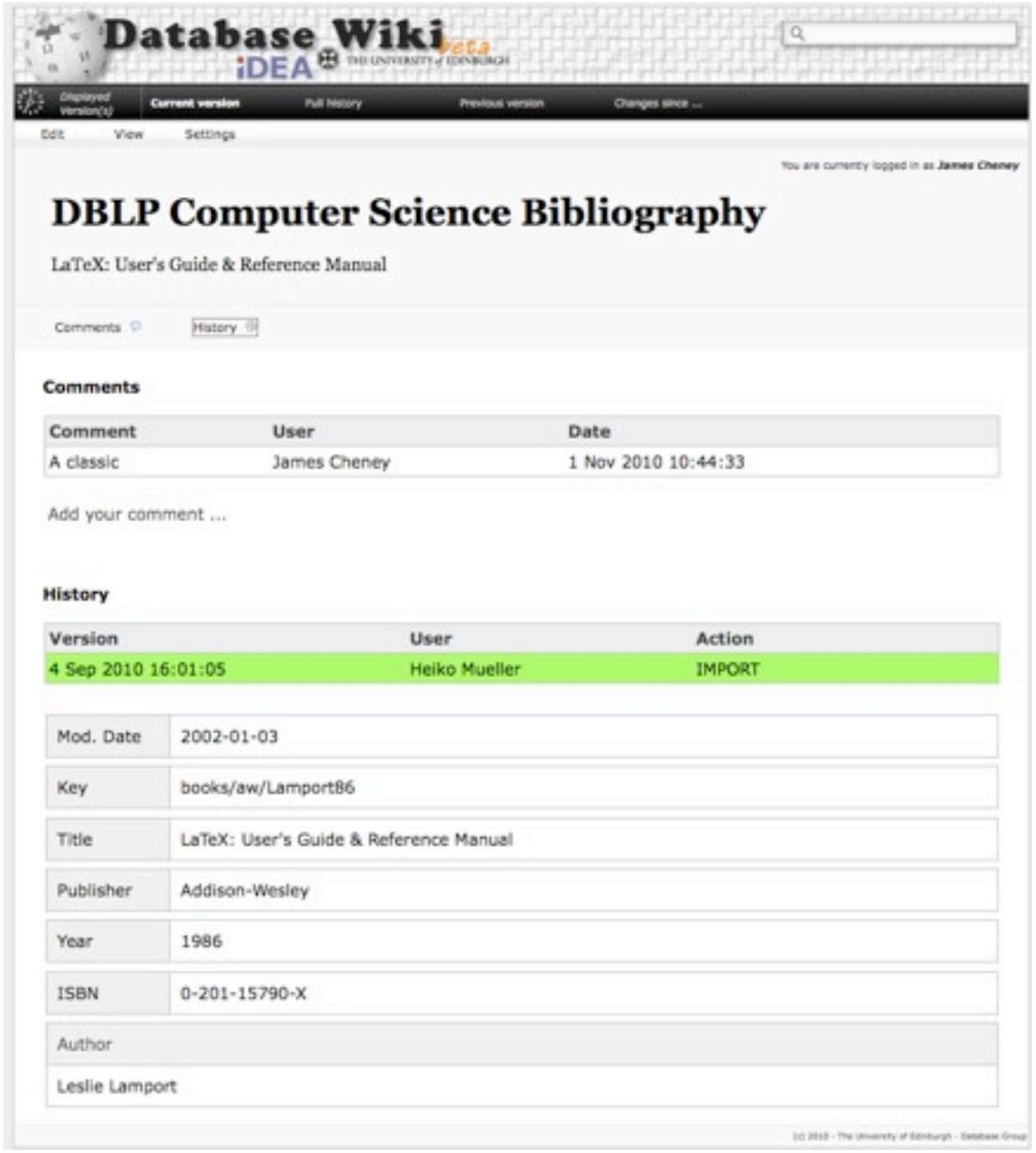
Demo

Data in DBWiki

Queries embedded in wiki pages

Updating data through forms

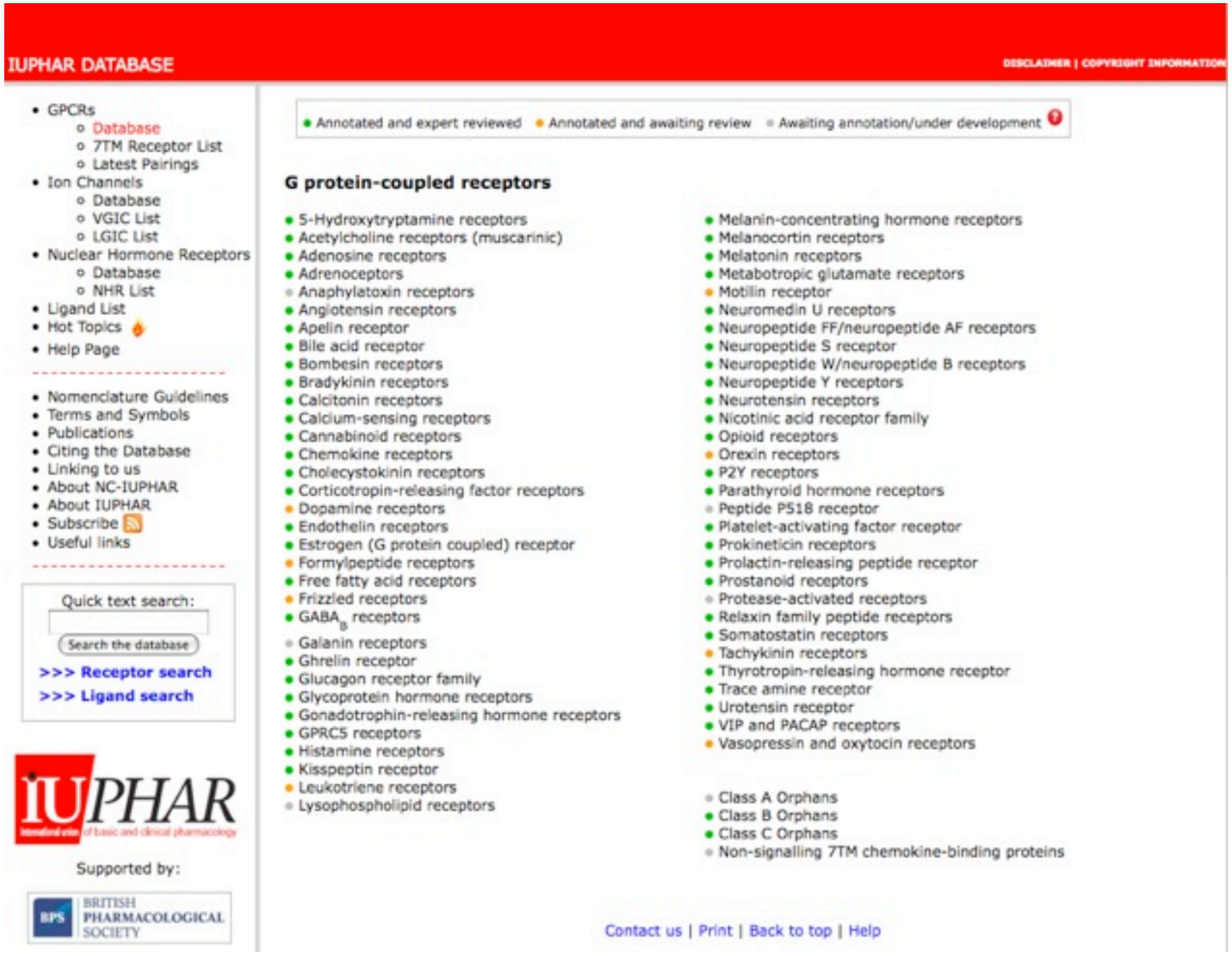
