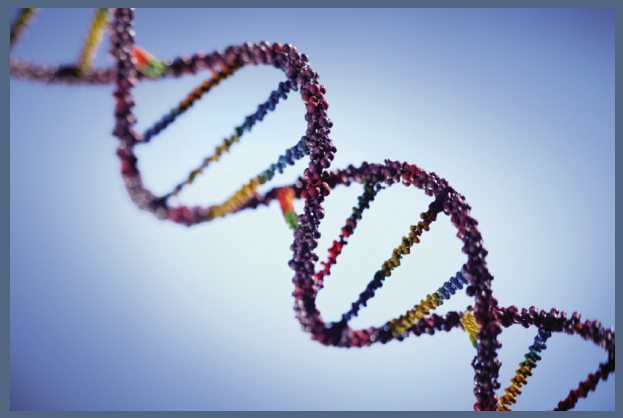
Centre for Molecular Medicine and Therapeutics


# The wiki-based Transcription Factor Encyclopedia and a model for robust community participation

Wyeth W. Wasserman

30 November 2010

http://www.cisreg.ca/tfe

# Transcription Factors 101

- **Diverse set of proteins sharing common functional role in the regulation and production of RNA transcripts**
- **Present in all species**
- **For the purpose of today's talk, TF will refer to the subset of the proteins that bind to DNA in a sequence-specific manner**

# Transcription Factors 102

- **Act cooperatively with other TFs to confer specific patterns of gene activity in response to developmental, physiological and environmental conditions**
- **Much of their function is defined by protein-protein interactions**
- **Understanding TFs is key to understanding developmental and tissue differentiation**

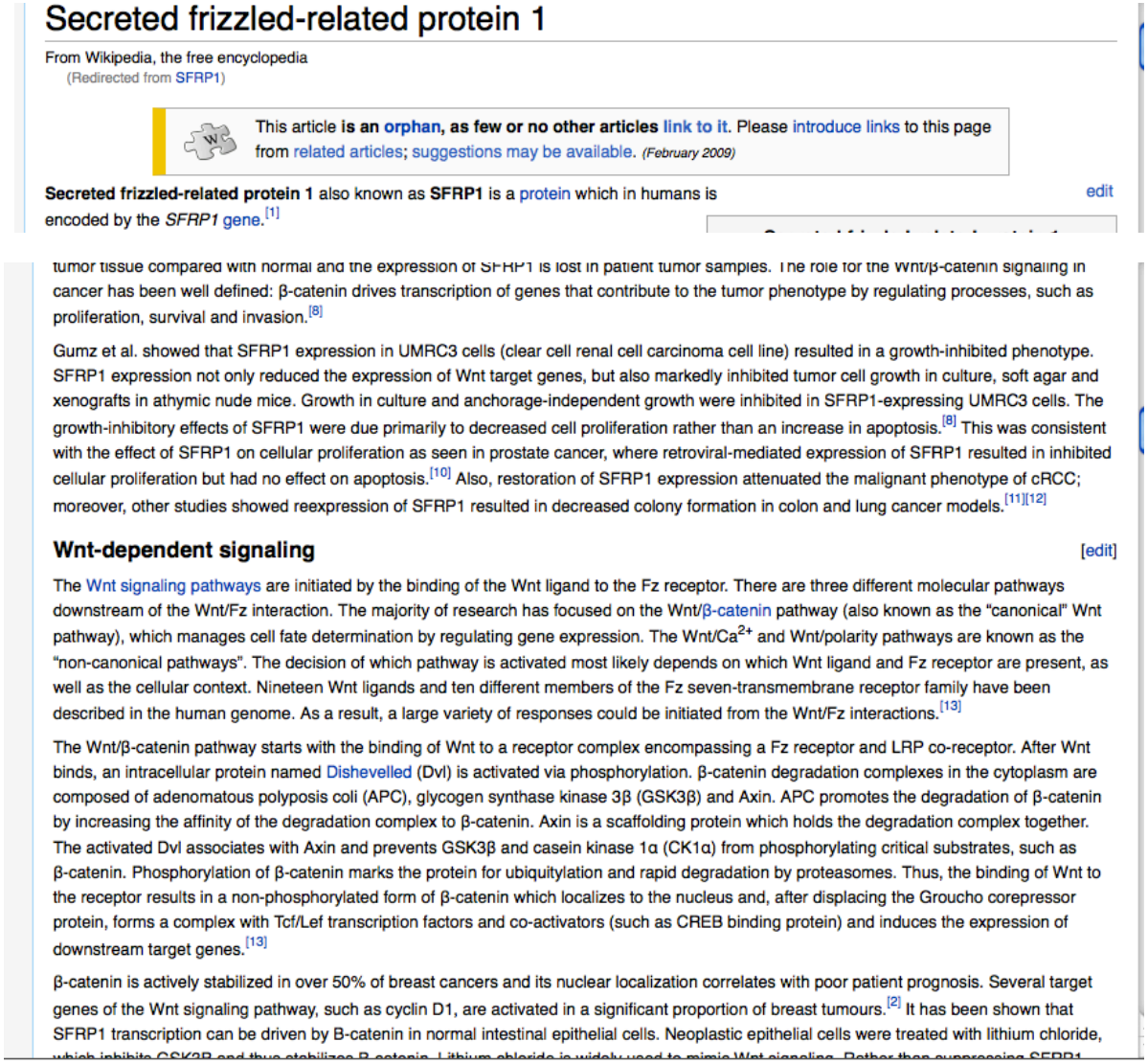
# Resources for the Study of TFs



TFCAT


CMMT

# Wiki-in-the-classroom



## Secreted frizzled-related protein 1

From Wikipedia, the free encyclopedia
(Redirected from SFRP1)

This article **is an orphan, as few or no other articles link to it**. Please introduce links to this page from related articles; suggestions may be available. *(February 2009)*

**Secreted frizzled-related protein 1** also known as **SFRP1** is a protein which in humans is encoded by the *SFRP1* gene.[1] edit

tumor tissue compared with normal and the expression of SFRP1 is lost in patient tumor samples. The role for the Wnt/β-catenin signaling in cancer has been well defined: β-catenin drives transcription of genes that contribute to the tumor phenotype by regulating processes, such as proliferation, survival and invasion.[8]

Gumz et al. showed that SFRP1 expression in UMRC3 cells (clear cell renal cell carcinoma cell line) resulted in a growth-inhibited phenotype. SFRP1 expression not only reduced the expression of Wnt target genes, but also markedly inhibited tumor cell growth in culture, soft agar and xenografts in athymic nude mice. Growth in culture and anchorage-independent growth were inhibited in SFRP1-expressing UMRC3 cells. The growth-inhibitory effects of SFRP1 were due primarily to decreased cell proliferation rather than an increase in apoptosis.[8] This was consistent with the effect of SFRP1 on cellular proliferation as seen in prostate cancer, where retroviral-mediated expression of SFRP1 resulted in inhibited cellular proliferation but had no effect on apoptosis.[10] Also, restoration of SFRP1 expression attenuated the malignant phenotype of cRCC; moreover, other studies showed reexpression of SFRP1 resulted in decreased colony formation in colon and lung cancer models.[11][12]

### Wnt-dependent signaling [edit]

The Wnt signaling pathways are initiated by the binding of the Wnt ligand to the Fz receptor. There are three different molecular pathways downstream of the Wnt/Fz interaction. The majority of research has focused on the Wnt/β-catenin pathway (also known as the "canonical" Wnt pathway), which manages cell fate determination by regulating gene expression. The Wnt/$Ca^{2+}$ and Wnt/polarity pathways are known as the "non-canonical pathways". The decision of which pathway is activated most likely depends on which Wnt ligand and Fz receptor are present, as well as the cellular context. Nineteen Wnt ligands and ten different members of the Fz seven-transmembrane receptor family have been described in the human genome. As a result, a large variety of responses could be initiated from the Wnt/Fz interactions.[13]

