Available online at www.sciencedirect.com
SCIENCE DIRECT
Mechanisms of Ageing and Development 126 (2005) 209–215
mechanisms of ageing and development
www.elsevier.com/locate/mechagedev

From worm to human: bioinformatics approaches to identify FOXO target genes

Zhenyu Xuan, Michael Q. Zhang*

Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

Available online 26 October 2004

Abstract

Longevity regulatory genes include the Forkhead transcription factor FOXO, in addition to NAD-dependent histone deacetylase silent information regulator 2 (Sir2). The FOXO/DAF-16 family of transcription factors constitute an evolutionarily conserved subgroup within a larger family known as winged helix or Forkhead transcriptional regulators. Here we demonstrate how to identify FOXO target genes and their potential cis-regulatory binding sites in the promoters via bioinformatics approaches. These results provide new testable hypotheses for further experimental verifications.

Keywords: Foxo target genes; Insulin pathway; Comparative genomics; Aging

1. Introduction

FOXO (Forkhead box, class O) subfamily of Forkhead transcription factors has been identified as direct downstream targets of phosphoinositide 3-kinase (PI3K) in the insulin/insulin-like growth factor signaling pathway. Insulin-mediated activation of PI3K increases 3′-phosphorylated phosphoinositide lipids' (PIP_3) production, and PIP_3 then recruits protein kinase Akt to phosphorylate FOXO proteins at serine/threonine residues. This phosphorylation induces a shuttling mechanism that retains FOXO factors in the cytoplasm, thereby activating or suppressing target gene expression (Burgering and Kops, 2002) that affect cell cycle progression (Kops et al., 1999; Alvarez et al., 2001), apoptosis (Brunet et al., 1999; Dijkers et al., 2000), and metabolism (Ayala et al., 1999; Durham et al., 1999; Hall et al., 2000; Nakae et al., 2001; Nadal et al., 2002). Thus, FOXO protein is a key component in this insulin-signaling cascade. In invertebrates, this pathway apparently plays an essential role in regulating life span as well as body, organ, and cell size (Finch and Ruvkun, 2001; Tatar et al., 2001). A recent study of the mouse insulin-like growth factor type 1 receptor (Igf1r) showed that Igf1r might be a central regulator of mammalian life span (Holzenberger et al., 2003).

Genetic analyses in the nematode worm *C. elegans* have found that inhibition of the worm FOXO transcription factor, DAF-16, in the *daf-2*/insulin-like signaling pathway can regulate organism lifespan (Kenyon et al., 1993). Further research has resulted in the discovery that many DAF-16 target genes mediate distinct aspects of *daf-16* function, including longevity, metabolism, and development (Lee et al., 2003). The identification of more DAF-16/FOXO target genes in worm and higher eukaryotes will likely contribute to our understanding of basic mechanisms of aging.

Comparative genomics method is to use known functional information in one species to discover potential function of gene, metabolism mechanism, or signal transduction pathway in other related species, based on the evolutionary conservation information. The basic assumption is that many essential regulation mechanisms are conserved from lower species to higher ones during the evolution, such as cell cycle, cell growth and survival, chromatin assembly and gene activation, and so on. This implies that across a wide variety of animal species, pathways sharing a common function may also share a common origin in their ancestor.

Many bioinformatics approaches and tools were developed for comparative genomics analysis. Dynamic pro-

* Corresponding author. Tel.: +1 516 367 8393; fax: +1 516 367 8461.
E-mail address: mzhang@cshl.edu (M.Q. Zhang).