Annotation and provenance



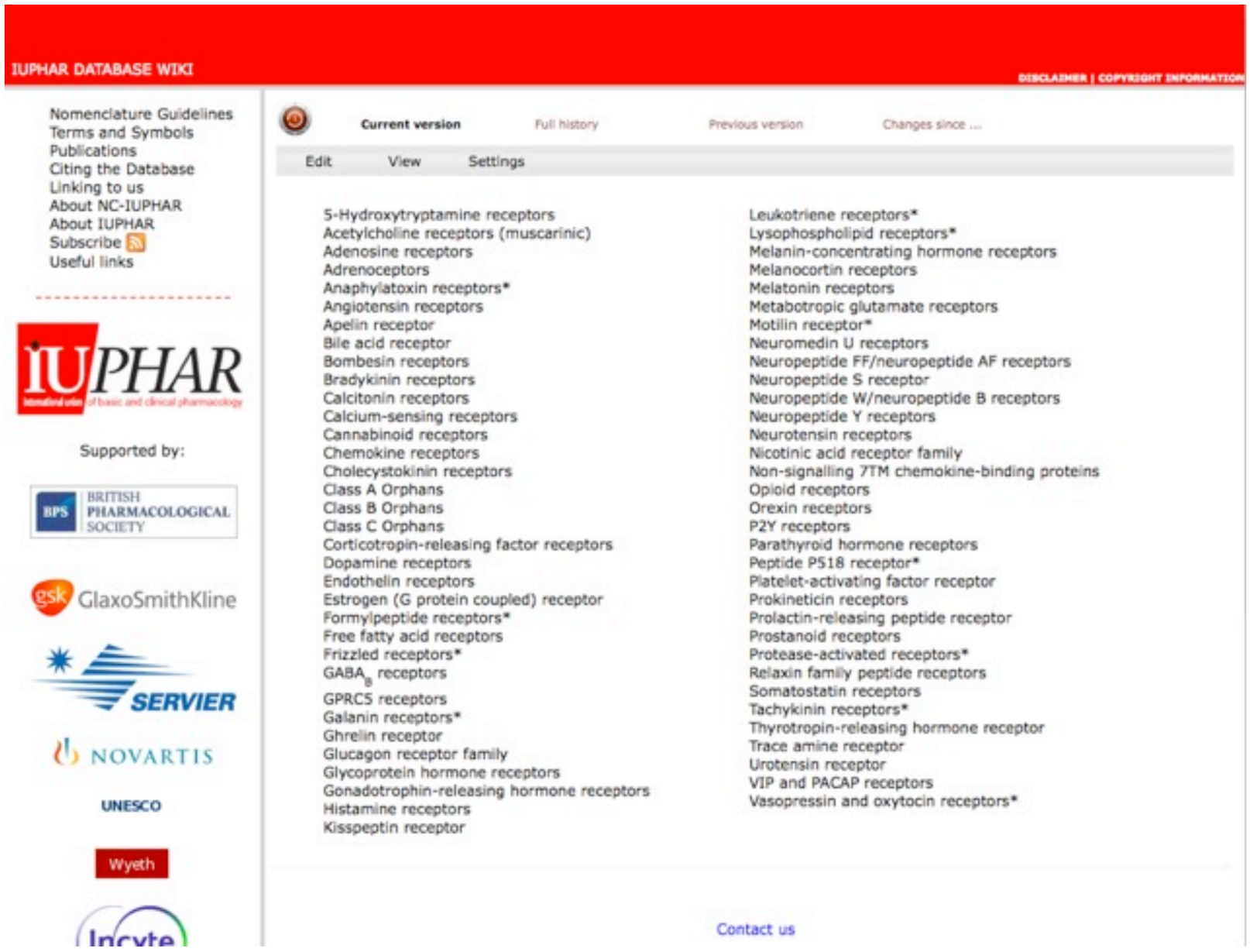
Case study

- Applied to IUPHAR database (www.iuphar-db.org)
- DBWiki appearance can be customized through templates, style sheets
- Close to look and feel of original
  - not an exact match
  - but provides archiving, annotation, provenance for free

