Bioinformatics analysis of Zebrafish gene *npdc1*

By G07：张雨田 杨磊 郭前进 程振朝

Speaker：程振朝

10th Jan., 2014 @ PKU

Outline

1. Introduction to zebrafish

2. Introduction to Zebrafish gene *npdc1*

NPDC-1, a regulator of neural cell proliferation and differentiation, interacts with E2F-1, reduces its binding to DNA and modulates its transcriptional activity

--NPDC-1 is specifically expressed in neural cells when they stop to divide and begin to differentiate.

Research

Open Access

A neuronal-specific differentiation protein that directly modulates retinoid receptor transcriptional activation

Kenneth W Henry II, Michael L Spencer, Maria Theodosiou, Dingyuan Lou

NPDC-1 is transported in vesicles from the Golgi apparatus to the cell membrane and is then likely internalized into endosomes.
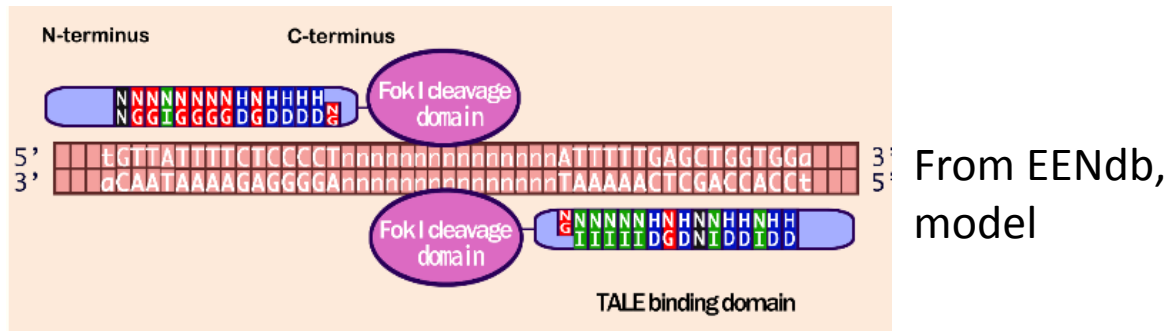
Introduction to Zebrafish gene *npdc1*

- Our research:
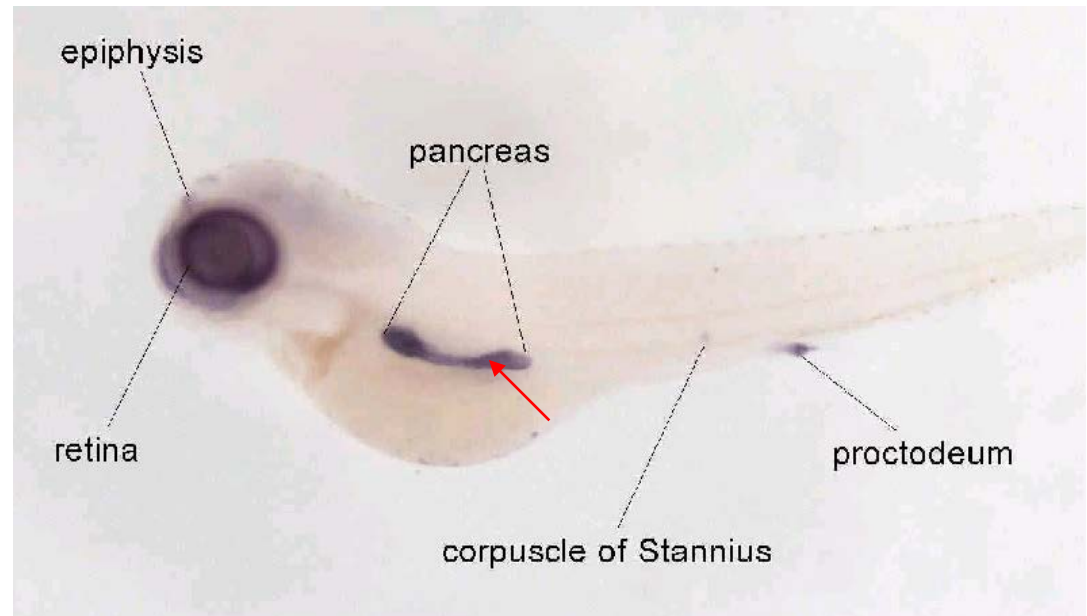
This in-situ hybridation figure (from ZFIN)showed that *npdc1* gene was largely expressed in pancreas on 5dfp, which inspired us to explore the role of npdc1 in zebrafish pancreas by gene targeting.

Gene targeting

- TALEN: DNA binding domain - **TALE** and DNA cleavage domain- **FokI** endonuclease