0047-6374/$ – see front matter
doi:10.1016/j.mad.2004.09.021

gramming was used to align two or more sequences to find the conservation, such as BLAST (Altschul et al., 1990), FASTA (Pearson, 2000), CLUSTALW (Higgins et al., 1996), mLAGAN (Brudno et al., 2003), etc. One can find orthologous relationship of genes by analyzing their protein sequences with these tools. Conserved sequence blocks, such as protein domains, transcription factor binding sites, can also be found by using these programs. The second class of approaches is to use statistics methods. Based on known functional information, one can build statistical models and apply them to discover novel targets related to this function. These methods include motif-finding algorithm, such as Gibbs Sampler (Lawrence et al., 1993), MEME (Bailey and Elkan, 1995); promoter and gene finders, such as FirstEF (Davuluri et al., 2001), Twinscan (Korf et al., 2001). With the rapid increasing of the biological data, especially those high-throughput data, like genome sequences, microarray and chromatin-immunoprecipitation (ChIP) data, using statistical methods will greatly strengthen the power to analyze them in a systematic and comprehensive manner.

Because the pathway from *daf-2*/insulin receptor to *daf-16*/FOXO is evolutionarily conserved from nematode to vertebrates (Finch and Ruvkun, 2001; Birkenkamp and Coffer, 2003; Holzenberger et al., 2003), here, we will use these comparative genomics approaches to identify potential mammalian FOXO target genes. The mammalian DAF-16 orthologues include FOXO1 (FKHR), FOXO3a (FKHR-like 1), and FOXO4 (AFX). These FOXO proteins interact preferentially with a common core consensus motif called daf-16 family protein-binding element (DBE) (Furuyama et al., 2000; Biggs et al., 2001). By searching DBEs on the mammalian promoters in our mammalian promoter databases (Zhang, 2003), we can identify mammalian FOXO target genes either through finding mammalian orthologues of worm DAF-16 target genes, or by analyzing motif conservation in the mammalian gene promoters.

2. Approach I: finding mammalian orthologs of worm DAF-16 target genes

Promoters of higher organisms usually have a longer distance to ATG than those of lower species, due to the longer 5′-UTR and potential introns. We have to use known or predicted promoters to search for transcription factor binding sites, such as DBE, instead of upstream of ATG. Our mammalian promoter database has collected all of human, mouse, and rat known promoters from Eukaryotic Promoter Database (EPD) (Cavin Perier et al., 1998), DataBase of Transcriptional Start Site (DBTSS) (Suzuki et al., 2002), and Genbank (Benson et al., 2003), and has predicted promoters by FirstEF, which will facilitate the motif analysis and target gene identification.

The DAF-16/FOXO binding sites were searched for upstream of the *C. elegans* gene and its orthologous promoters, to identify functional DAF-16 sites in conserved components of the DAF-16 transcriptional cascade (Lee et al., 2003). The mammalian orthologues of worm genes were identified based on their protein sequence similarity. We used the reciprocal BLAST analysis to find the potential orthologues, which has been used for orthologue prediction in Ensembl (http://www.ensembl.org/) and HomoloGene (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?DB=homologene). DBE is a degenerated motif among different species. The first motif consensus sequence found is TTGTTTAC, which is located in worm *sod3* and human *SOD2* promoters. But it cannot be found in promoter of the mouse orthologue. Comparing 17 *C. elegans* genes containing DBE with their *C. briggsae* orthologues, we found that five *C. briggsae* genes' upstream have this motif, and nine have motifs with only one mismatch from the consensus. Further in vitro PCR-assisted binding site selection experiments showed that the degeneration of the first, 6–8th sites are not random (Furuyama et al., 2000; Biggs et al., 2001). Therefore, we use DBEs one-site degenerated patterns (called DBEm1) based on Biggs and Furuyama's works, to search the promoter sequences of orthologous genes in different species. The orthologues of known DAF-16 target genes were assumed to also have DBEm1 sites in their promoters.