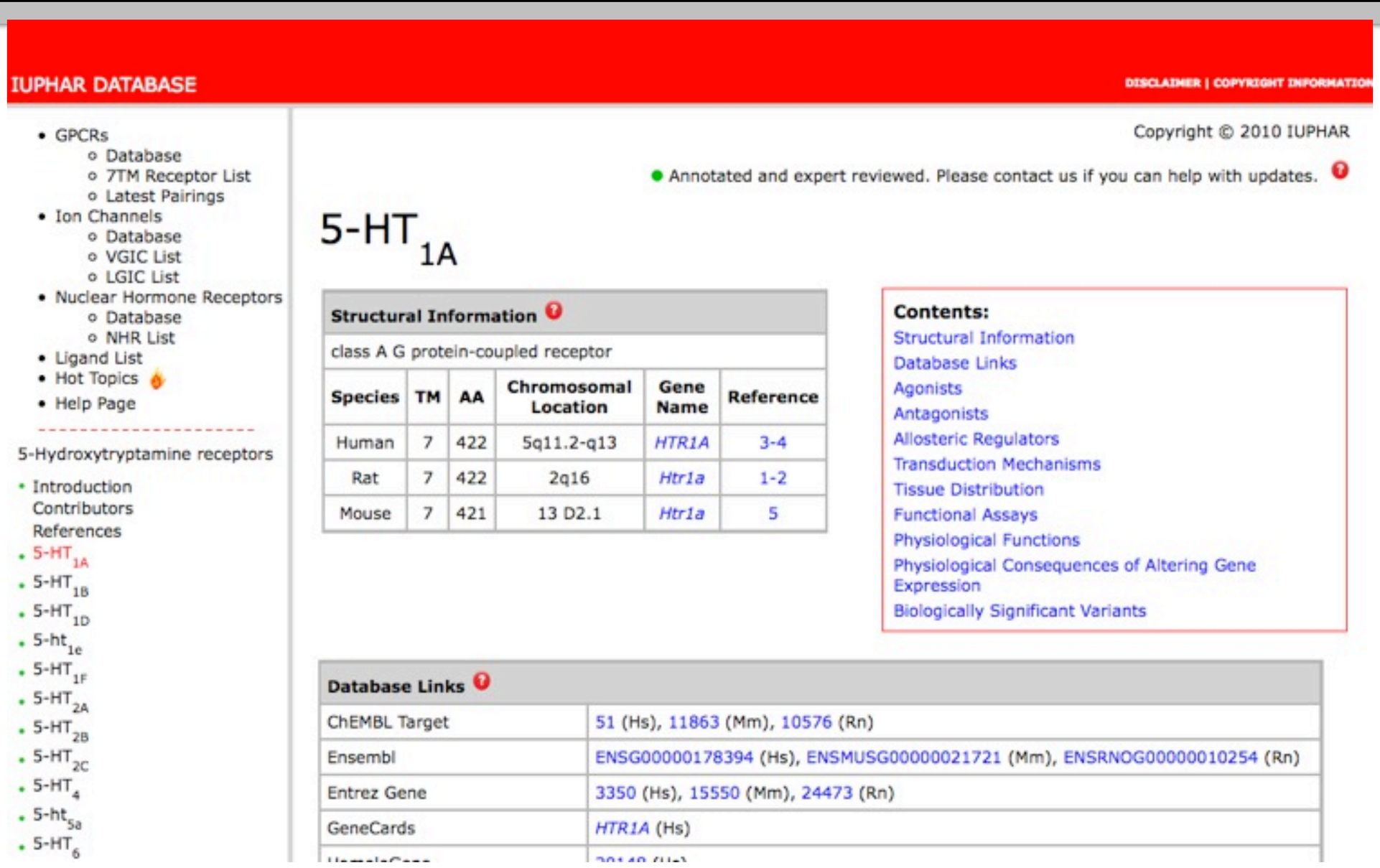
Real IUPHAR


IUPHAR DATABASE
DISCLAIMER | COPYRIGHT INFORMATION
GPCRs
Database
7TM Receptor List
Latest Pairings
Ion Channels
Database
VGIC List
LGIC List
Nuclear Hormone Receptors
Database
NHR List
Ligand List
Hot Topics
Help Page
Nomenclature Guidelines
Terms and Symbols
Publications
Citing the Database
Linking to us
About NC-IUPHAR
About IUPHAR
Subscribe
Useful links
Quick text search:
Search the database
>>> Receptor search
>>> Ligand search
Supported by:
BRITISH PHARMACOLOGICAL SOCIETY
Annotated and expert reviewed
Annotated and awaiting review
Awaiting annotation/under development
G protein-coupled receptors
5-Hydroxytryptamine receptors
Acetylcholine receptors (muscarinic)
Adenosine receptors
Adrenoceptors
Anaphylatoxin receptors
Angiotensin receptors
Apelin receptor
Bile acid receptor
Bombesin receptors
Bradykinin receptors
Calcitonin receptors
Calcium-sensing receptors
Cannabinoid receptors
Chemokine receptors
Cholecystokinin receptors
Corticotropin-releasing factor receptors
Dopamine receptors
Endothelin receptors
Estrogen (G protein coupled) receptor
Formylpeptide receptors
Free fatty acid receptors
Frizzled receptors
GABA$_B$ receptors
Galanin receptors
Ghrelin receptor
Glucagon receptor family
Glycoprotein hormone receptors
Gonadotrophin-releasing hormone receptors
GPRC5 receptors
Histamine receptors
Kisspeptin receptor
Leukotriene receptors
Lysophospholipid receptors
Melanin-concentrating hormone receptors
Melanocortin receptors
Melatonin receptors
Metabotropic glutamate receptors
Motilin receptor
Neuromedin U receptors
Neuropeptide FF/neuropeptide AF receptors
Neuropeptide S receptor
Neuropeptide W/neuropeptide B receptors
Neuropeptide Y receptors
Neurotensin receptors
Nicotinic acid receptor family
Opioid receptors
Orexin receptors
P2Y receptors
Parathyroid hormone receptors
Peptide P518 receptor
Platelet-activating factor receptor
Prokineticin receptors
Prolactin-releasing peptide receptor
Prostanoid receptors
Protease-activated receptors
Relaxin family peptide receptors
Somatostatin receptors
Tachykinin receptors
Thyrotropin-releasing hormone receptor
Trace amine receptor
Urotensin receptor
VIP and PACAP receptors
Vasopressin and oxytocin receptors
Class A Orphans
Class B Orphans
Class C Orphans
Non-signalling 7TM chemokine-binding proteins
Contact us | Print | Back to top | Help

IUPHAR in DBWiki



Real IUPHAR



IUPHAR DATABASE

DISCLAIMER | COPYRIGHT INFORMATION

- GPCRs
  - Database
  - 7TM Receptor List
  - Latest Pairings
- Ion Channels
  - Database
  - VGIC List
  - LGIC List
- Nuclear Hormone Receptors
  - Database
  - NHR List
- Ligand List
- Hot Topics
- Help Page

5-Hydroxytryptamine receptors

- Introduction
  Contributors
  References
- 5-HT$_{1A}$
- 5-HT$_{1B}$
- 5-HT$_{1D}$
- 5-ht$_{1e}$
- 5-HT$_{1F}$
- 5-HT$_{2A}$
- 5-HT$_{2B}$
- 5-HT$_{2C}$
- 5-HT$_{4}$
- 5-ht$_{5a}$
- 5-HT$_{6}$

Annotated and expert reviewed. Please contact us if you can help with updates.