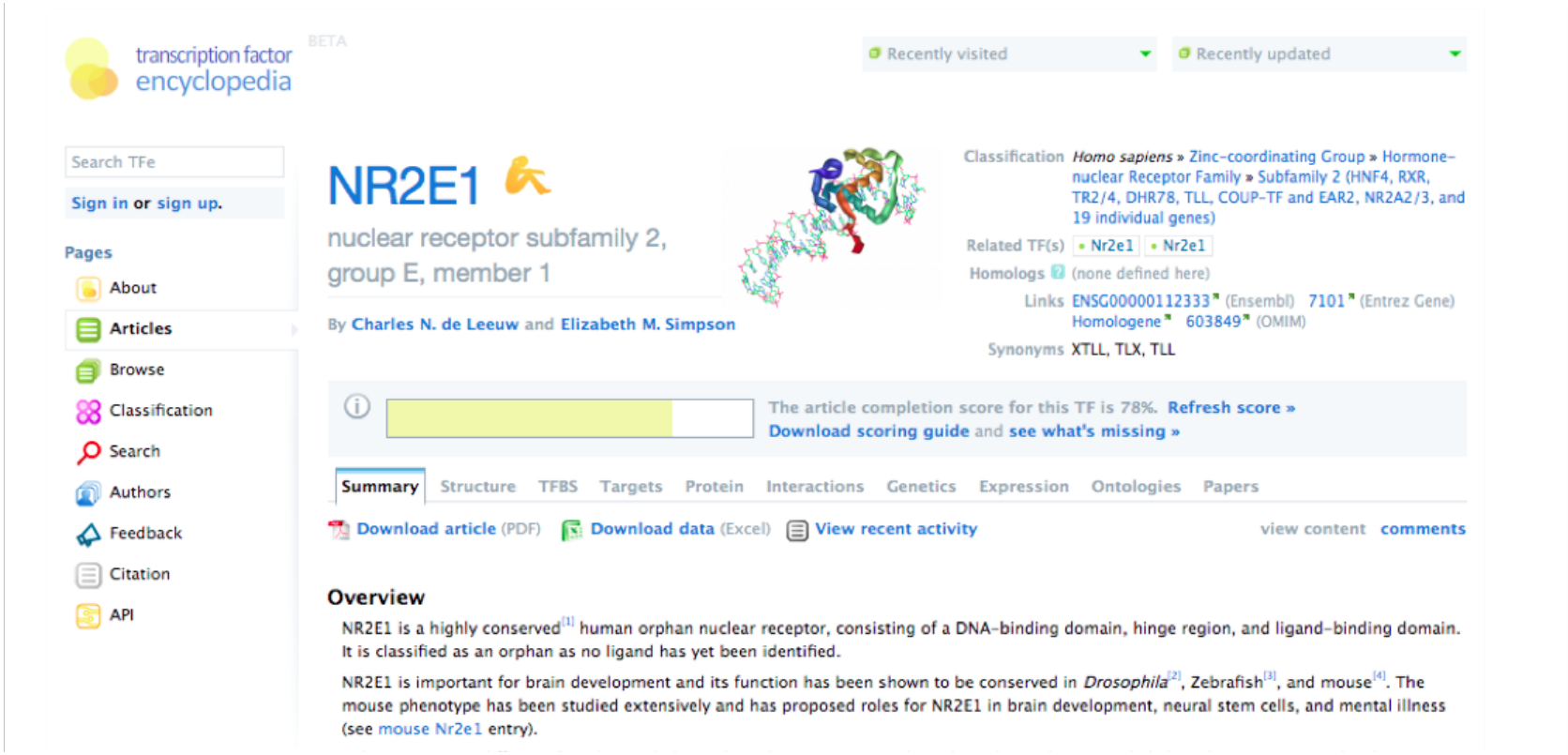
The Wnt/β-catenin pathway starts with the binding of Wnt to a receptor complex encompassing a Fz receptor and LRP co-receptor. After Wnt binds, an intracellular protein named Dishevelled (Dvl) is activated via phosphorylation. β-catenin degradation complexes in the cytoplasm are composed of adenomatous polyposis coli (APC), glycogen synthase kinase 3β (GSK3β) and Axin. APC promotes the degradation of β-catenin by increasing the affinity of the degradation complex to β-catenin. Axin is a scaffolding protein which holds the degradation complex together. The activated Dvl associates with Axin and prevents GSK3β and casein kinase 1α (CK1α) from phosphorylating critical substrates, such as β-catenin. Phosphorylation of β-catenin marks the protein for ubiquitylation and rapid degradation by proteasomes. Thus, the binding of Wnt to the receptor results in a non-phosphorylated form of β-catenin which localizes to the nucleus and, after displacing the Groucho corepressor protein, forms a complex with Tcf/Lef transcription factors and co-activators (such as CREB binding protein) and induces the expression of downstream target genes.[13]

β-catenin is actively stabilized in over 50% of breast cancers and its nuclear localization correlates with poor patient prognosis. Several target genes of the Wnt signaling pathway, such as cyclin D1, are activated in a significant proportion of breast tumours.[2] It has been shown that SFRP1 transcription can be driven by B-catenin in normal intestinal epithelial cells. Neoplastic epithelial cells were treated with lithium chloride, which inhibits GSK3B and thus stabilizes B-catenin. Lithium chloride is widely used to mimic Wnt signaling. Rather than suppressing SFRP1

transcription factor
encyclopedia

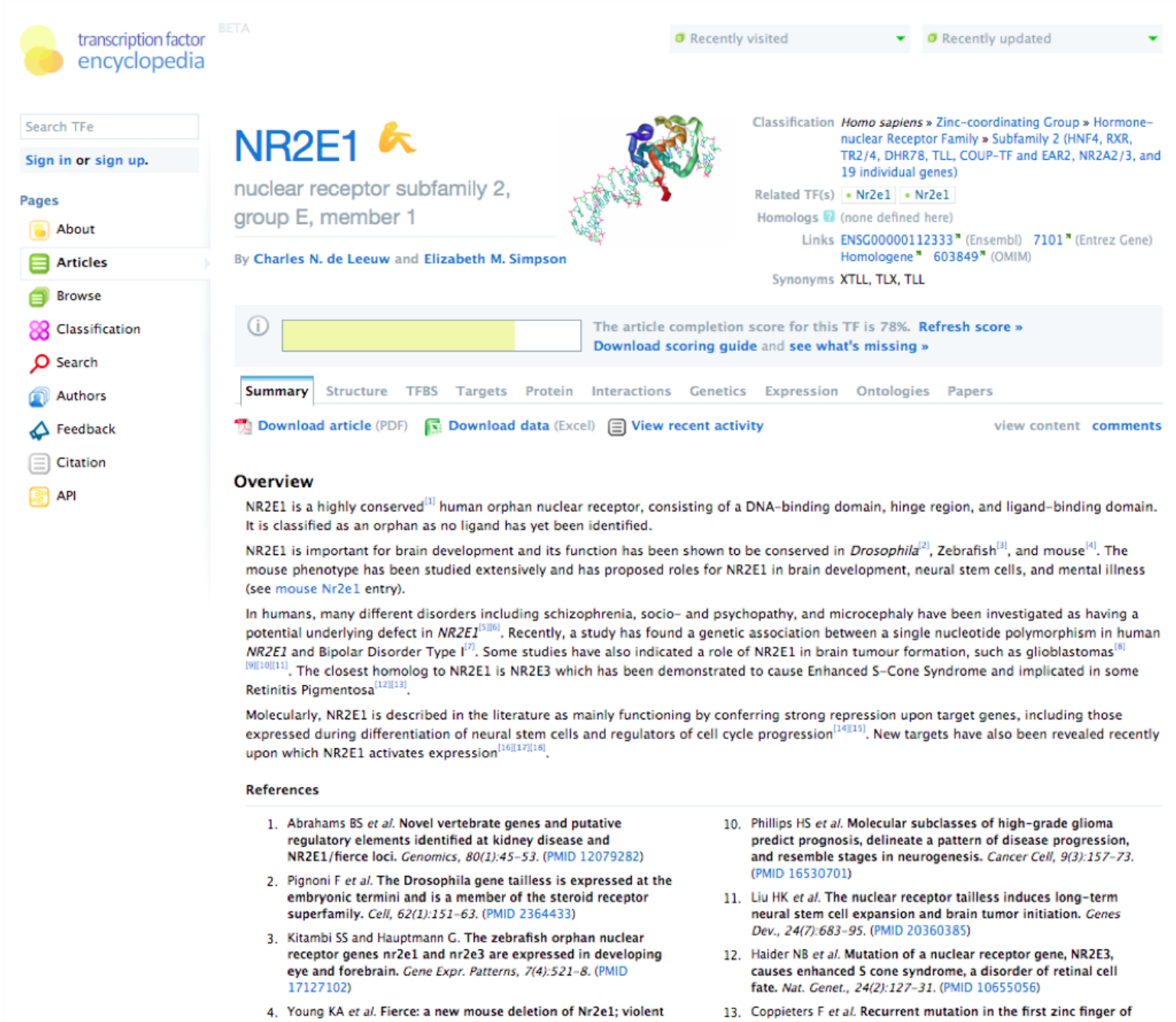
# TFe

**Goal:** create an online encyclopedic collection of reviews about well-studied TFs, combining a mixture of **expert-curated** and automatic content.




**http://www.cisreg.ca/tfe**

# Reader Interface



transcription factor encyclopedia BETA

Recently visited | Recently updated

Search TFe

Sign in or sign up.

Pages

- About
- Articles
- Browse
- Classification
- Search
- Authors
- Feedback
- Citation
- API

## NR2E1

nuclear receptor subfamily 2, group E, member 1

By Charles N. de Leeuw and Elizabeth M. Simpson

Classification *Homo sapiens* » Zinc-coordinating Group » Hormone-nuclear Receptor Family » Subfamily 2 (HNF4, RXR, TR2/4, DHR78, TLL, COUP-TF and EAR2, NR2A2/3, and 19 individual genes)

Related TF(s) • Nr2e1 • Nr2e1

Homologs (none defined here)

Links ENSG00000112333 (Ensembl) 7101 (Entrez Gene) Homologene 603849 (OMIM)

Synonyms XTLL, TLX, TLL

The article completion score for this TF is 78%. Refresh score »
Download scoring guide and see what's missing »

Summary | Structure | TFBS | Targets | Protein | Interactions | Genetics | Expression | Ontologies | Papers

Download article (PDF) Download data (Excel) View recent activity view content comments

### Overview

NR2E1 is a highly conserved[1] human orphan nuclear receptor, consisting of a DNA-binding domain, hinge region, and ligand-binding domain. It is classified as an orphan as no ligand has yet been identified.

NR2E1 is important for brain development and its function has been shown to be conserved in *Drosophila*[2], Zebrafish[3], and mouse[4]. The mouse phenotype has been studied extensively and has proposed roles for NR2E1 in brain development, neural stem cells, and mental illness (see mouse Nr2e1 entry).

In humans, many different disorders including schizophrenia, socio- and psychopathy, and microcephaly have been investigated as having a potential underlying defect in *NR2E1*[5][6]. Recently, a study has found a genetic association between a single nucleotide polymorphism in human *NR2E1* and Bipolar Disorder Type I[7]. Some studies have also indicated a role of NR2E1 in brain tumour formation, such as glioblastomas[8][9][10][11]. The closest homolog to NR2E1 is NR2E3 which has been demonstrated to cause Enhanced S-Cone Syndrome and implicated in some Retinitis Pigmentosa[12][13].

Molecularly, NR2E1 is described in the literature as mainly functioning by conferring strong repression upon target genes, including those expressed during differentiation of neural stem cells and regulators of cell cycle progression[14][15]. New targets have also been revealed recently upon which NR2E1 activates expression[16][17][18].

### References

1. Abrahams BS *et al.* **Novel vertebrate genes and putative regulatory elements identified at kidney disease and NR2E1/fierce loci.** *Genomics, 80(1):45-53.* (PMID 12079282)
2. Pignoni F *et al.* **The Drosophila gene tailless is expressed at the embryonic termini and is a member of the steroid receptor superfamily.** *Cell, 62(1):151-63.* (PMID 2364433)
3. Kitambi SS and Hauptmann G. **The zebrafish orphan nuclear receptor genes nr2e1 and nr2e3 are expressed in developing eye and forebrain.** *Gene Expr. Patterns, 7(4):521-8.* (PMID 17127102)
4. Young KA *et al.* **Fierce: a new mouse deletion of Nr2e1; violent**
10. Phillips HS *et al.* **Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis.** *Cancer Cell, 9(3):157-73.* (PMID 16530701)
11. Liu HK *et al.* **The nuclear receptor tailless induces long-term neural stem cell expansion and brain tumor initiation.** *Genes Dev., 24(7):683-95.* (PMID 20360385)
12. Haider NB *et al.* **Mutation of a nuclear receptor gene, NR2E3, causes enhanced S cone syndrome, a disorder of retinal cell fate.** *Nat. Genet., 24(2):127-31.* (PMID 10655056)
13. Coppieters F *et al.* **Recurrent mutation in the first zinc finger of**

# Reader Interface -2

schizophrenia, and aggression through resequencing. *Am. J. Med. Genet. B Neuropsychiatr. Genet., 147B(6):880–9.* (PMID 18205168)

8. Parsons DW *et al.* **An integrated genomic analysis of human glioblastoma multiforme.** *Science, 321(5897):1807–12.* (PMID 18772396)
9. Sim FJ *et al.* **Neurocytoma is a tumor of adult neuronal progenitor cells.** *J. Neurosci., 26(48):12544–55.* (PMID 17135416)

*Biophys. Res. Commun., 386(4):671–5.* (PMID 19555662)

17. Qu Q *et al.* **Orphan nuclear receptor TLX activates Wnt/beta-catenin signalling to stimulate neural stem cell proliferation and self-renewal.** *Nat. Cell Biol., 12(1):31–40; sup pp 1–9.* (PMID 20010817)
18. Elmi M *et al.* **TLX activates MASH1 for induction of neuronal lineage commitment of adult hippocampal neuroprogenitors.** *Molecular and cellular neurosciences* (PMID 20599619)

## Figures

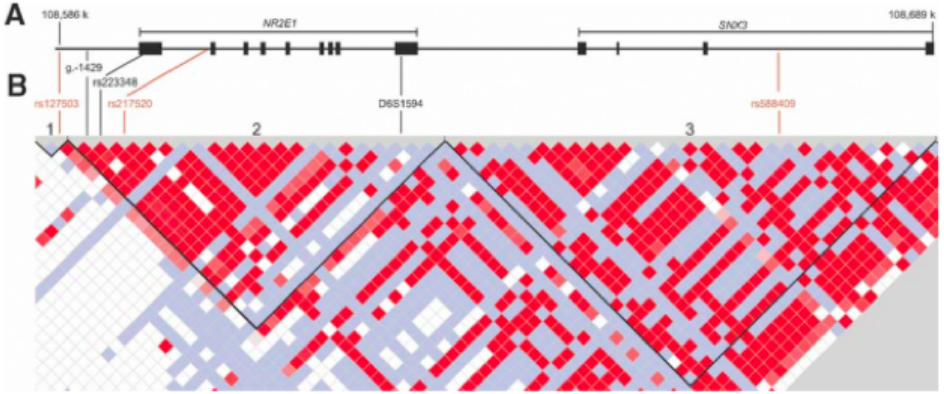



**FIGURE 1** Linkage disequilibrium map with six markers chosen for association analysis in bipolar disorder and schizophrenia

Genomic structure of *NR2E1* and location of six markers selected for association analyses in bipolar disorder and schizophrenia.
A: Schematic of *NR2E1* and closest neighboring gene *SNX3*.
B: Linkage disequilibrium (LD) map generated from the HAPMap data CEU population set. LD blocks (1, 2, and 3) were generated using the "Solid spine of LD" method. Markers that are tag SNPs are indicated in red. (Kumar *et al.* 2008)

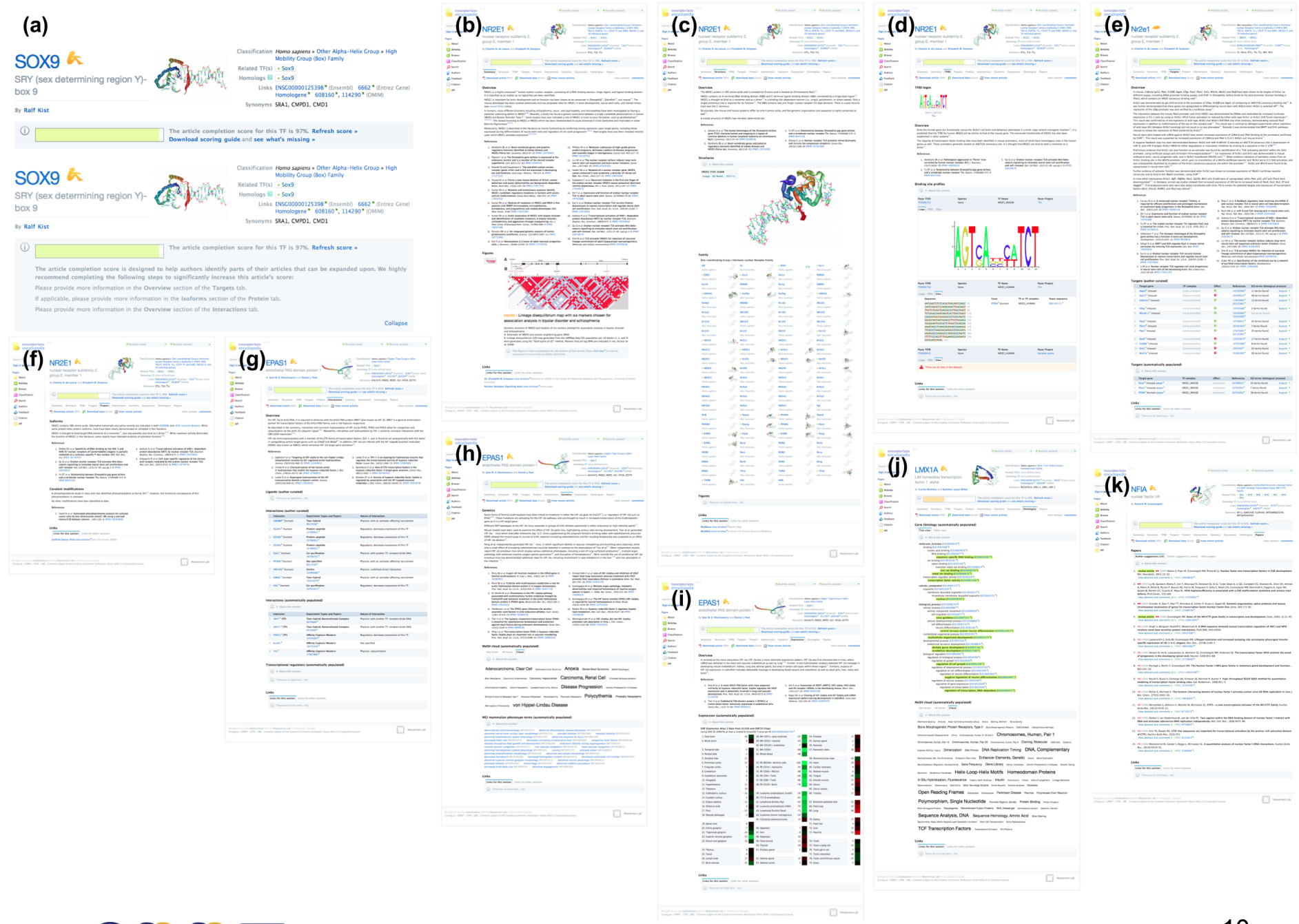


## Links

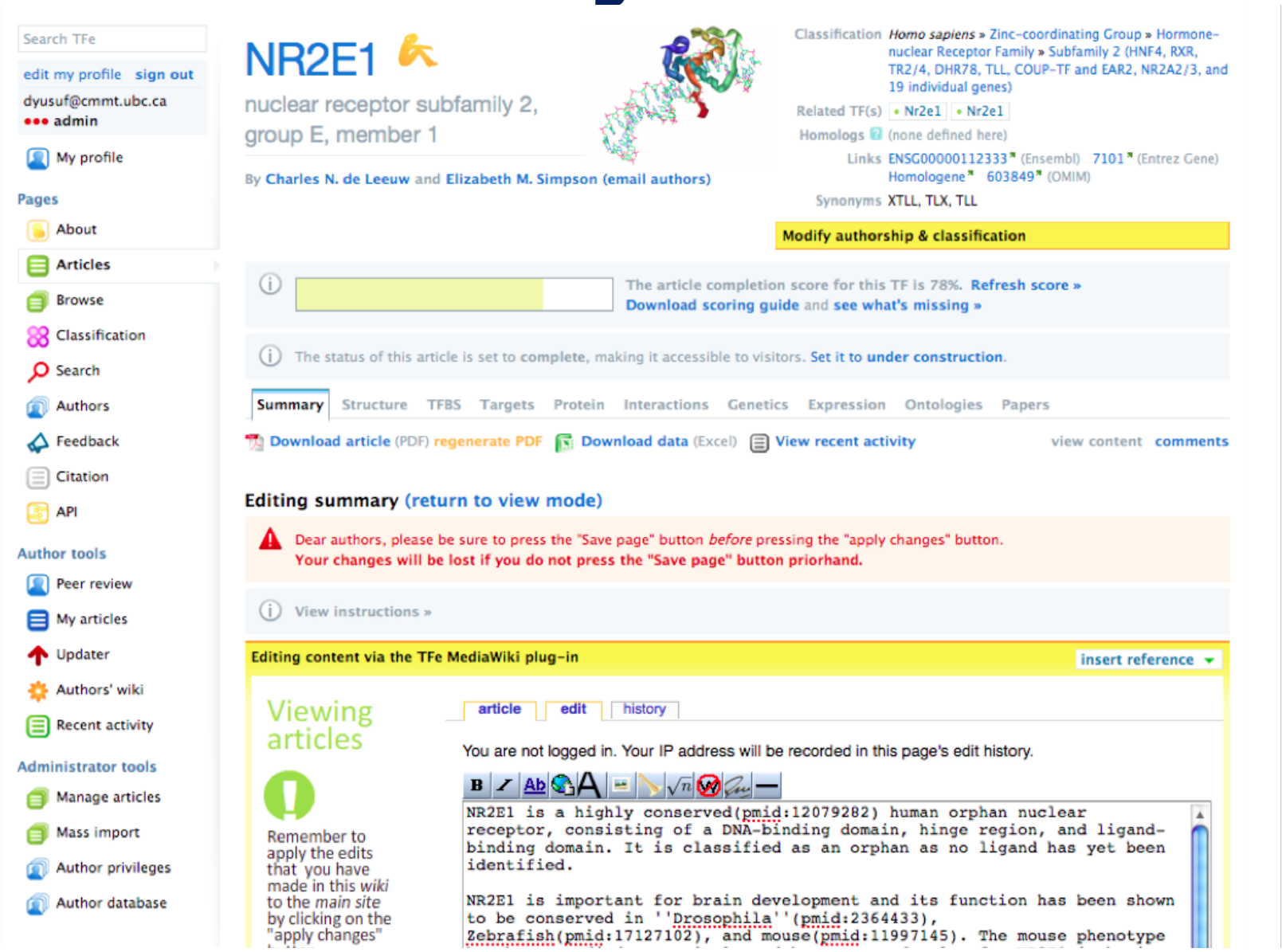
Links for this section | Links for other sections

Dr. Elizabeth M. Simpson (new window) Research on NR2E1 at the Centre for Molecular Medicine and Therapeutics, University of British Columbia

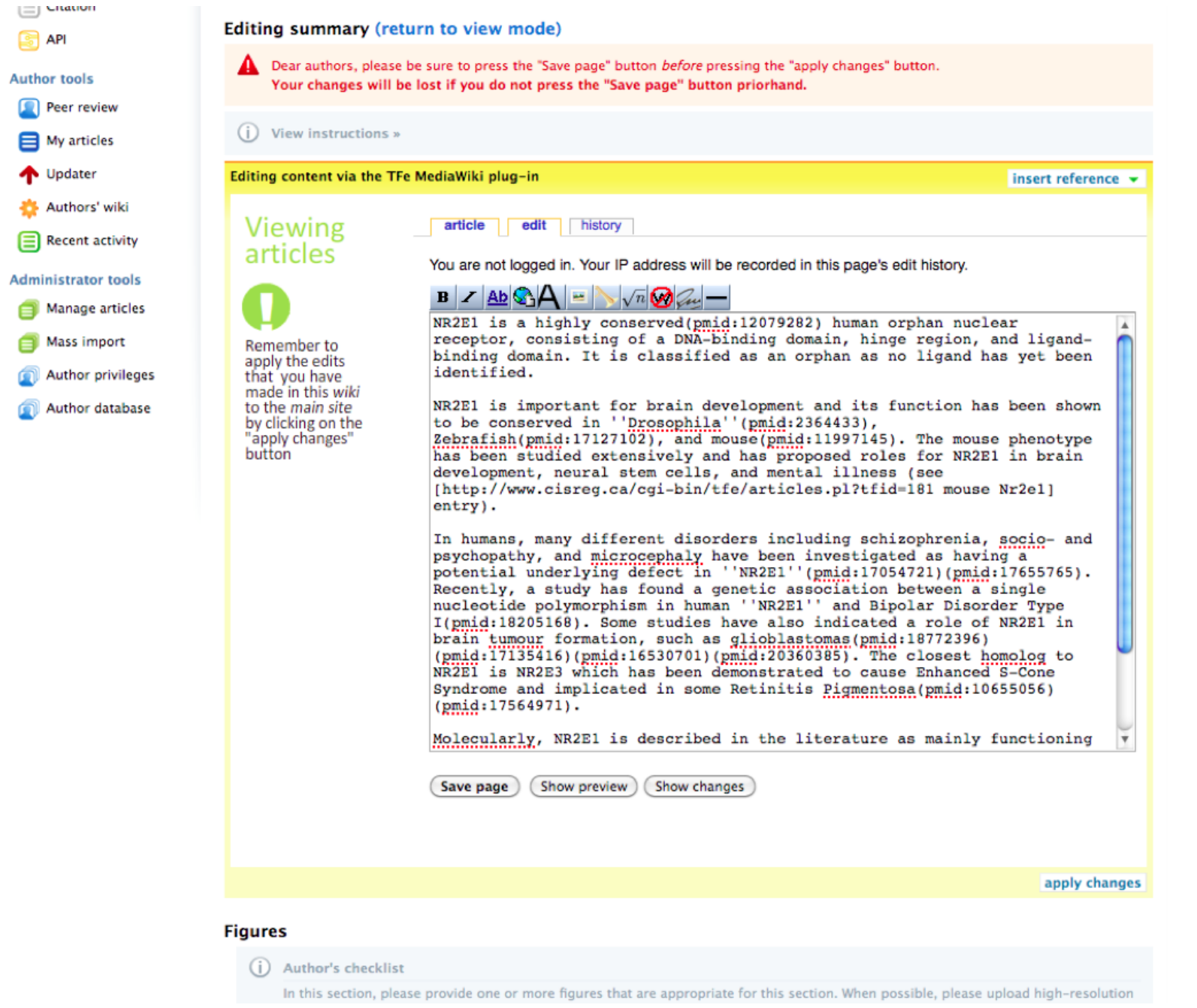
Nuclear Receptor Signaling Atlas (new window) Home page


(a)
SOX9
SRY (sex determining region Y)-box 9
By Ralf Kist
Classification Homo sapiens » Other Alpha-Helix Group » High Mobility Group (Box) Family
Related TF(s) Sox9
Homologs Sox9
Links ENSG00000125398 (Ensembl) 6662 (Entrez Gene) 608160, 114290 (OMIM)
Synonyms SRA1, CMPD1, CMD1
The article completion score for this TF is 97%. Refresh score »
Download scoring guide and see what's missing »
SOX9
SRY (sex determining region Y)-box 9
By Ralf Kist
Classification Homo sapiens » Other Alpha-Helix Group » High Mobility Group (Box) Family
Related TF(s) Sox9
Homologs Sox9
Links ENSG00000125398 (Ensembl) 6662 (Entrez Gene) 608160, 114290 (OMIM)
Synonyms SRA1, CMPD1, CMD1
The article completion score for this TF is 97%. Refresh score »
The article completion score is designed to help authors identify parts of their articles that can be expanded upon. We highly recommend completing the following steps to significantly increase this article's score:
Please provide more information in the Overview section of the Targets tab.
If applicable, please provide more information in the Isoforms section of the Protein tab.
Please provide more information in the Overview section of the Interactions tab.
Collapse
(b)
NR2E1
(c)
NR2E1
(d)
NR2E1
(e)
Nr2e1
(f)
NR2E1
(g)
EPAS1
(h)
EPAS1
(i)
EPAS1
(j)
LMX1A
(k)
NFIA

# Authoring Interface



# Authoring Interface - 2



# Authoring Interface - 3

**Figures**

Author's checklist

In this section, please provide one or more figures that are appropriate for this section. When possible, please upload high-resolution images that are at least 2,400 pixels across. **For the figure caption, please limit your description to 100 words per figure.** Feel free to add references to your figure caption by using the syntax: (pmid:123456) at the location where you want the reference to appear. In this case, "123456" refers to the PMC ID of your desired reference. Please ensure that you have copyright or permission to reproduce the figure here and for later publication in an open-access, peer-reviewed journal.