We checked the six *C. elegans* genes, which show different expression in the wildtype and *daf-2*(−) animals, and one gene that extends the life span of worm after RNAi inactivation. Worm *sod3* was used as the control. DBE was found in promoters of all of eight genes. We searched DBEm1 in the different regions in the promoters of these

Table 1
DAF-16 binding site (DBE) and its one-site degenerated pattern (DBEm1) in promoters of mammalian orthologues of eight known worm DAF-16 target genes

Worm gene	Human gene	2 kb promoter			3 kb promoter			4 kb promoter			5 kb promoter		
		H	M	R	H	M	R	H	M	R	H	M	R
C08A9.1	*MnSod*	+	+	+	+	+	+	+	+	+	+	+	+
F52H3.5	Novel gene	+	+	+	+	+	+	+	+	+	+	+	+
F43G9.5	*CPSF5*	−	+	+	+	+	+	+	+	+	+	+	+
F14F4.3	*ABCC5*	−	+	+	−	+	+	+	+	+	+	+	+
C39F7.5	*FLJ10648*	+	−	−	+	−	−	+	+	+	+	+	+
ZK593.4	*RBBP2*	−	−	+	−	−	+	+	−	+	+	+	+
T21C12.2	*HPD*	−	+	−	−	+	+	−	+	+	−	+	+
C10G11.5	*PANK4*	−	−	−	−	−	−	−	−	−	−	−	+

H: human; M: mouse; R: rat; +: motif locates in the certain promoter region; −: otherwise.

Table 2
Thirty-two genes with DBE located in all of the human, mouse and rat 1 kb promoter regions with function described in Gene Ontology (The Gene Ontology and Consortium, 2000) or literature

Human gene	Definition and function
Known FOXO target genes	
INSR	Insulin receptor
PDK4	Pyruvate dehydrogenase kinase, isoenzyme 4
Potential FOXO target genes with known function	
TXNIP	Thioredoxin interacting protein, GO: biological process unknown
TLP19	Endoplasmic reticulum thioredoxin superfamily member, 18 kDa, GO: electron transport
FEN1	Flap structure-specific endonuclease 1, GO: DNA replication/repair
BTG1	B-cell translocation protein 1, GO: cell proliferation
PLXNC1	Plexin C1, GO: cell adhesion/development
MLH3	Mismatch repair gene MLH3, GO: meiotic recombination/mismatch repair
IGF1R	Insulin receptor signaling pathway
SLC12A6	Solute carrier family 12 (potassium/chloride transporters), member 6
TFAP4	Transcription factor AP-4 (activating enhancer binding protein 4)
SSB3	SPRY domain-containing SOCS box protein SSB-3, GO: intracellular signaling cascade
PER1	Period (*Drosophila*) homolog 1, GO: regulation of transcription, DNA-dependent
DHX8	DEAH (Asp-Glu-Ala-His) box polypeptide 8, GO: RNA splicing
FLJ10597	Function unknown, GO: protein ubiquitination.
STK11	Serine/threonine protein kinase 11, related to growth suppression (Tiainen et al., 2002)
TFDP2	Transcription factor Dp-2 (E2F dimerization partner 2), GO: regulation of cell cycle
ELOVL6	ELOVL family member 6, elongation of long chain fatty acids (FEN1/Elo2)
LCP2	Lymphocyte cytosolic protein 2, GO: regulation of blood and lymphatic vascular separation
CITED2	Cbp/p300-interacting transactivator, GO: regulation of transcription from Pol II promoter
AP4M1	Adaptor-related protein complex 4, mu 1 subunit, GO: intracellular protein transport
ASBABP2	Pregnancy-associated plasma protein A, pappalysin 1. (Chen et al., 2003)
Potential FOXO target genes with function unknown	
FLJ12168	TBC1 domain family, member 17, unknown function
FLJ12221	Zinc finger, SWIM domain containing 4, unknown function
CNNM3	Cyclin M3,ancient conserved domain protein 3, unknown function
FLJ23142	Secernin 3, unknown function
KIAA1155	Unknown function
DJ465N24.2.1	NPD014 protein, function unknown
C5orf6	Unknown function
FLJ13611	Hypothetical protein
FLJ23209	PDZK7 protein, hypothetical protein, function unknown
LOC220213	OTUD1: OTU-like cysteine protease, function unknown