From EENdb, model

NHEJ (Non-homologous end joining)
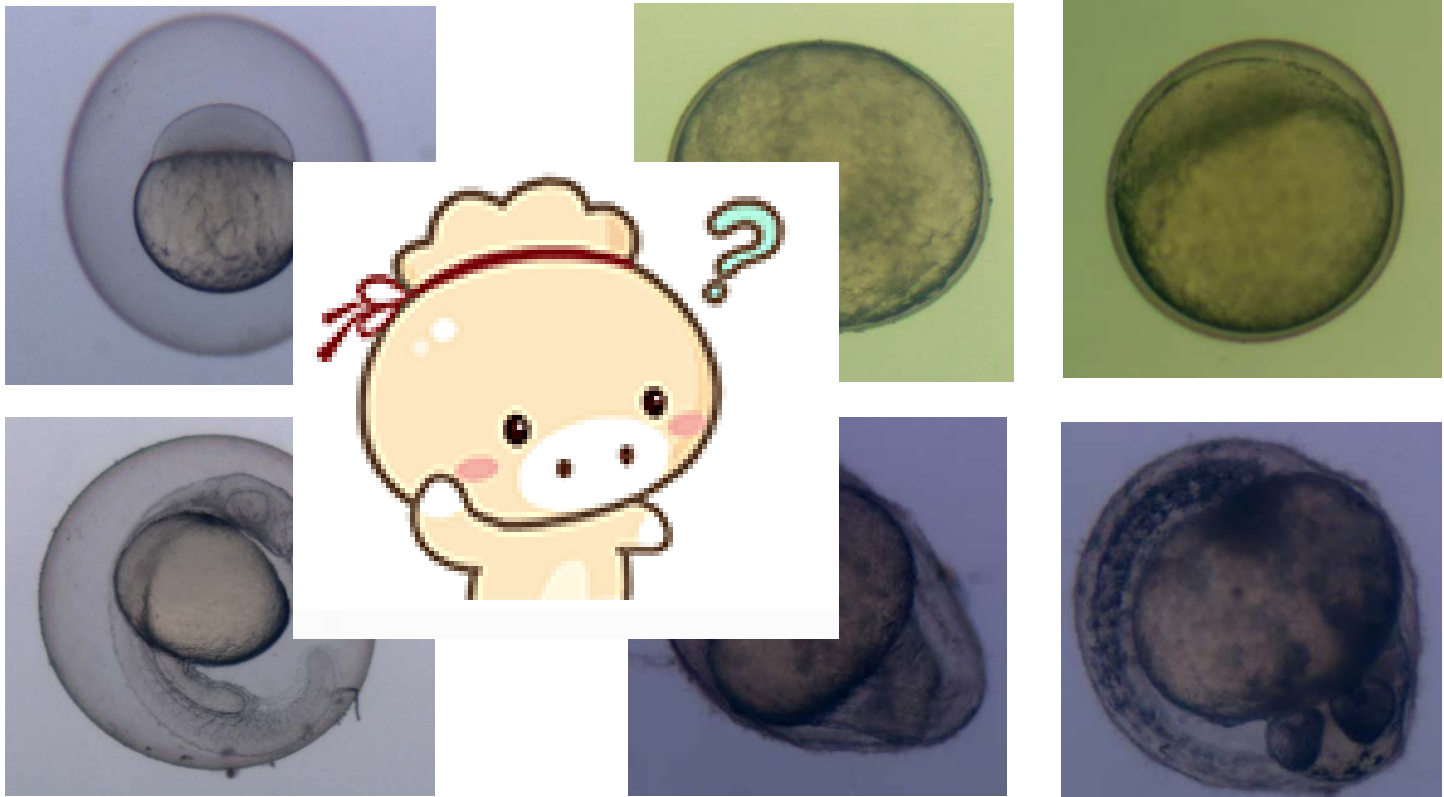
Indel (insertion/deletion)

breeding and screening

Homozygous mutation

Phenotype of zebrafish npdc1$^-$/npdc1$^-$

- Just finding that zebrafish egg cannot absorb water to expand its oolemma, compared with WT.
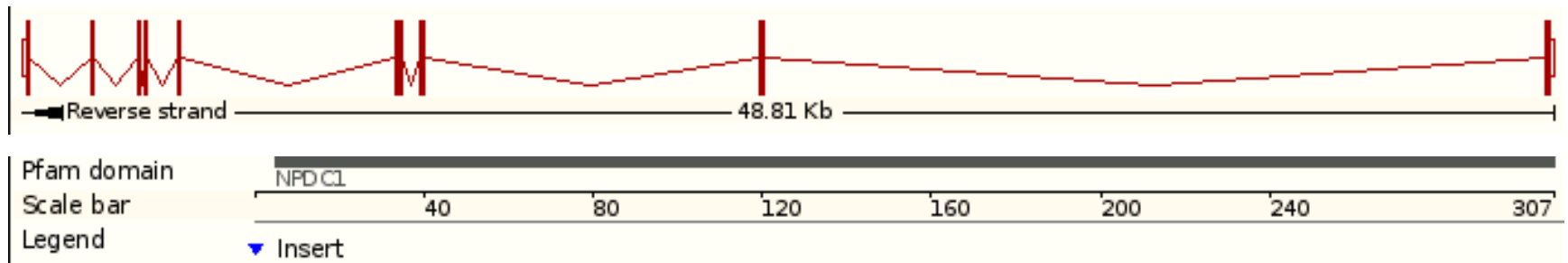

WT   npdc1$^-$/npdc1$^-$ mutants (F3)

3. Analysis of gene *npdc1*

- Location: 21: 11,904,663-11,953,632 -.

- npdc1$^-$/$^-$ zebrafish (exon2 -7bp)

Paralogous of *npdc1*

- An amazing NPDC1 gene in zebrafish

Zfin	Ensembl	NCBI
npdc1	npdc1	npdc1
	NPDC1	NPDC1-L

- DNA Sequence alignment (Ensembl)

Protein sequence similarity is 44%, but N-terminal 40 Aas are high conserved. NPDC1 may be a pseudogene.