View figures | Upload a new figure

up • down

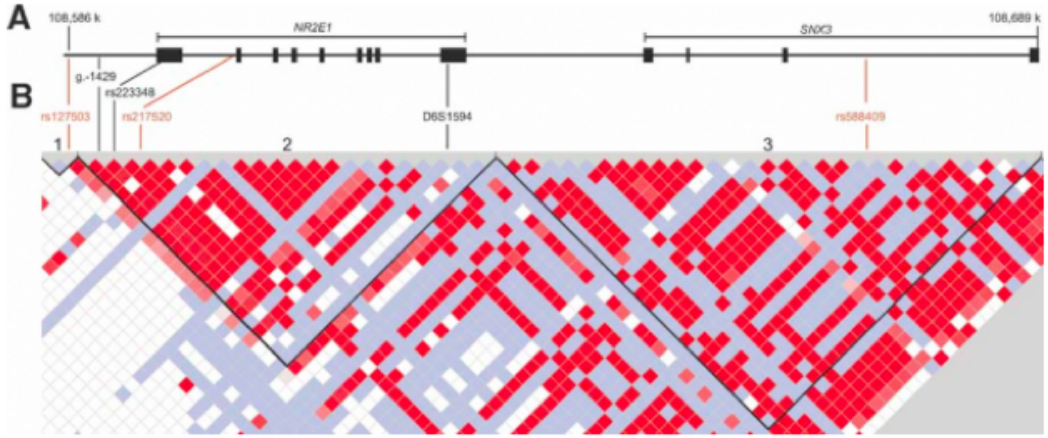



**FIGURE 1** Linkage disequilibrium map with six markers chosen for association analysis in bipolar disorder and schizophrenia

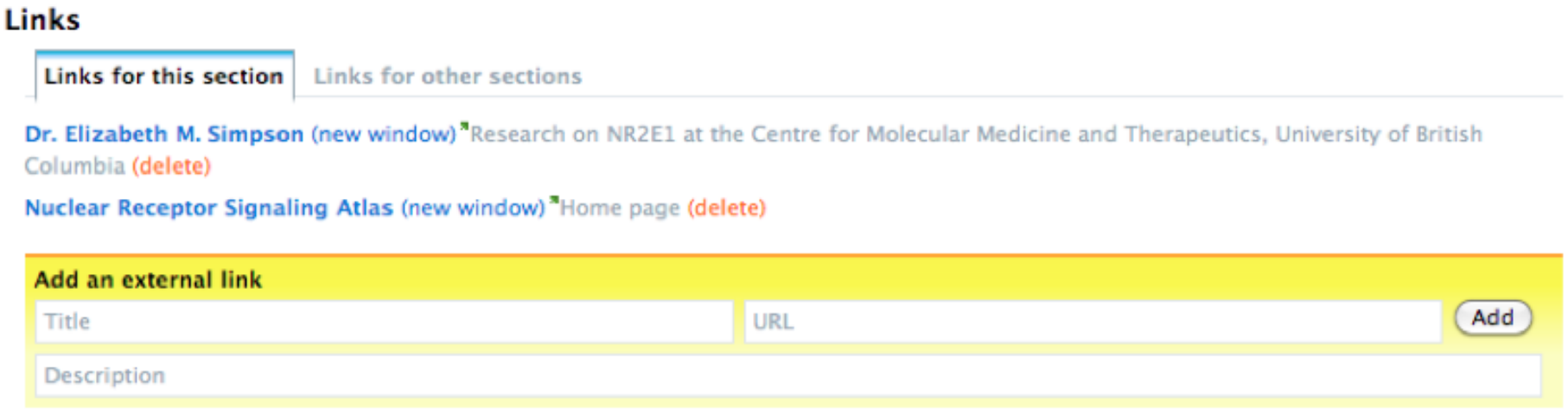
Genomic structure of *NR2E1* and location of six markers selected for association analyses in bipolar disorder and schizophrenia.
A: Schematic of *NR2E1* and closest neighboring gene *SNX3*.
B: Linkage disequilibrium (LD) map generated from the HAPMap data CEU population set. LD blocks (1, 2, and 3) were generated using the "Solid spine of LD" method. Markers that are tag SNPs are indicated in red. (Kumar *et al.* 2008)

This figure is from a publication by the authors of this article. Please click here to view the publication's entry in Pubmed (PMID 18205168).

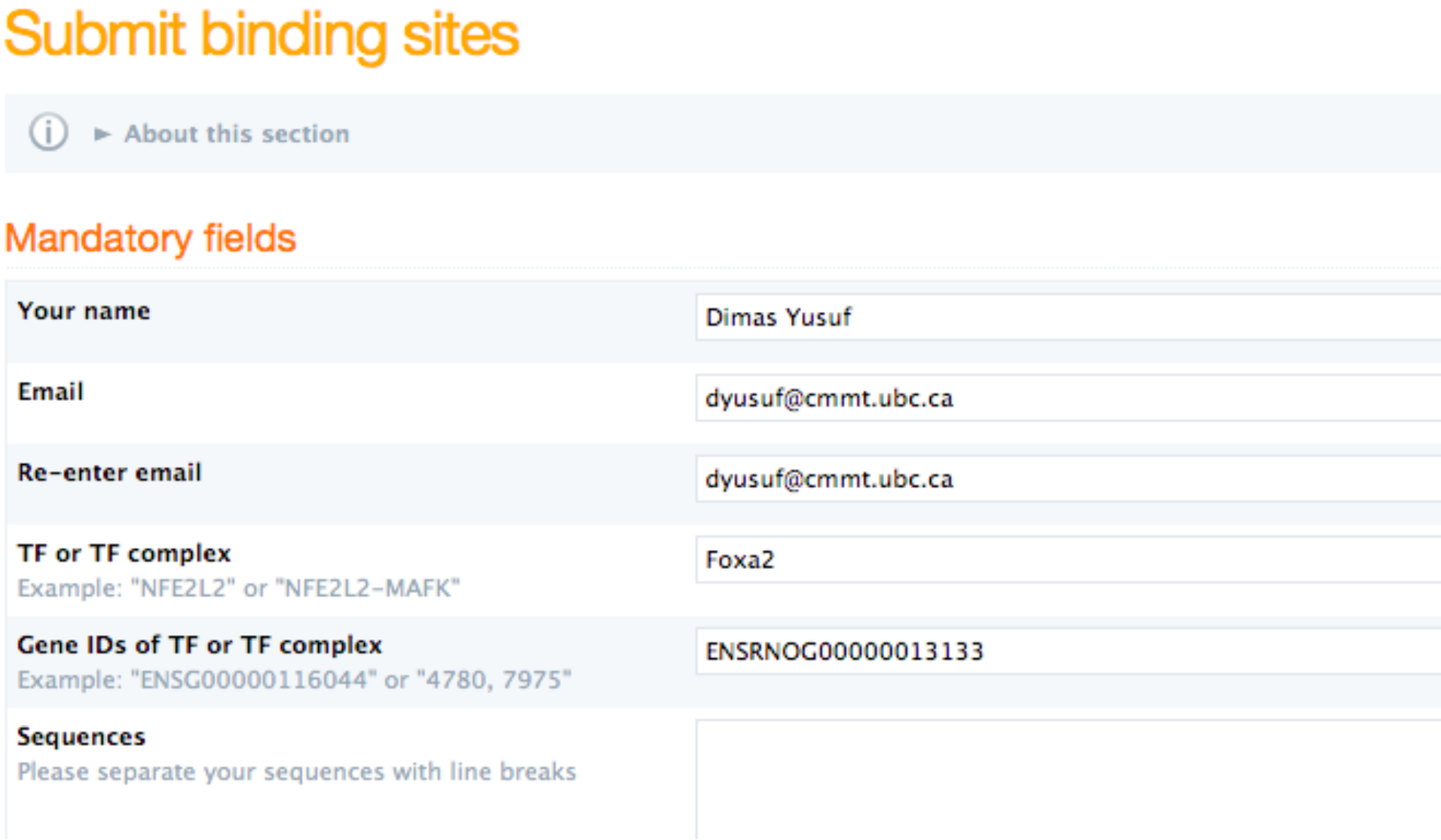
modify annotation | modify image | delete figure

**Links**

# Authoring Interface - 4

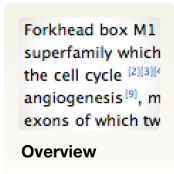


# TFBS Data Submission



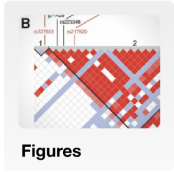
# Data within TFe

## Summary



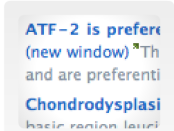

Text provided by the author with citations, ideally 500 words

WEB PDF XLS API




Diagrams provided by the author with caption text and citations

WEB PDF XLS API




Web links to relevant resources provided by the author

This section is present in all tabs

WEB PDF XLS API

## Structure




Text provided by the author with citations, ideally 200 words

WEB PDF XLS API



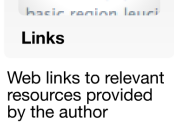

Prediction of DBD structure as an image and downloadable PDB file

**image**

WEB PDF XLS API

**PDB file**

WEB PDF XLS API




Tile of other TFs that belong to the same family based on DNA binding domain characteristics

WEB PDF XLS API




Diagrams provided by the author with caption text and citations

WEB PDF XLS API

## TFBS




Tile of mini logos of TFBS models available in TFe

WEB PDF XLS API




Text provided by the author with citations, ideally 150 words

WEB PDF XLS API




The logos, position frequency matrix (PFM), and sequence table of TFBS models

**logos**

WEB PDF XLS API

**PFM**

WEB PDF XLS API

**sequence table**

WEB PDF XLS API

## Targets




Text provided by the author with citations, ideally 200 words

WEB PDF XLS API




Genomic targets provided by the author

WEB PDF XLS API




Genomic targets provided by PAZAR

WEB PDF XLS API

## Protein




Text provided by the author with citations, ideally 200 words

WEB PDF XLS API




Text provided by the author with citations, ideally 200 words

WEB PDF XLS API

## Interactions




Text provided by the author with citations, ideally 200 words

WEB PDF XLS API




Interacting ligands (*i.e.* messengers) from the author

WEB PDF XLS API




Protein interactors provided by the author

WEB PDF XLS API


Protein interactors provided by BioGRID with additional author annotation

WEB PDF XLS API




## Genetics




Text provided by the author with citations, ideally 250 words

WEB PDF XLS API




TF-to-disease links based on NCBI data, from MeSHOP

WEB PDF XLS API




Associated mouse phenotype terms from MGI

WEB PDF XLS API

## Expression




Text provided by the author with citations, ideally 200 words

WEB PDF XLS API




Microarray data from the GNF Expression Atlas 2, as provided by the UCSC Genome Browser

WEB PDF XLS API




A link to the mouse brain expression data from Allen Brain Atlas

WEB PDF XLS API

## Ontologies



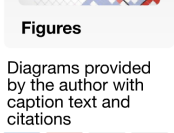

Associated GO terms from Entrez Gene and Gene Ontology in tree and table view

**tree view**

WEB PDF XLS API

**table view**

WEB PDF XLS API




TF-to-MeSH term associations based on NCBI data, from MeSHOP

WEB PDF XLS API

## Papers




A list of relevant papers provided by the author

WEB PDF XLS API

## Other information




TF classification information (group,



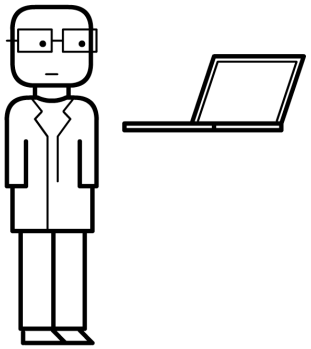

Links to external resources such as




Authors' names, contact information,

# URL-based Data Extraction

**1** Bioinformatician writes a small "script" program to automatically retrieve the desired information from the TFe Web API



Get all **TFe TFIDs**
For each **TFe TFID**,
  Get the TFID's **symbol**
  If symbol equals "ATF3",
    Get **MeSH disease term**

**2** The program is set to run on its own

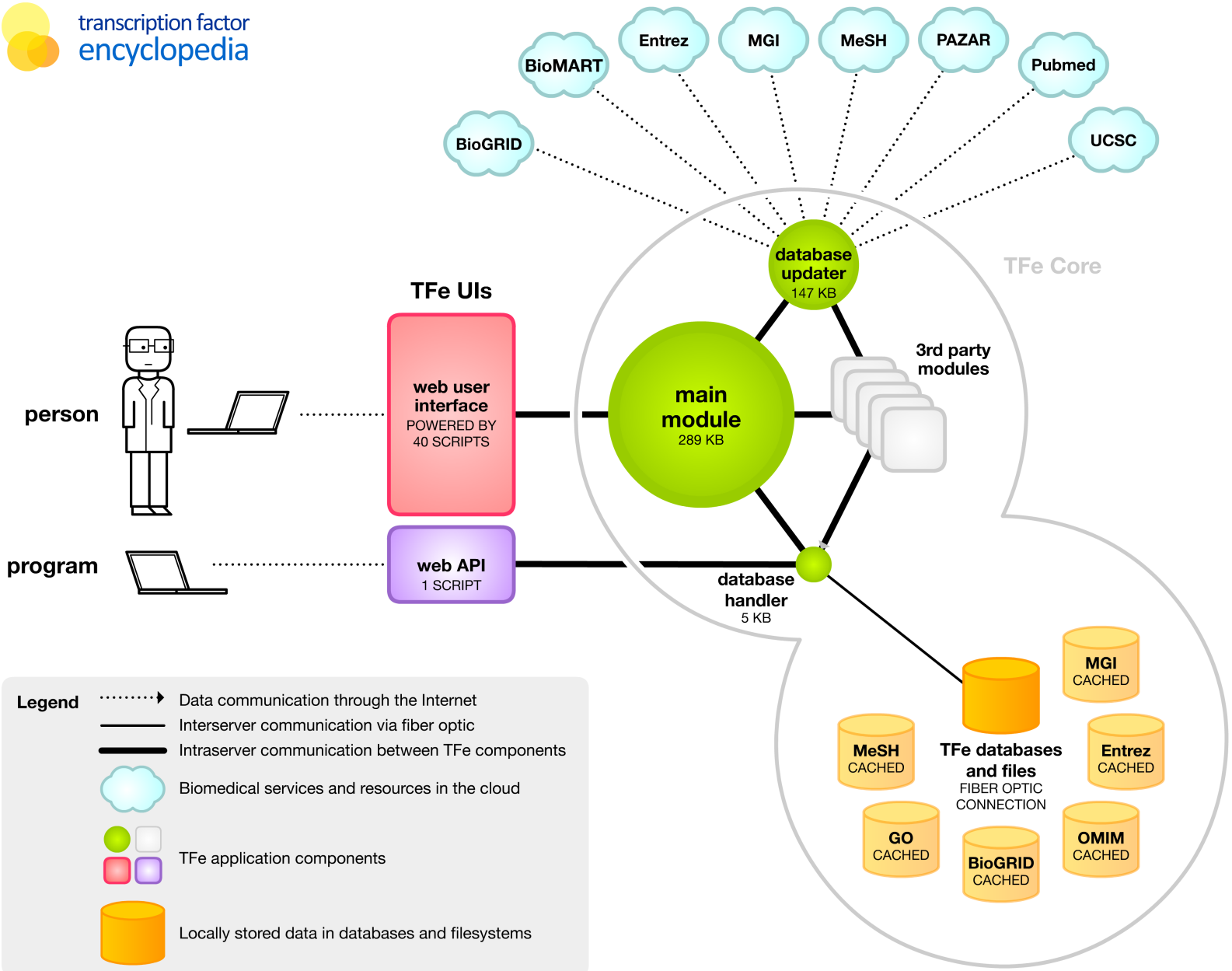

Get all **TFe TFIDs**
http://www.cisreg.ca/cgi-bin/tfe/api.pl?code=all-tfids

118
119
120
...

**Web API**

# System Map

transcription factor encyclopedia



# Author Recruitment

- **Stage 1 (alpha)**
  - Friends
  - Friends of Friends
  - ~10
- **Stage 2 (beta)**
  - PubMed mining to identify experts (> 10 published articles about a TF) with email solicitation
  - ~100

# TFe Authors


+1 Australia

~15% of contacted experts agreed to participate

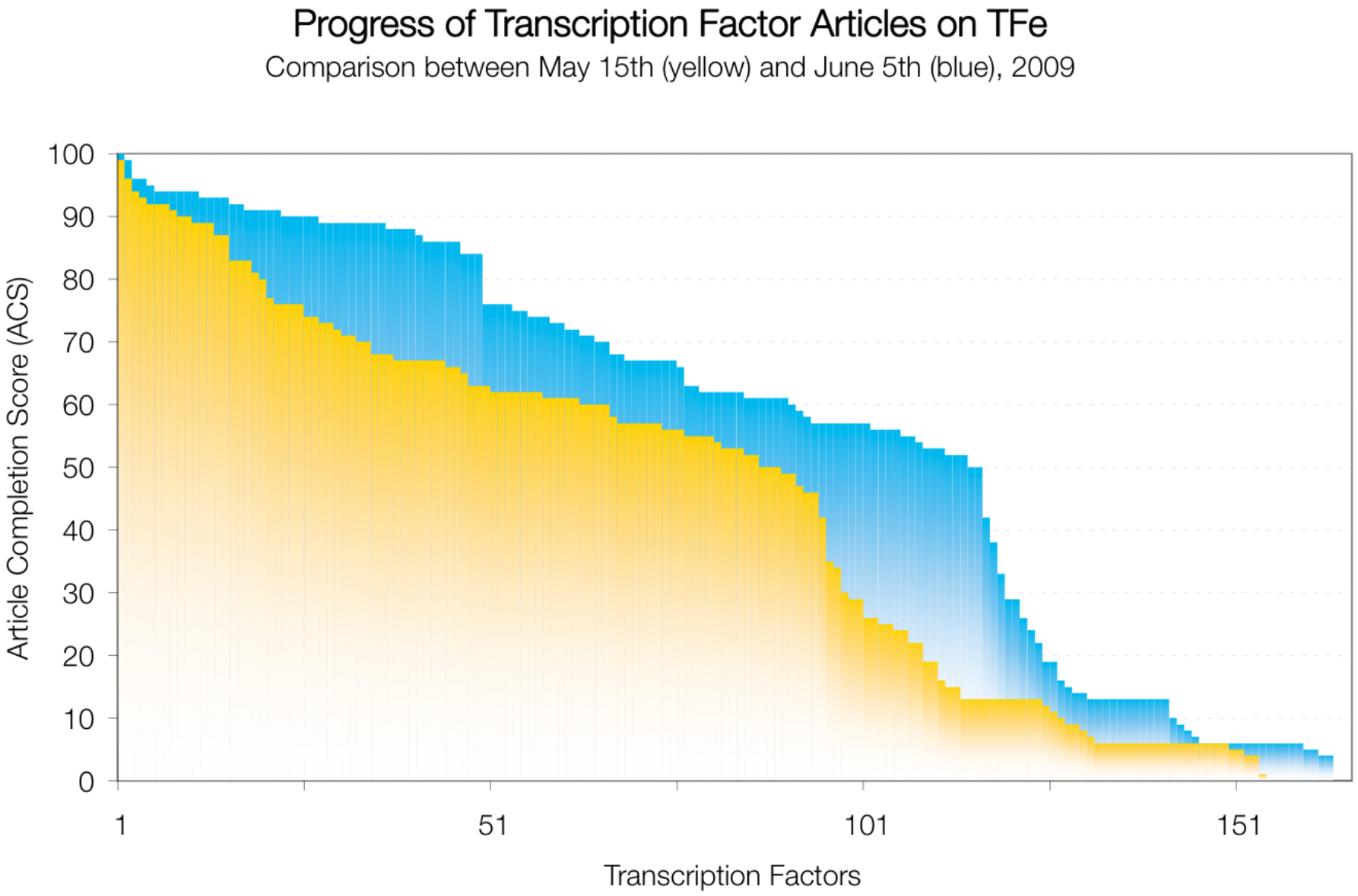
~60% of participants completed an entry

# Author Motivation

- **Sponsoring Journal**
  - Scientists more willing to commit time and effort to projects that give authorships
- **Progress Scores**
  - Many people are competitive and like to see their work achieve the best possible grade
- **Peer Review**
  - Respect of peers and fixing weaknesses

# TFe Article Scoring Procedure

| Tab | Scoring element | Target | Maximum points | Weight |
|---|---|---|---|---|
| Summary | Overview text | 500 words | 10 points | 8.333% |
| Summary | References in overview text | 3 references | 5 points | 4.167% |
| Summary | Figures | 1 figure | 10 points | 8.333% |
| Structure | Overview text | 200 words | 5 points | 4.167% |
| TFBS | Overview text | 150 words | 5 points | 4.167% |
| TFBS | Binding site profiles | 1 binding site profile | 10 points | 8.333% |
| Targets | Overview text | 200 words | 5 points | 4.167% |
| Targets | Targets | 10 targets in total (both author and auto) | 10 points | 8.333% |
| Protein | Isoforms text | 200 words | 5 points | 4.167% |
| Protein | Covalent modifications text | 200 words | 5 points | 4.167% |
| Interactions | Overview text | 200 words | 5 points | 4.167% |
| Interactions | Ligands | 1 ligand | 1 point | 0.833% |
| Interactions | Interactions | 10 interactors in total (both author and auto) | 10 points | 8.333% |
| Interactions | Interactions | All “nature of interaction” fields annotated | 10 points | 8.333% |
| Genetics | Overview text | 250 words | 5 points | 4.167% |
| Expression | Overview text | 200 words | 5 points | 4.167% |
| Papers | Papers | 15 papers | 10 points | 8.333% |
| Papers | Papers | 3 papers marked as “recommended” | 3 points | 2.500% |
| (all) | Links | 1 link | 1 point | 0.833% |
| | | | **120 points** | **100%** |


Progress of Transcription Factor Articles on TFe
Comparison between May 15th (yellow) and June 5th (blue), 2009
Article Completion Score (ACS)
100
90
80
70
60
50
40
30
20
10
0
1
51
101
151
Transcription Factors

**PDFs**



June 29th, 2010

http://www.cisreg.ca/cgi-bin/tfe/articles.pl?tfid=444

# FOXO4

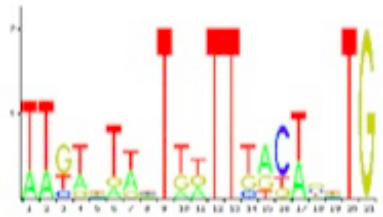
*Homo sapiens* forkhead box O4

**By Astrid Eijkelenboom[1] & Marrit Putker[2*]**

FOXO4 (also known as AFX) is a member of the Forkhead familiy of transcription factors[1] and forms a subclass with FOXO1 (FKHR), FOXO3A (FKHRL1) and FOXO6. The FOXO transcription factors are key players in regulation of cell-fate decisions (cell death, cell proliferation and cell metabolism, see FIGURE 1) and are considered to be tumor suppressors[2]. In model organisms FOXO was shown to increase longevity[3] and this function was shown to be dependent on FOXO's ability to induce oxidative stress inducing genes[4]. The FOXO subclass members 1, 3a and 4 are ubiquitously expressed, but their respective levels differ per cell type or organ[5], whereas FOXO6 expression seems to be restricted to the brain. All FOXOs consist of a forkhead DNA binding domain, nuclear localization signal (NLS), nuclear export sequence (NES) and transactivation domain (TA). The DNA binding domain is highly conserved within the FOXO family and is shown to bind the core consensus DNA sequence 5` TTGTTTAC 3`[5]. Overlap in target genes is thus expected and has been shown. Indeed studies in FOXO1, FOXO3a and FOXO4 knockout mice show that FOXO tumor suppressive function is intact in mice lacking any combination of two, but not all three, FOXO genes[6] . Any functional specificity in function is likely to be obtained through posttranslational modifications and interaction with specific co-factors. The activity of FOXO family members FOXO1, FOXO3 and FOXO4 is regulated by cellular localization and several posttranslational modifications like phosphorylation, acetylation and ubiquitination (see FIGURE 2). Some of these modifications induce a change in subcellular localization of FOXO. For instance, FOXO activity is negatively regulated by PI3-K via PKB/Akt in response to insulin through phosphorylation, resulting in translocation of the transcription factors from nucleus to cytoplasm[7,8,9,10], whereas stress induced kinases like JNK positively regulate FOXO nuclear localization and transcriptional activity[11,12,13](see FIGURE 3 for a model of the regulation of FOXO transcriptional activity by posttranslational modifications). For more information and reviews, we would like to refer to an issue of Oncogene Reviews, completely dedicated to the FOXO family of forkhead transcription factors (volume 27, number 16 ? April 7, 2008).

## Binding sites

Using in vitro experiments, the DBD of the FOXO proteins has been shown to bind the consensus sequence 5` TTGTTTAC 3`[5,16]. Although overlap in target genes between the different members of the FOXO family has been shown, specificity in function is likely to be obtained through posttranslational modifications and interaction with specific co-factors. The TFBS below is based upon data obtained for FOXO4.

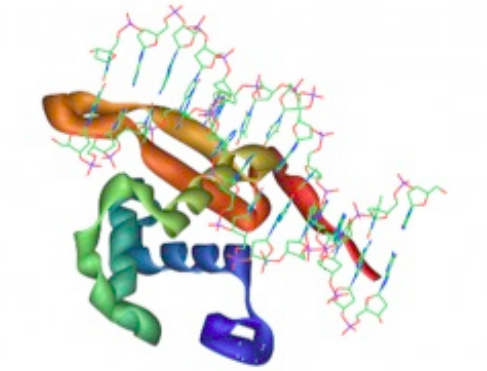


**Binding profile from Pazar**

| | |
|---|---|
| Project name | TFe |
| TF name | FOXO4_HUMAN |
| TF species | *None* |
| Pazar ID | TF0000771 |
| Ensembl ID | ENST00000399704 |

This data is sourced from Pazar, a public database of transcription factor and regulatory sequence annotation. http://www.pazar.info/

| PFM | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 2 | 2 | 0 | 2 | 2 | 1 | 3 | 2 | 0 | 1 | 2 | 0 | 0 | 0 | 4 | 0 | 3 | 1 | 2 | 0 | 0 |
| C | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 5 | 0 | 2 | 1 | 0 | 0 |
| G | 0 | 0 | 4 | 1 | 2 | 1 | 1 | 1 | 0 | 2 | 2 | 0 | 0 | 2 | 2 | 1 | 0 | 3 | 2 | 0 | 7 |
| T | 5 | 5 | 2 | 4 | 2 | 5 | 3 | 2 | 7 | 4 | 3 | 7 | 7 | 4 | 1 | 1 | 4 | 1 | 2 | 7 | 0 |



**Protein structure of FOXO4**

FOXO proteins consist of four domains: a highly conserved DNA binding domain (DBD), a nuclear localization signal (NLS), a nuclear export sequence (NES) and a C-terminal transactivation domain. The structure of FOXO4 DBD with DNA has been resolved and shows high similarity with other forkhead DBDs[14]. We would like to refer to a recent review for more information on structural studies of forkhead transcription factors and the potential effects of posttranslational modifications on DNA binding[15].

**Classification**

| | |
|---|---|
| Group | Winged Helix-Turn-Helix |
| Family | Forkhead Domain Family |
| Subfamily | Not specified |

**Resources**

| | |
|---|---|
| Entrez Gene | 4303 |
| Ensembl | ENSG00000184481, |
| Refseq | NP_005929, NP_001164402 |
| Uniprot | P98177 |
| OMIM | 300033 |
| Synonyms | MLLT7, AFX1, MGC120490, AFX |

**About**



**Contact**

Affiliation(s): (1) The Netherlands. (2) University Medical Centre Utrecht, Utrecht, Utrecht, The Netherlands. (*) Correspondence: Marrit Putker. Email: m.putker@umcutrecht.nl

Displaying the first 3 of 4 figures. See more on site »

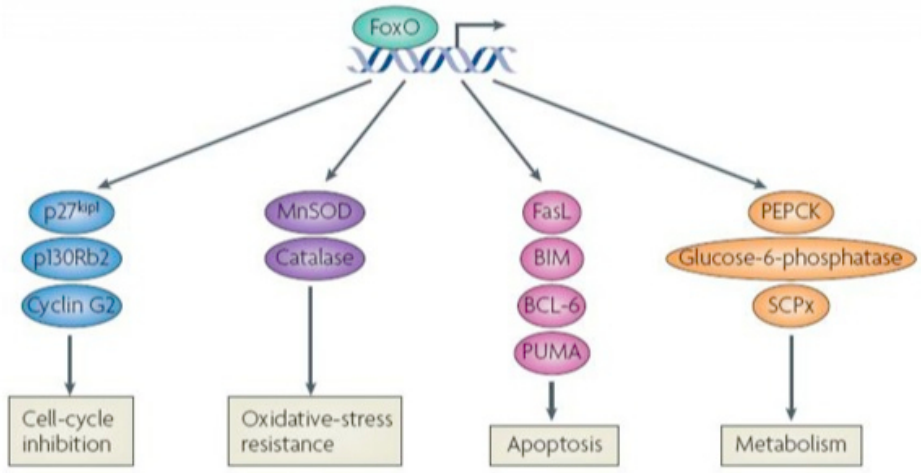



FIGURE 1 | **Figure 1: Transcriptional outputs of FoxO activity.** Increased class O forkhead box transcription factor (FoxO) activity participates in several cellular processes, most notably inhibition of the cell cycle, regulation of cell death, protection from cellular (oxidative) stress and regulation of cellular metabolism (gluconeogenesis and fatty-acid oxidation). FoxO-regulated genes that are linked to these processes are indicated. The outcome of FoxO function is likely to be determined in conjunction with other genetic determinants, for example, the function of the tumour-suppressor protein p53. MnSOD, manganese superoxide dismutase; PEPCK, phosphenolpyruvate carboxykinase; SCPx, sterol carrier protein-x.

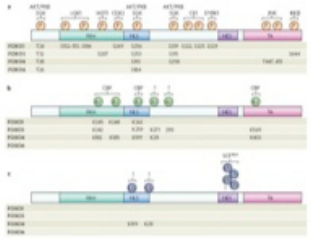


FIGURE 2 | **Figure 2: Summary of post-translational modifications on the various FoxO isoforms.** The class O forkhead box transcription factor (FoxO) isoforms contain a forkhead domain (FKH, the DNA-binding domain), a nuclear localization signal (NLS), a nuclear export sequence (NES) and a transactivation domain (TA). Modified residues are indicated and amino-acid numbers are given for the individual isoforms. Enzymes responsible for these modifications are indicated above the modification. However, when a residue is conserved but no positive evidence for the modification has been obtained in other isoforms, the residue is not specified for those isoforms.

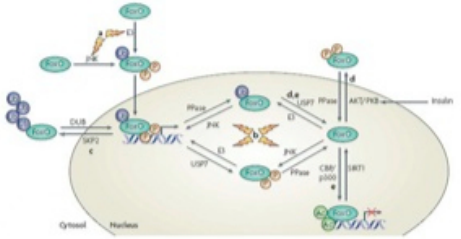


FIGURE 3 | **Figure 3: An integrated model for FoxO activity control through stress-induced post-translational modifications.** Following increased cellular oxidative stress (for example, by hydrogen peroxide treatment of cells) class O forkhead box transcription factors (FoxOs) translocate from the cytosol to the nucleus. This process correlates with c-Jun N-terminal kinase (JNK)-mediated phosphorylation (residues Thr447 and Thr451 in FOXO4) and monoubiquitylation of FoxO. The cellular location at which these modifications occur is not known but two possibilities are indicated: in the cytosol (a) or in the nucleus (b). Monoubiquitylation might occur in either location but the identity of the E3 ligase is unknown. *(continued on site)*

**Isoforms**

Two isoforms of FOXO4 (AFX) have been described in non-cancer cells: AFX? (regular) and AFX? (lacking aminoacid 58-112 and thus the first 16 aminoacids of the forkhead domain)). The AFX? splice variant shows a somewhat different transcriptional activity compared to the regular form of AFX. It is even proposed that the two isoforms antagonize each other[20,21]. In cancer cells, three other isoforms of FOXO4 have been described. Two of them are produced by aberrant splicing: AFXtr1 and AFXtr2 are short N-terminal FOXO4 proteins of respectively 90 and 101 amino acids long. However, no protein expression could be detected of these splice variants. The third isoform identified is translated from a downstream start site and is thus a N-terminally deleted isoform: AFX?(?198-505). This isoform of FOXO4 shows similar transcriptional activity as AFX?, not being able to induce apoptosis and interfering with AFX? function when both overexpressed[21].

**Covalent modifications**

FOXO transcription factors are regulated by several posttranslational modifications including phosphorylation, acetylation and ubiquitination. The precise mechanisms and the effects of these modifications are still not completely understood. Several modifications might interfere with DNA binding or affect binding with other proteins involved in the localization of the transcription factor. Upon insulin signaling, active PKB (AKT) phosphorylates FOXO4 on three sites: T28, S193 and S258[8]. Close relative of PKB SGK is able to phosphorylate FOXO on the same sites[22]. These phosphorylations lead to inhibition of FOXO4 activity due to reduced DNA binding and to binding to 14-3-3, leading to relocalization of FoxO4 to the cytoplasm[23,24,25]. In the cytoplasm, FOXO4 can be polyubiquitinated by Skp2, leading to proteosomal degradation[26]. Elevated cellular levels of ROS activate JNK, which in turn phosphorylates FOXO4 at at least two sites: T447, T451. Phosphorylation at these sites activates FOXO4 by inducing its nuclear localization[11]. FOXO4 is monoubiquitinated at several sites by MDM2, leading to nuclear localization and thus induction of transcritpional activity[27]. Deubiquitination occurs by DUB USP7/HAUSP[28]. P300 and CBP (Creb binding protein) regulate FOXO4 activity by acetylating it on several sites. Deacetylation occurs by HDAC and Sirt1[29]. *(continued on site)*

# Other Features

- **Predicted Binding Domain Structures (Phil Bradley, FHCRC)**
- **MeSH Over-representation Profiles – Attribute Clouds (Warren Cheung)**
- **TF Binding Site Profiles (Elodie Portales-Casamar)**
- **Parent-Child Relationships (inherit ortholog content if no species-specific version)**

# Sustainability

- **Unclear if it can be sustained – not proven**
- **Continuing journal sponsorships to motivate authors to update articles (or for new authors to take on abandoned articles)**
- **Focused TF Family-based papers (e.g. Nuclear Receptors)**

# Lessons Learned

- **Motivation for authors – Sponsor Journals**
- **Simple interfaces with extremely short learning periods to keep authors' attention**
- **Dedicated manager to engage authors, suggest changes and resolve issues**
- **Focus allows the system to be tailored to the needs of the field**

# Thanks!

**THE LAB**

- Dimas Yusuf
- Elodie Portales-Casamar
- Warren Cheung
- Stefanie Butland
- Magdalena Swanson

- Virginie Bernard
- Rebecca Hunt-Newbury
- Andrew Kwon
- Miroslav Hatas
- David Arenillas
- Jonathan Lim
- Dora Pak

•And many alumni

**COLLABORATORS**

- Phil Bradley and Amy Ticoll (FHCRC)
- Frances Sladek (UC – Riverside)

•~100 TFe Authors – The people most responsible for the success of the TFe

**THE APPRECIATED FUNDERS!**

- National Institutes of Health
- Genome Canada / Genome BC
- Canadian Institutes for Health Research
- Canada Foundation for Innovation

CIHR/MSFHR BIOINFORMATICS TRAINING PROGRAM FOR HEALTH RESEARCH

CAMT