worm genes' orthologues in the human, mouse, and rat. By searching the 2 kilobase pairs (kb) region upstream of the transcription start site (TSS), we found only mammalian *SOD2* and a novel gene that is orthologue of worm *F52H3.5* gene contain DBEm1 in all of three mammals (see Table 1). DBEm1 site cannot be found in the 5 kb upstream of TSS of *HPD* (orthologue of worm *T12C12.2*) and *PANK4* (orthologue of worm *C10G11.5*) genes in either human, mouse, or rat. So we thought that the mammalian orthologs of *C. elegance F52H3.5*, *F43G9.5* (*CPSF5*), *F14F4.3*, *C39F7.5*, and *ZK593.4* (*RBBP2*) genes are potential FOXO target genes. The reason for missing DBEm1 sites in *HPD* and *PANK4* may be due to the fact that either they are not target genes in mammal, or there are alternative promoters missing in the gene annotation, which may locate further than 5 kb upstream of the genes. Further analysis, such as identifying additional regulatory regions, and experiments to test the FOXO direct binding of these genes, could help to further clarify those results.

3. Approach II: finding mammalian FOXO target genes with conserved DBE motifs

Except starting from the worm target genes, we can also predict potential mammalian FOXO target genes by searching DBE in the mammalian orthologous genes. All of mammalian orthologous genes were identified by reciprocal BLAST analysis of their protein sequences, which has been done in Ensembl. Promoter information of each human, mouse, and rat gene in one orthologous gene group was extracted from our promoter database, and motif DBE and DBEm1 were searched in all of these promoters by using Perl regular expression pattern search function. We found 32 mammalian orthologous genes with DBE locating in 1 kb region upstream of TSS in all of three species (see Table 2). One of these genes, *PDK4* (pyruvate dehydrogenase kinase 4), were found that its expression level was up-regulated through the direct binding of FKHR to the promoter region of the gene in C2C12 cells (Furuyama et al., 2003).