Range 1: 1 to 277 Graphics   ▼Next Match ▲Prev

Score	Expect Method	Identities	Positives	Gaps
256 bits(654) 5e-88	Compositional matrix adjust.	141/282(50%)	186/282(65%)	23/282(8%)

```
Query   38   CPDHIECARRGRHFCKPGSSNCGPCLIPLEEDQMGHCVVRSVRYAAQSRDPVAPLSGIDE   97
             CP  ++CAR  RHFC+PGS +CGPCL P  E++ G CV+R    + A+    V+ L  +DE
Sbjct    1   CPRSLDCARERRHFCQPGSLHCGPCLDPFMENKRGKCVLRRRNHPAK----VSHLPELDE   56

Query   98   EIDFLSSVITKQQLSESKHSDSAPQS--PPQVQLKQKSLHKT-----TTVEPMT---TSP   147
             EID LSS+I+K + SE KHS  +P +    P+ +    S +K      T+++P T    S
Sbjct   57   EIDILSSIISKHRESEMKHSAPSPAASKAPENKSGSSSHYKAPPTAATSIQPPTRGLISS   116

Query  148   TAPTHTTN--SSNRHGPVISPNPSRD-SLLVLMISLCIIVGAMALILVTVCWVRLQRETR   204
             T+PT T  S+     P I P PS D S + + + ++VG++A++L  VCWVR+QR  R
Sbjct  117   TSPTSNTPFISAVHSAPFIIPYPSEDHSFITVFLGVFLMVGSVAMVLTGVCWVRMQRGCR   176

Query  205   LAQKVDYPAFQEAANRNNTS-----SGDKTLAHSAQMYHYQHQKQQMLSMEKHKAEPKVS   259
             LAQKVDYPAF    + S      SGDK LA SAQMYH+Q QKQQM+S+ K +++ K+
Sbjct  177   LAQKVDYPAFGLIGPNSYDSGMVSKSGDKKLAQSAQMYHFQLQKQQMMSL-KQRSDSKIP   235

Query  260   ESGGHSDEETEEGDFTVYECPGLAPTGEMEVKNPLFDDSTLH   301
             +SG   SDEE E+GDFTVYECPGLAPTGEMEVKNPLFDDST H  ◄───────
Sbjct  236   DSGATSDEENEDGDFTVYECPGLAPTGEMEVKNPLFDDSTFH   277
```
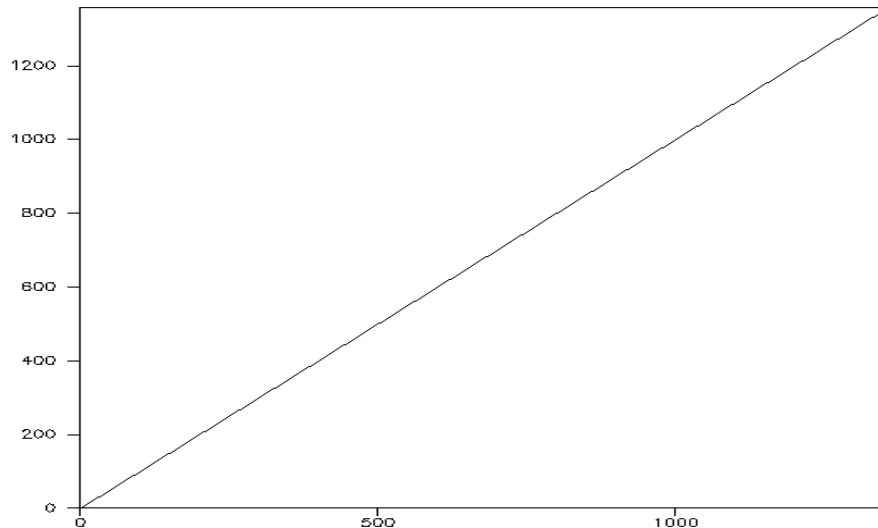
ORF and repeated sequence

- PlotORF: The result was consistent with truth.
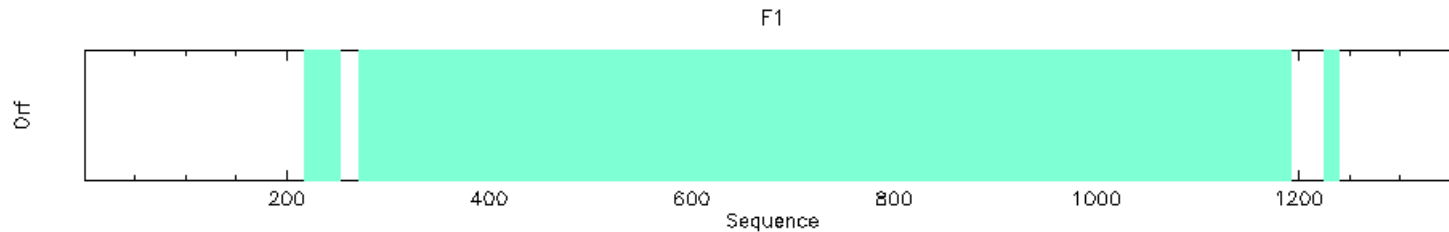
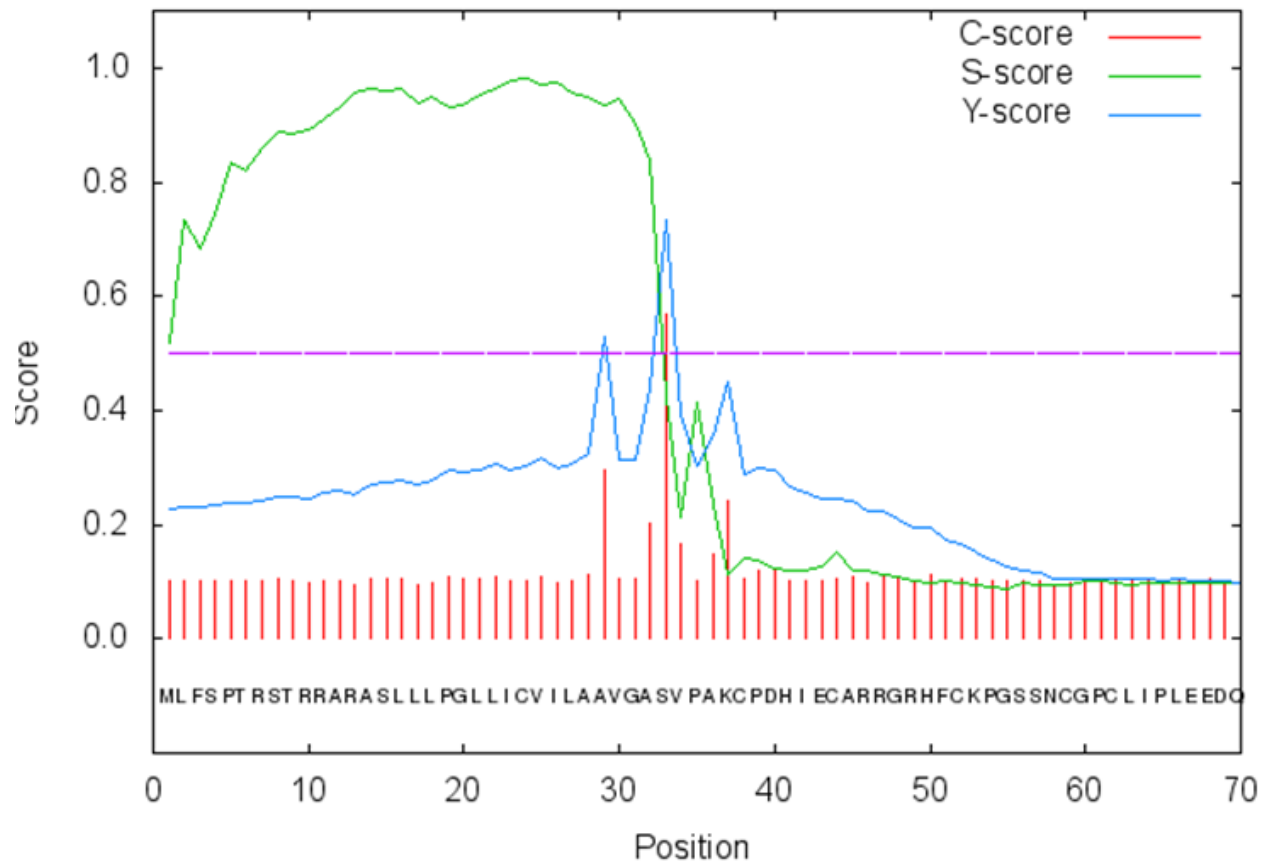- Dottpath: No repeated sequence was predicted.