5-HT$_{1A}$

Structural Information

class A G protein-coupled receptor

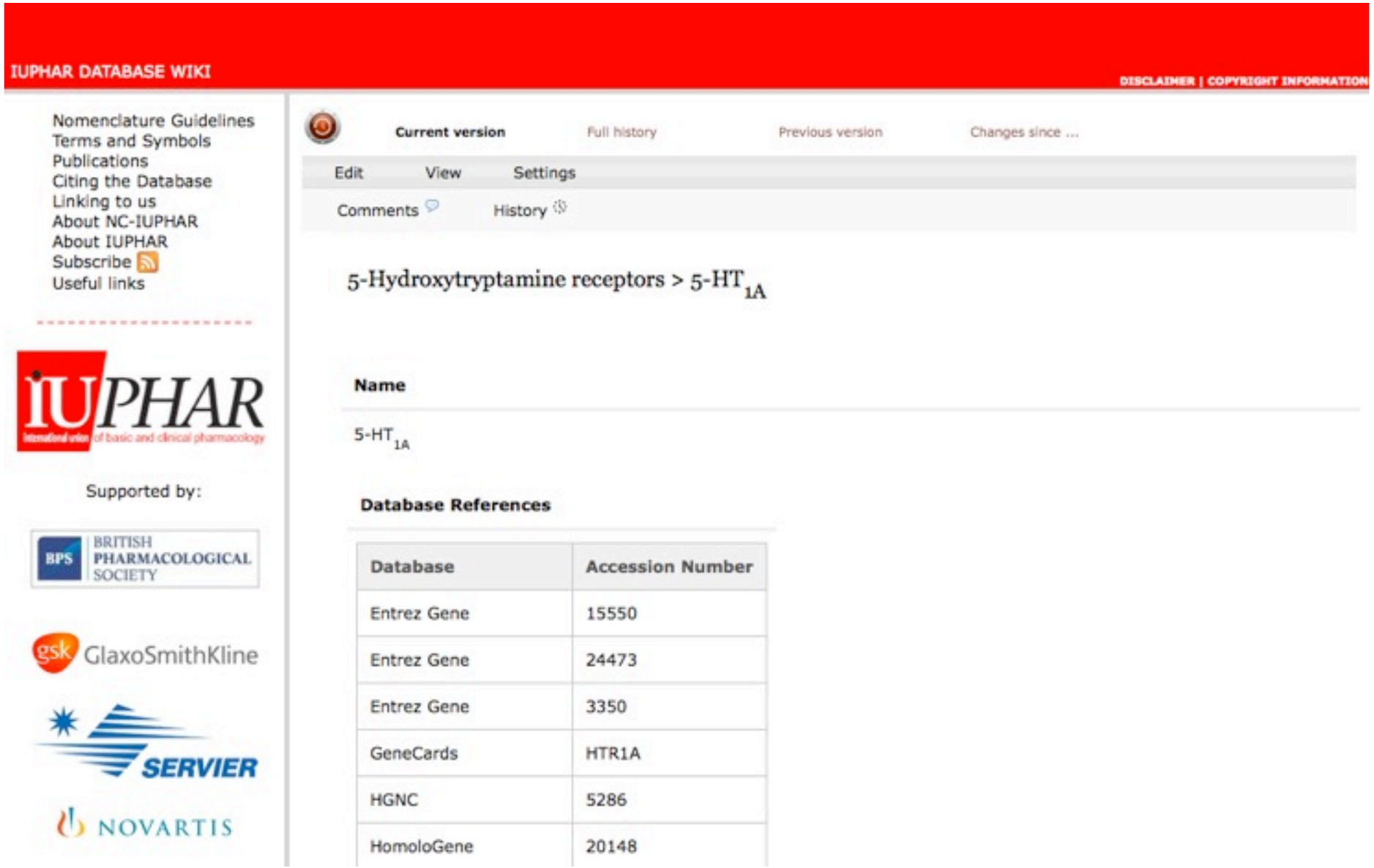
Species	TM	AA	Chromosomal Location	Gene Name	Reference
Human	7	422	5q11.2-q13	*HTR1A*	3-4
Rat	7	422	2q16	*Htr1a*	1-2
Mouse	7	421	13 D2.1	*Htr1a*	5

Contents:

Database Links

ChEMBL Target	51 (Hs), 11863 (Mm), 10576 (Rn)
Ensembl	ENSG00000178394 (Hs), ENSMUSG00000021721 (Mm), ENSRNOG00000010254 (Rn)
Entrez Gene	3350 (Hs), 15550 (Mm), 24473 (Rn)
GeneCards	*HTR1A* (Hs)

IUPHAR in DBWiki



Next steps

- Scaling up to big data, existing sources
  - efficient querying on unstructured sources
  - import/annotation vs. refresh issue
- Structure discovery/learning?
  - can we extract useful *structure* from unstructured wiki page contributions?
- Getting UI and functionality right
  - volunteers?

Would this be useful to you? Want to help?

- Play with prototype, let us know what breaks

  http://forsberg.inf.ed.ac.uk:8080

- mailing list: db-wiki@inf.ed.ac.uk

- plug: W3C Provenance Incubator Group
  - industry/bioinformatics involvement sought for WG
  - http://www.w3.org/2005/Incubator/prov/wiki/