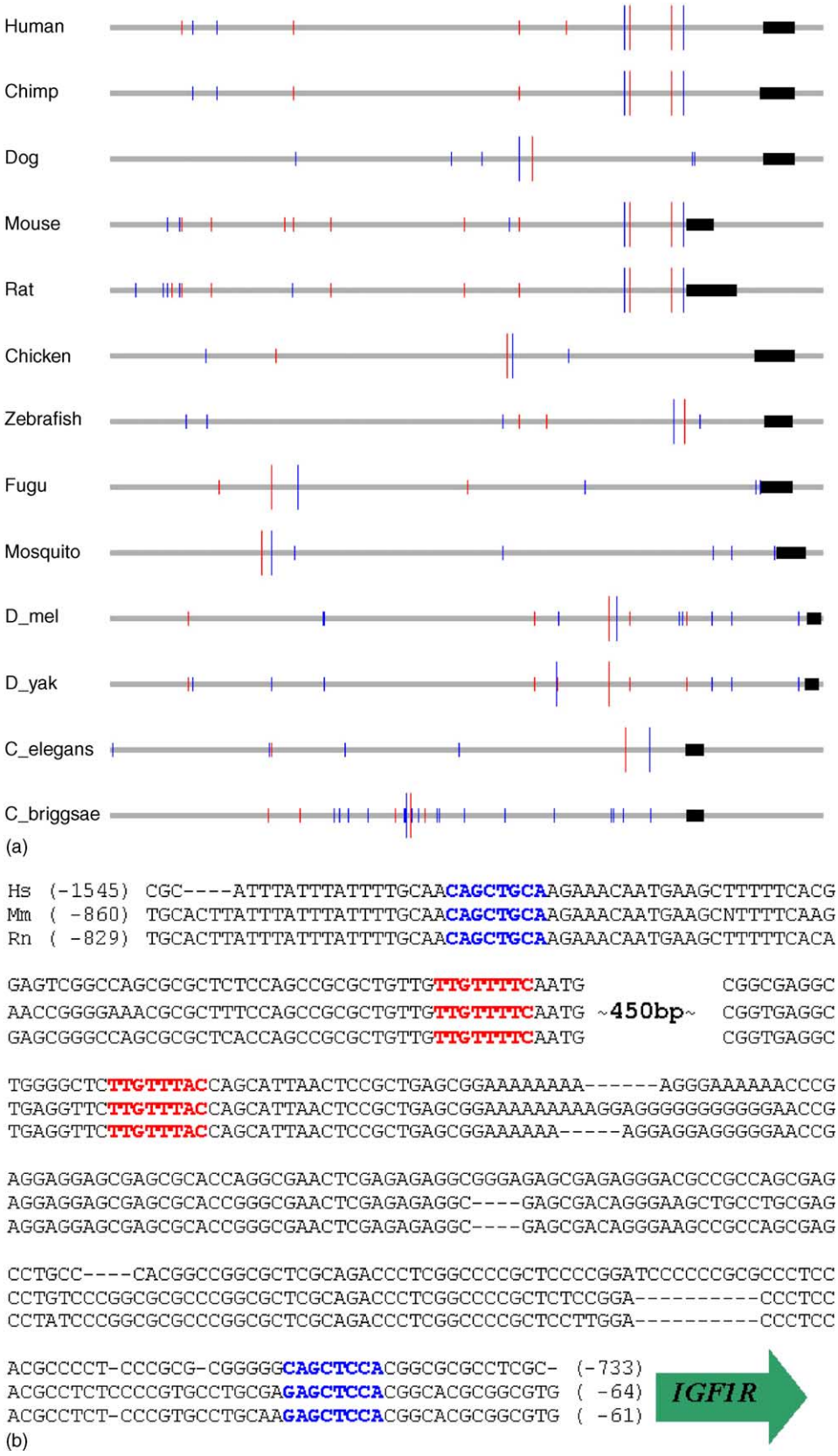

Fig. 1. (a) FOXO binding sites (DBEm1) and potential co-binding sites locations in the *IGF1R* promoters in 13 species: chimpanzee, human, mouse, rat, chicken, dog, zebrafish, fugu, mosquito, *D. melanogaster*, *D. yakuba*, *C. elegans* and *C. briggsae*. For human, mouse, and rat, the promoter sequences are 5 kb upstream of TSS, chimpanzee promoter is the matched region of human promoter. For the other species, the promoters are 5 kb upstream region of gene 5′-end

DBE was also found conserved in mammalian insulin receptor (*INSR*) and insulin-like growth factor type 1 receptor (*IGF1R*) genes. Both of them are homologues of *daf-2* (*C. elegans*) and *dInR* (*Drosophila*). Recent research in *Drosophila* has shown that dFOXO binds to the *dInR* promoter and activates its transcription (Puig et al., 2003). Interestingly, same as *daf-16* and *daf-2*, both *Drosophila* and mammalian FOXO are also regulated by the InR/PI3K/Akt pathway. This causes a feedback regulation of the insulin-signaling pathway by dFOXO. The conservation of the DBE in *INSR* and *IGF1R* promoters may implicate that this feedback regulation also exists in the mammals.

All three mammalian B-cell translocation protein 1 (*BTG1*) genes have DBE in their promoters. Microarray experiments showed that human *BTG1* expression levels were induced by Ad-FKHR;AAA but Ad-FKHR;HRAAA, which Ad-FKHR;AAA can bind DBE and Ad-FKHR;HRAAA cannot (Ramaswamy et al., 2002). Both motif analysis and mRNA expression data implicate that *BTG1* could be a direct target gene of FOXO.

We also found another 45 orthologous genes in human, mouse, and rat contain DBE within 2 kb upstream of their TSSs in all species (Table 1 in Supplementary information), which include FOXO3A, one member of FOXO subfamily (indicating a potential auto-regulation feedback loop). The relationship between most of these genes and longevity is unknown, and some genes are just novel genes without any known functions. Our DBE conservation analysis can shed light on their function discovery.

We also searched DBEm1 in promoters of mammalian orthologous genes. The known FOXO target genes, such as *SOD2*, *G6PC*, *BCL2L11/BIM*, and *CCNG2*, were found with DBEm1 in their promoters in all three species. Among all found genes, 16 human genes showed different expression level after infection with Ad-FKHR;AAA or Ad-FKHR;HRAAA. These 16 genes included *PAWR*, *GCA*, *SOX4*, and *MYO6*, which were only induced by Ad-FKHR;AAA (see Table 2 in Supplementary information). These genes in Supplementary information might be the most valuable candidates of FOXO target genes for further experimental test.

There are two questions need to be discussed. The first is the functional differences of mammalian FOXO members. There are three or more members of FOXO subfamily (Biggs et al., 2001) in mammal while only one in nematode worm. Although three of them, FOXO1, FOXO3, and FOXO4 can bind the same consensus sequence DBE as DAF-16 does, their tissue expression patterns are different. FOXO4 mRNA is expressed at a high level in muscle, and FOXO1 in adipose tissue, while FOXO3 may function in the mature tissues. The various FOXO members and isoforms can have distinct biological effects within a single cell (Medema et al., 2000; Ramaswamy et al., 2002). These data suggest that the mechanism by which FOXOs induce a blockage in proliferation is distinct from that inducing cell death. By comparing the hormone response of DAF-16 and its mammalian homologues in HepG2 cells, Nasrin and colleagues (Nasrin et al., 2000) found that DAF-16 and FOXO1 (FKHR) were most similar in their ability to activate gene transcription and modulate the response of the IGFBP-1 promoter to glucocorticoids and insulin. The second question is relationship between DBE and insulin response sequence (IRS). IRS is defined as TT(G/A)TTT(T/G)(G/T) (Streeper et al., 1997; O'Brien and Granner, 1996) and is different from DBE at the last two bases. FOXO1 can bind both of them, but the binding to DBE is stronger (Furuyama et al., 2000). It is unclear whether there is difference for FOXO1 to bind DBE or IRS when both are existed. One possibility is that the flanking regions of IRS and DBE may be critical for binding, such differential binding was also reported for another Forkhead family member, FREAC-3 (Pierrou et al., 1994). The binding-site selection showed that each FOXO member has a different optimal DNA sequence specificity in the 5′-end of DBE. This may also implicate the differential target gene recognition for each DAF-16 homologue. More detailed experiments are needed to answer these questions.

4. Detecting FOXO co-factor binding site with comparative genomics method

Forkhead proteins usually interact with other associated proteins, and together, they bind on the promoter to regulate gene expression. The co-binding factors, such as p300/CREB-binding protein (CBP), CCAAT/enhancer-binding protein (C/EBPβ), and dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A and B (DYRK1A, DYRK1B) (Chan and La Thangue, 2001; Christian et al., 2002; von Groote-Bidlingmaier et al., 2003), can interact with FOXO to make a complex, while some also bind DNA at the same time. To identify the potential binding sites in the promoter region will help to find the corresponding co-factor. If the regulation of FOXO and its target gene expression is evolutionary conserved, we could also use comparative genomics method to find conserved motifs close to DBE as the potential co-factor binding sites.

We analyzed more promoter sequences of orthologous genes to predict those potential binding sites. Here, we use *IGF1R* as an example. We identified human *IGF1R* orthologues in other 12 species, including chimpanzee, dog, chicken, mouse, rat, fugu, zebrafish, mosquito, *Drosophila*

defined by cDNA or predicted transcript. *MLAGAN* was used to do the multiple sequence alignment. Grey line present genomic DNA, and the black blocks present 5′-exon. DBEm1 site is shown as red vertical lines, with co-bind site in blue. The taller lines mean either the closest pair of DBEm1 and co-binding sites in one species, or aligned site pairs in three or more species. (b) Multiple sequence alignment by *MLAGAN* with aligned site pairs in human, mouse, and rat. DBEm1 is in red and co-binding site is in blue. Green arrow presents the gene.

Table 3
Potential co-binding sites in all 13 insulin-like growth factor receptor (IGF1R) orthologues' promoters

Species	DBE/DBEm1	Potential co-binding site sequences	
C. elegans	**TTGTTTAC**	GAGCTGCA (−104)	AAGCTGAA (152)
C. briggsae	**TTGTTTAC**	CAGCTGGA (−162)	CAGCTGGT (−140)
			GAGCTTGT (−92)
			TAGCTTCA (1)
			TAGCTGGT (155)
D. melanogaster	**TTGTTTAC**	TAGCTTAT (−69)	AAGCTGAA (−92)
D. yakuba	TTGTTTGC	AAGCTTGT (77)	
Mosquito	TTGTTTGC	CAGCTGGT (10)	
Zebrafish	**TTGTTTAC**	GAGCTGCT (130)	CAGCTTGA (−124)
Fugu	**TTGTTTAC**	AAGCTGCA (188)	
Chicken	TTGTTGAC	CAGCTGCA (118)	
Mouse	**TTGTTTAC**	GAGCTCCA (191)	
Rat	**TTGTTTAC**	GAGCTCCA (185)	
Dog	TTGTTTAT	TAGCTCCT (−143)	
Human	**TTGTTTAC**	CAGCTCCA (193)	
Chimpanzee	**TTGTTTAC**	CAGCTTCA (193)	

The distance between two motifs are given in parentheses, which is negative when DBE/DBEm1 in the downstream. DBE sites are shown in bold.

melanogaster, *Drosophila yakuba*, *C. elegans* and *C. briggsae*. For human, mouse, and rat genes, we collected their promoter information from our promoter database. Chimpanzee promoters were found by aligning corresponding human promoters with the chimpanzee genome. For other species, we just extracted 5 kb upstream regions of the gene as the promoters. Nine of these 13 promoters have DBE sites, while the other four only have one-site degenerated motif DBEm1. Because the co-factor also interacts with FOXO protein, we assumed that the potential co-factor binding site should be close to FOXO binding site and also evolutionary conserved. We then extracted the DBE/DBEm1 site with 200 bp flanking in both ends from each promoter sequences. Gibbs Sampler was used to detect the significantly distributed motifs as the protein binding sites in all 13 sequences, each of which is 408 bp long. We searched motifs with length from 6 to 8 bp. As expected, DBEm1 was found first. After excluding those motifs very similar with single- or di-nucleotide repeats, we only found one motif AGCT(C/G/T)(A/C/G)(A/T) existed in all 13 sequences. Some sequences have more than one site (see Fig. 1a). We also found that the distance between this motif and DBEm1 is not conserved in all species, although two pairs of these motifs are conserved and aligned in the human, mouse, and rat (see Fig. 1b). The reason that we did not find the aligned motif pairs in dog promoter may due to the incompleteness of the genome. However, these motif sites did reveal some evolutionary traces: The number of mismatches between the pair of motifs in two species is positively correlated to the evolutionary distance of these species (see Table 3). After searching the TRANSFAC (Matys et al., 2003) database for all known binding sites, we did not find any motif similar to this one. We thought that the motif AGCT(C/G/T)(A/C/G)(A/T) could be a novel co-factor binding site which await further experimental verification.

5. Summary

Insulin/insulin-like growth factor pathway plays an essential role in regulating life span and metabolism. The conservation of this pathway during evolution impels us to find the components of the pathway by using comparative genomics methods. Here we present two theoretical approaches to identify potential FOXO direct downstream targets. We also used motif conservation in the promoters of orthologous genes to predict potential FOXO and co-factor binding sites. All of the predictions could be the valuable starts for experimental test. The combination of theoretical and experimental methods will greatly benefit the discovery of mechanism of aging, as well as the understanding of biological and pathological metabolic pathways.

Acknowledgement

This work is supported by NIH grants HG002600 and HG001696.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at 10.1016/j.mad.2004.09.021.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Alvarez, B., Martinez, A.C., Burgering, B.M., Carrera, A.C., 2001. Forkhead transcription factors contribute to execution of the mitotic programme in mammals. Nature 413, 744–747.

Ayala, J.E., Streeper, R.S., Desgrosellier, J.S., Durham, S.K., Suwanichkul, A., Svitek, C.A., Goldman, J.K., Barr, F.G., Powell, D.R., O'Brien, R.M., 1999. Conservation of an insulin response unit between mouse and human glucose-6-phosphatase catalytic subunit gene promoters: transcription factor FKHR binds the insulin response sequence. Diabetes 48, 1885–1889.

Bailey, T.L., Elkan, C., 1995. The value of prior knowledge in discovering motifs with MEME. Proc. Int. Conf. Intell. Syst. Mol. Biol. 3, 21–29.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Wheeler, D.L., 2003. Genbank. Nucl. Acid Res. 31, 23–27.

Biggs, W.H., , W.K., Cavenee, K.C., Arden, J., 2001. III. Mamm. Genome 12, 416–425.

Birkenkamp, K.U., Coffer, P.J., 2003. Regulation of cell survival and proliferation by the FOXO (Forkhead box, class O) subfamily of Forkhead transcription factors. Biochem. Soc. Trans. 31, 292–297.

Brudno, M., et al., 2003. LAGAN and multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. Genome Res. 13, 721–731.

Brunet, A., Bonni, A., Zigmond, M.J., Lin, M.Z., Juo, P., Hu, L.S., Anderson, M.J., Arden, K.C., Blenis, J., Greenberg, M.E., 1999. Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. Cell 96, 857–868.

Burgering, B.M., Kops, G.J., 2002. Cell cycle and death control: long live Forkheads. Trend Biochem. Sci. 27, 352–360.

Cavin Perier, R., Junier, T., Bucher, P., 1998. The Eukaryotic Promoter Database EPD. Nucl. Acids Res. 26, 353–357.

Chan, H.M., La Thangue, N.B., 2001. p300/CBP proteins: HATs for transcriptional bridges and scaffolds. J. Cell Sci. 114, 2363–2373.

Chen, B.K., Leiferman, K.M., Pittelkow, M.R., Overgaard, M.T., Oxvig, C., Conover, C.A., 2003. Localization and regulation of pregnancy-associated plasma protein a expression in healing human skin. J. Clin. Endocrinol. Metab. 88, 4465–4471.

Christian, M., Zhang, X., Schneider-Merck, T., Unterman, T.G., Gellersen, B., White, J.O., Brosens, J.J., 2002. Cyclic AMP-induced Forkhead transcription factor, FKHR, cooperates with CCAAT/enhancer-binding protein beta in differentiating human endometrial stromal cells. J. Biol. Chem. 277, 20825–20832 (e-publication: March 13, 2002).

Davuluri, R., Grosse, I., Zhang, M.Q., 2001. Computational identification of promoters and first exons in the human genome. Nat. Genet. 29, 412–417.

Dijkers, P.F., Medema, R.H., Lammers, J.W., Koenderman, L., Coffer, P.J., 2000. Expression of the pro-apoptotic Bcl-2 family member Bim is regulated by the Forkhead transcription factor FKHR-L1. Curr. Biol. 10, 1201–1204.

Durham, S.K., Suwanichkul, A., Scheimann, A.O., Yee, D., Jackson, J.G., Barr, F.G., Powell, D.R., 1999. FKHR binds the insulin response element in the insulin-like growth factor binding protein-1 promoter. Endocrinology 140, 3140–3146.

Finch, C.E., Ruvkun, G., 2001. The genetics of aging. Annu. Rev. Genomics Hum. Genet. 2, 435–462.

Furuyama, T., Kitayama, K., Yamashita, H., