Daphnia Genes Encoding Multiple Tandem Zn-finger Proteins

Yang Bai and Gary W. Stuart

Department of Life Sciences, Indiana State University, Terre Haute, IN 47809

Introduction

Zinc-finger proteins probably represent the largest family of DNA-binding transcription factors in eukaryotes. A recent comparison of the full complement of C2H2 zinc-finger protein observed or predicted within worm, fly, and human genomes has lead to the tentative identification of nearly 40 orthologous groups shared between humans and invertebrates. From a phylogenetic perspective, the Daphnia genome would be expected to contain identifiable members for most of these groups. Likely orthologs already identified include those of the Sp, KLF, and MTF-1 groups. In addition, potential members of the Egr, Zfh, Gli, YY1, Evi-1, Spalt, and Zfam-2 groups have been identified. A homolog for the quintessential zinc-finger protein, TFIIIA, is also apparent. Our work to date suggests that while some early gene models are accurate, a high fraction of the C2H2 zinc-finger containing gene models identified by automatic annotation should probably include additional nearby C2H2 zinc-fingers missing from the models. As a result, ongoing annotation of putative C2H2 zinc-finger orthologs requires the detailed inspection and modification of current gene models.

The zinc-finger protein families Gli and YY1 considered in this poster are two examples of the roughly 40 zinc-finger protein families we are currently exploring in Daphnia genome.

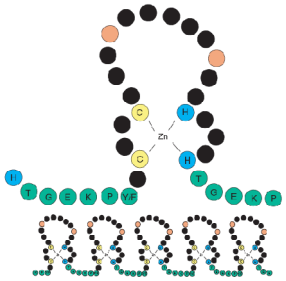

Figure 1. Schematic diagram of a C2H2 zinc-finger motif. Yellow – cysteines (C); Blue – histidines (H); Red – hydrophobic residues; Green – linker sequence; Black – non-structurally important residues.

Methods

Table 1. The 9 groups of orthologous C2H2 ZNF genes defined by our analyses

Gene Family	Human	Drosophila	C. elegans	Daphnia
Gli (5)	Gli1 (SP:P08151); Gli2 (SP:P10070); Gli3 (SP:P10071)	Ci (CT6641; SPTR:P19538)	Tra-1 (Y47D3A.6; SPTR:Q9U2C0)	fgenesh1_pg.C_scaffold_3000142
YY1 (4)	TYY1 (SPTR:P25490); SPTR:O15391	Pho (CT39329; SPTR:O76247); CT11601 (SPTR:Q9VSZ3)		E_GW1.78.11.1
Blimp (5)	BLIMP-1 (SPTR:O95914)	CT16759 (SPTR:Q9VRN4)	F25D7.3 (SPTR:Q93560)	e_gw1.29.264.1; e_gw1.247.2.1
Fez (6)	SPTR:Q9NWB9	CT22557 (SPTR:Q9VQ56)	Y38H8A.5 (SPTR:O62425)	e_gw1.2.698.1
Spalt (7)	HSAL1 (SPTR:Q99881); HSAL2 (SPTR:Q9Y467); SALL3 (SPTR:Q9UGH1)	Spalt-major (CT20082; SPTR:P39770); Spalt-related (CT15643; SPTR:Q24163)	SEM-4 (SPTR:Q17396)	fgenesh1_pg.C_scaffold_88000006
CTCF (11)	CTCF (SP:P49711) CTCF (ref:NP_542185)	CTCF (ref:NP_648109); GA21187-PA (gb:EAL30192)		estExt_Genewise1Plus.C_1580013
GFI (6)	GFI-1, GFI-1B (SPTR:Q99684)	CT31381 (SPTR:Q9VM77)	F45B8.4 (SPTR:O02265)	e_gw1.106.63.1; e_gw1.106.61.1
MTF (6)	MTF-1 (SPTR:Q14872)	CT12477 (SPTR:Q9NFS1)		estExt_fgenesh1_pg.C_720028
TFIIIA (9)	TFIIIA (sp:Q92664)	CG9609-PA (gb:AAF48663.1)		estExt_Genewise1.C_940065

- Choose one of the pre-identified zinc finger protein sequences from Human, Drosophila, or C. elegans as standard sequences (http://www.ncbi.nlm.nih.gov/)
- Use Blastp from JGI to search for the orthologs in Daphnia genome sequences (http://shake.jgi-psf.org/)
- Manually analyze and choose the best gene model in Daphnia genome
- Get orthologous sequences from different species and do clustalw
- Mark the zinc finger protein sequences with clustalx colors in Jalview 2.3 (http://www.jalview.org/)