4. Analysis of Protein

- **General feature：**
- ➢ Average residue weight: 109.721 g/mol
- ➢ Isoelectric point: 7.5587
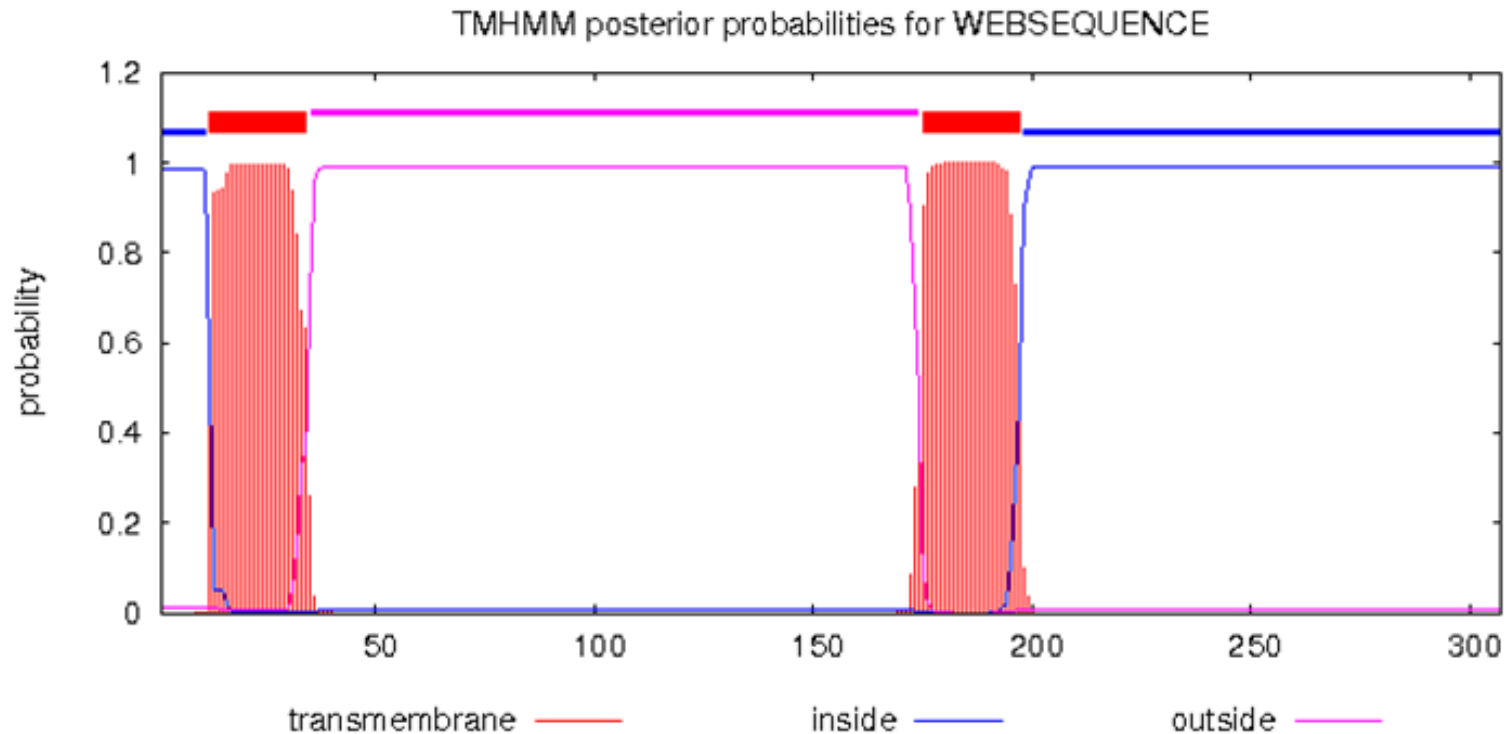- ➢ Molecular weight: 33,684.30 g/mol
- ➢ Number of residues: 307 aa

Signal peptide prediction

- 1-32 AA region may be a signal peptide

Transmembrane perdiction

TMHMM posterior probabilities for WEBSEQUENCE

- Red: Transmembrane region

SMART Prediction
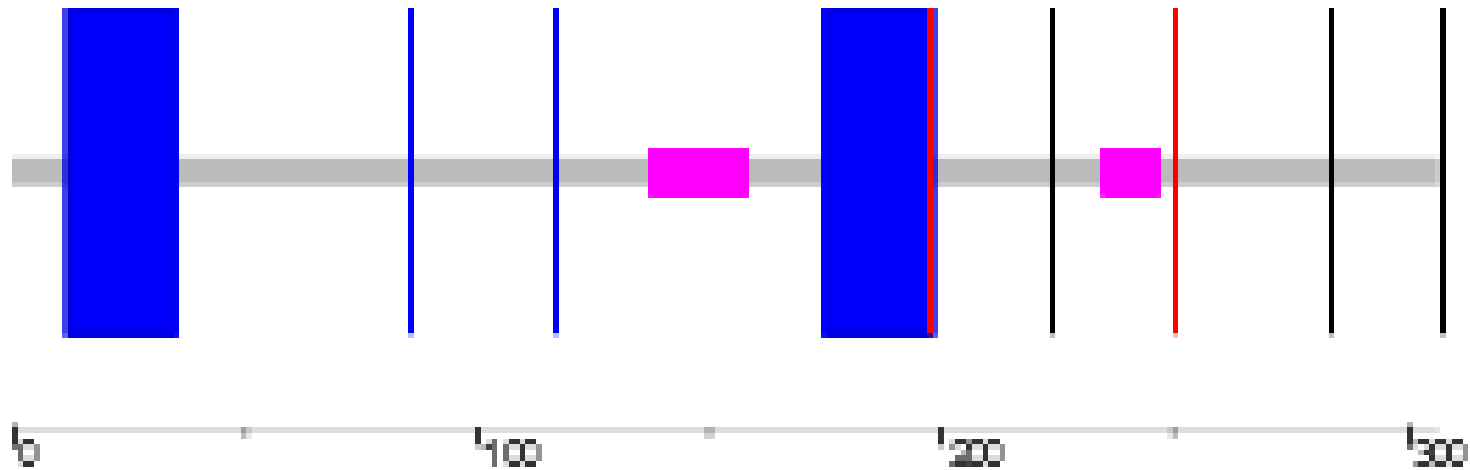
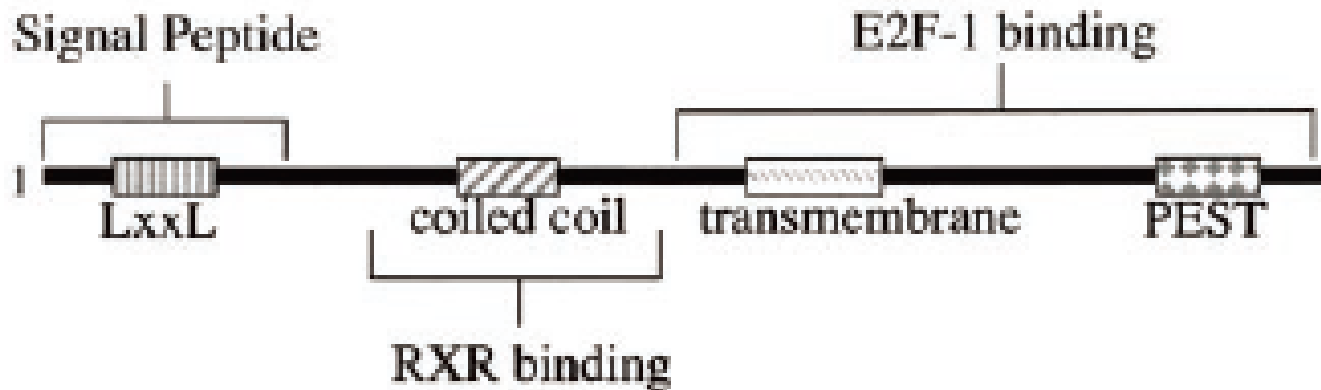
- Blue Box: Transmembrane domain
- line: Intron
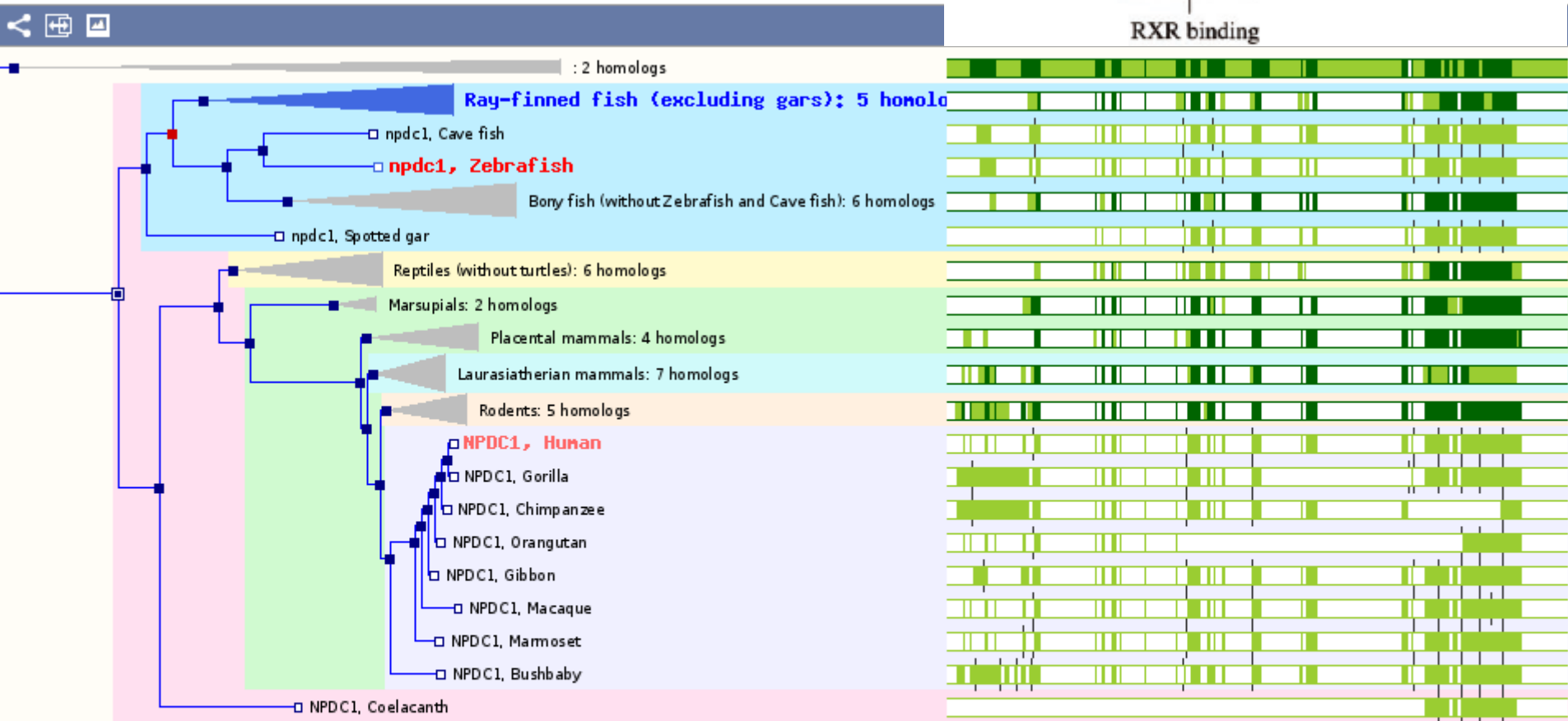- Pink box: Low complexity region
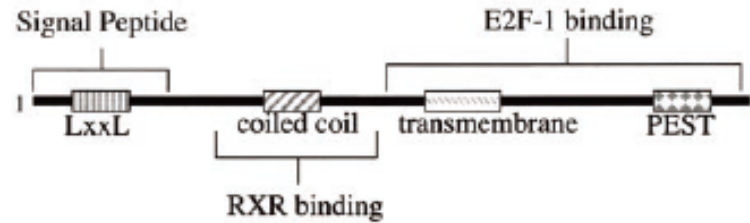
Motif of npdc1 protein

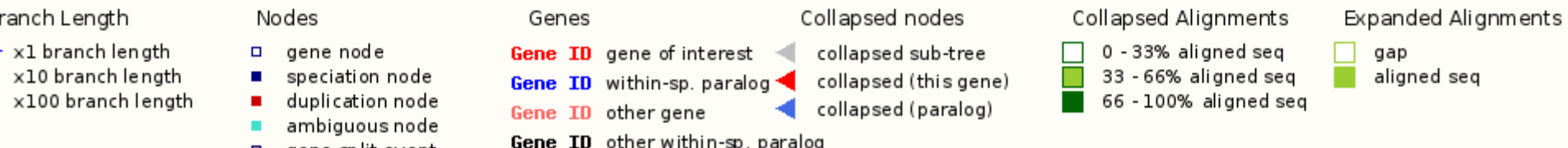Fig. 1. Representation of the NPDC-1 protein structure.

C. Evrard et al., 2005

Gene tree of npdc1 (Ensembl)

Protein Interaction

- **Protein function in human:**
- **Ddr1**: regulates cell migration, differentiation, survival and cell proliferation.
- **Nenf**: May play a role on neuronal differentiation and may have a transient effect on neural cell proliferation in neural precursor cells.
- **Ddx41**: Probable ATP-dependent RNA helicase. Is required during post-transcriptional gene expression. May be involved in pre-mRNA splicing.
- **Arhgap23**: GTPase activator for the Rho-type GTPases by converting them to an inactive GDP-bound state
- **Prmt2**: Represses E2F1 transcriptional activity (in a RB1-dependent manner). May be involved in growth regulation
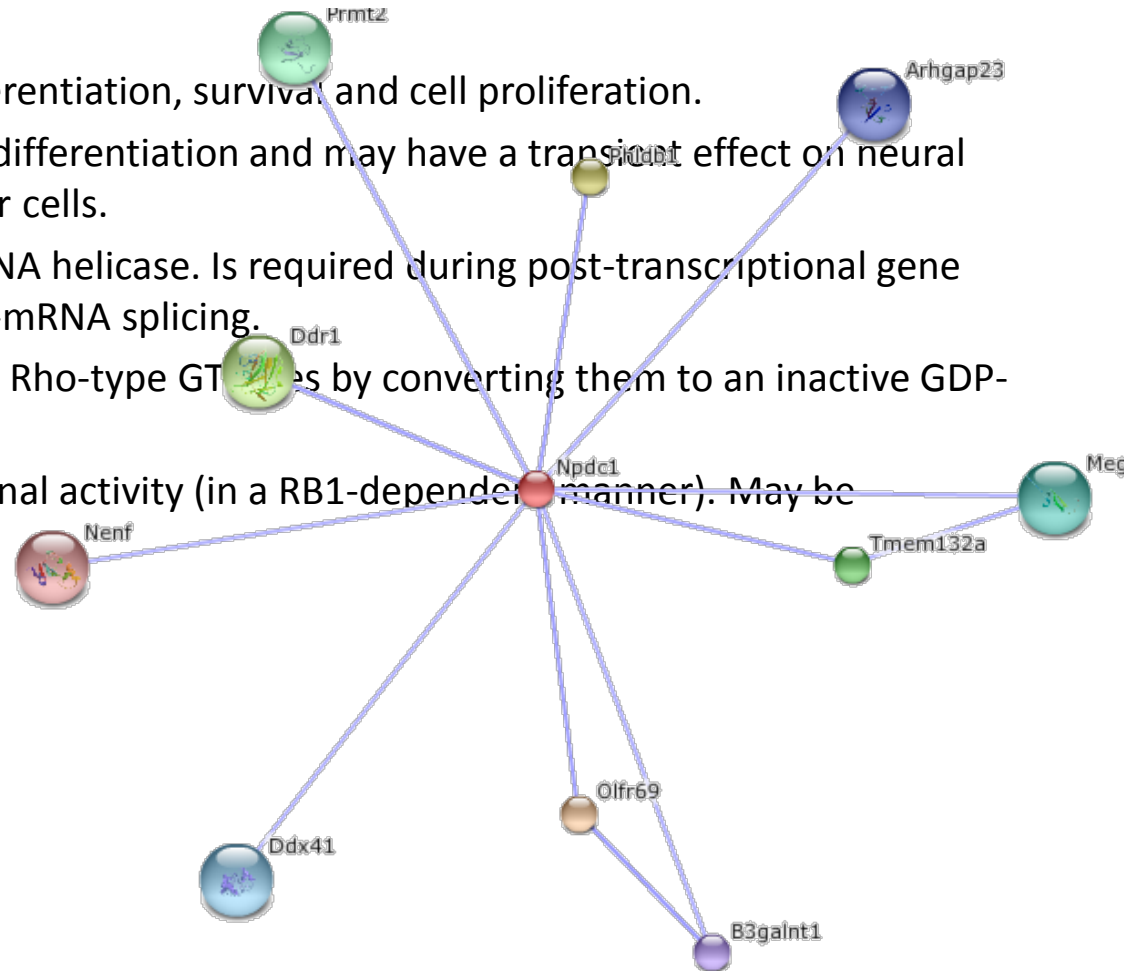
5. Blast and Phylogenetic tree

- Gene sequence blast

 Just finding that 800-1000bp regions were conserved.
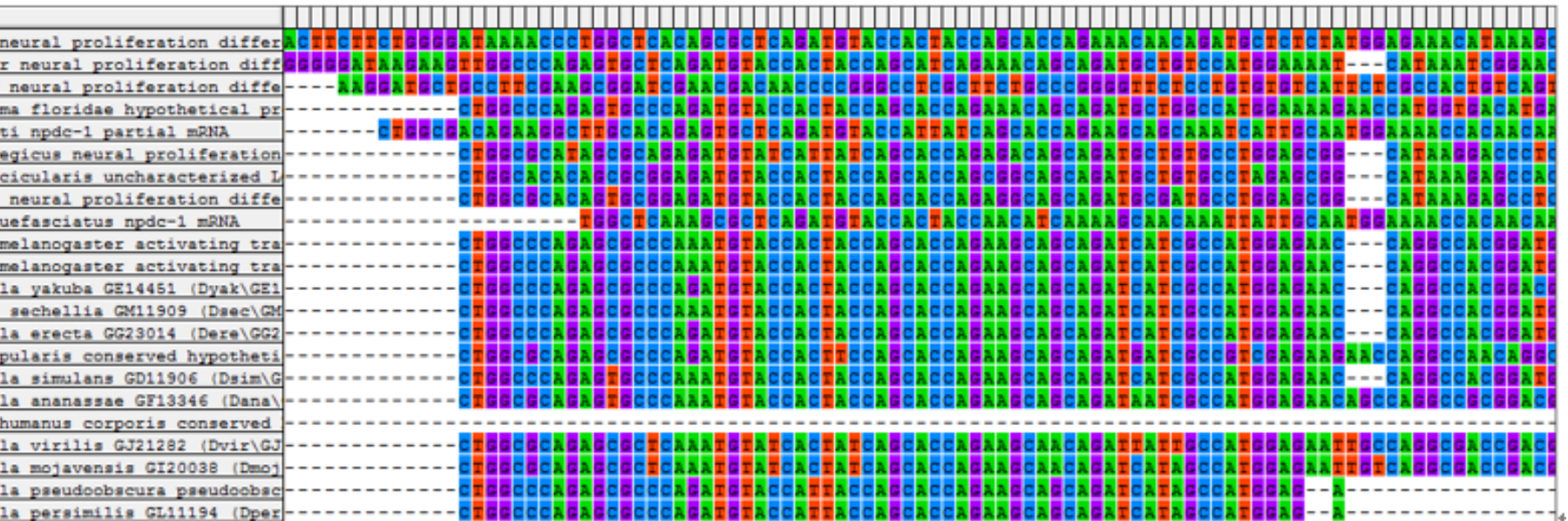
Blastp

- Use npdc1 protein sequence to Blastp, with BLOSUM62, finding that only Rat, Mouse protein npdc1 sequence were found.

Alignments ☰Download ⌄ GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession
RecName: Full=Neural proliferation differentiation and control protein 1; Short=NPDC-1; Flags: Precursor	162	162	88%	2e-45	42%	gi|22261810|Q9NQX5.2
RecName: Full=Neural proliferation differentiation and control protein 1; Short=NPDC-1; Flags: Precursor	161	161	85%	5e-45	41%	gi|341941179|Q64322.2
RecName: Full=Protein cab-1	63.9	63.9	42%	3e-10	35%	gi|6226923|Q93249.3
RecName: Full=Zinc metalloproteinase nas-9; AltName: Full=Nematode astacin 9; Flags: Precursor	34.7	34.7	31%	0.94	29%	gi|57012912|P91137.2

Uniprot annotation

- So we use Uniprot to get npdc1 protein sequences, but most of them were unreviewed.

- Multiple sequence alignment showed that C-terminal sequences were conserved.

6. Protein structure prediction and function

- Swissmodle: --No suitable templates found
- Phyre2:

---Just an a-helix (165-202 AA, 12%) was predicted
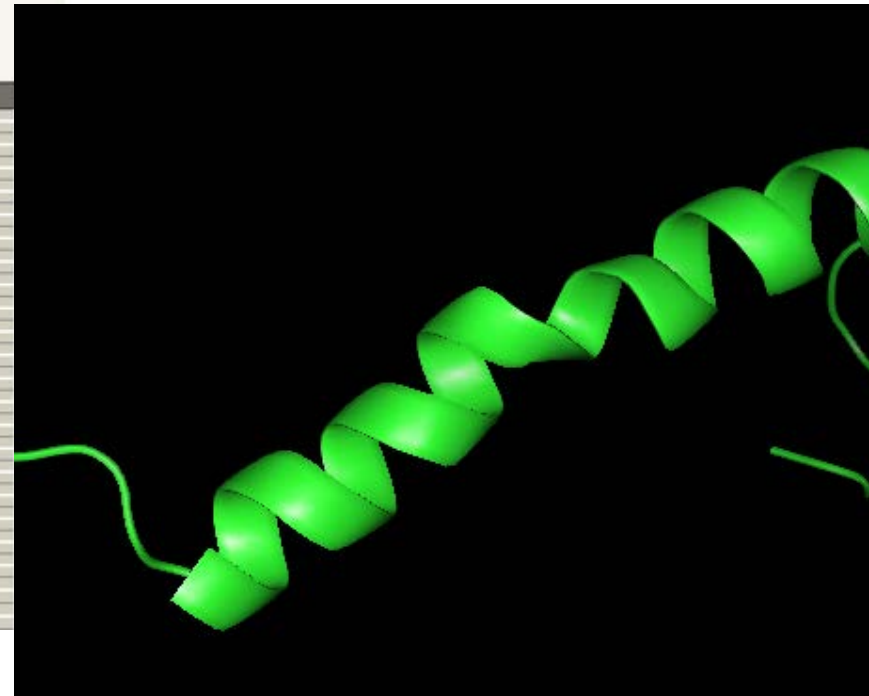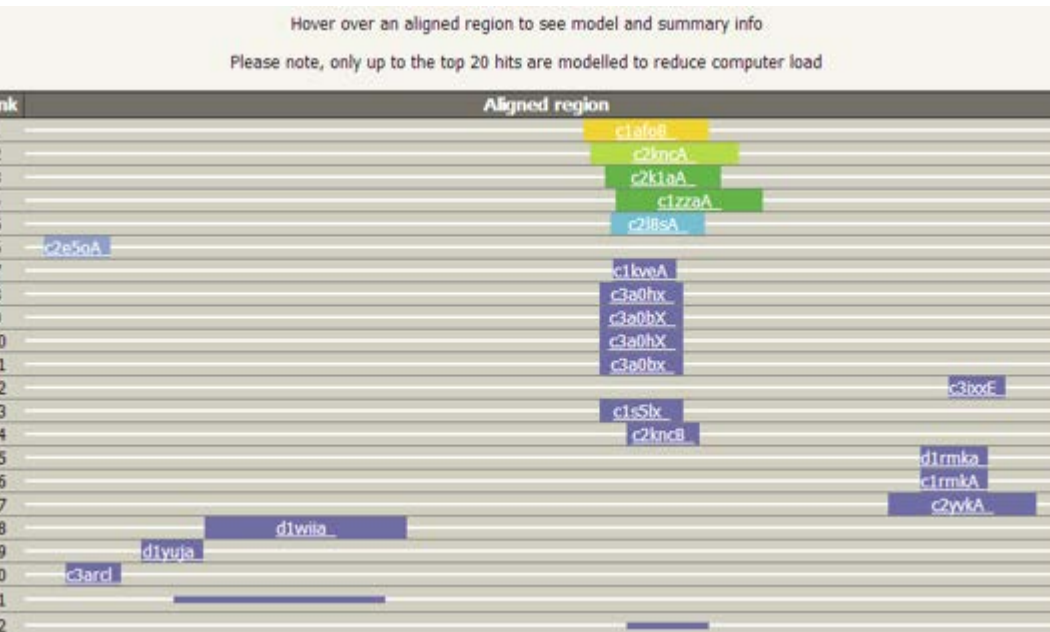
7. Conclusion

- Npdc1 is likely to be a transmembrane protein and interact with E2F-1,to reduce E2F-1 binding to DNA and transcription activity. In this way npdc1 may regulate cell proliferation and differentiation.

- Without npdc1 gene,E2F-1 may activate some gene expression, which may be involved in membrane generation or disintegration.

- Npdc1 may interact with Retinoid Acid pathway.

8. Acknowledgement

- We appreciate Professor Luo JC for teaching and discussion;

- We thank Professor Zhang Bo's guide in this project;

- We thank Wu Qian for providing materials; Zhang Ying for providing video of zebrafish development.

Thanks!