Results

1.Gli

The Gli protein subclass of Krüppel-like zinc finger proteins consists of three members, Gli1, Gli2, and Gli3. Each member contains five tandem Cys2-His2 zinc finger motifs. Gli proteins are related to Cubitus interruptus (Ci), a Krüppel-like zinc finger protein important in wing development in *Drosophila melanogaster*. Ci and Gli proteins function as downstream regulators of transcription in the Sonic hedgehog (SHH)-Patched signal transduction pathway in Drosophila and vertebrates. Gli proteins can function as activators and repressors of gene transcription and play multiple roles in the regulation of prenatal mammalian development particularly in limb and craniofacial development. Gli1 has been reported to function as a proto-oncogene. Aberrant expression and amplification of Gli1 is associated with cell carcinomas of the skin, sarcomas, and glioblastomas. Gli2 and Gli3 have been reported to be essential in the organogenesis of various tissues, including lung, trachea, and esophagus, and Gli3 has been implicated in several genetic disorders in humans, such as Glioblastoma, Osteosarcoma and Greig Syndrome.

Figure 2. Gli zinc-finger region alignment Hsap–Homo sapiens; Agam–Anopheles gambiae; Aaeg–Aedes aegypti; Dmel–Drosophila melanogaster; Dpul–Daphnia pulex; Amel–Apis mellifera.

2.YY1

YY1 (Yin Yang 1) is a transcription factor found in both vertebrates and invertebrates. It is a Gli-Krüppel type zinc finger protein that contain four C2H2 zinc-fingers, and function as repressors, activators and transcription initiators depending upon the sequence context of YY1-binding sites with respect to other regulator elements. The proteins has a DNA-binding domain at the C-terminus and other modulating domains at the N-terminus displaying repression, activation and protein/protein interaction activities. YY1 can interact with several key transcription factors, including TBP, TAFs, TFIIB and Sp1, as well as histone-modifying complexes, such as p300, HDACs, PRMT1 and Polycomb complexes.

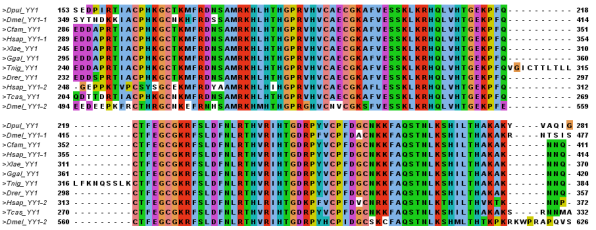

Figure 3. YY1 zinc-finger region alignment Cfam–Canis familiaris; Tcas–Tribolium castaneum; Drer–Danio rerio; Xlae–Xenopus laevis; Tnig–Tetraodon nigroviridis; Ggal–Gallus gallus.

Conclusions

- Some of the Daphnia gene models need further inspection and manual modification
- Zinc-finger proteins Gli and YY1 are highly conserved in Daphnia genome
- Molecular phylogenetic reconstruction of gene family history can help distinguish speciation from gene duplication, and reveal orthologous and paralogous relationships

References

1. Knight RD and Shimeld SM. (2001). Identification of conserved C2H2 zinc-finger gene families in the Bilateria. *Genome Biol*. 2(5), 0016.1 – 0016.8 (http://genomebilogy.com/2001/2/5/research/0016).
2. Zhang F, Nakanishi G, Kurebayashi S, Yoshino K, Perantoni A, Kim YS, and Jetten AM. (2002) Characterization of Glis2, a novel gene encoding a Gli-related, Krüppel-like transcription factor with transactivation and repressor functions. *J Biol Chem*. 277, 10139-10149.
3. Hosking CR, Ulloa F, Hogan C, Ferber EC, Figueroa A, Gevaert K, Birchmeier W, Briscoe J and Fujita Y. (2007). The transcription repressor Glis2 is a novel binding partner for p120 catenin. *Mol Biol Cell*. 18, 1918 – 1927.
4. Kim JD, Faulk C, and Kim J. (2007). Retroposition and evolution of the DNA-binding motifs of YY1, YY2 and REX1. *Nucleic Acids Res*. 35, 3442 – 3452.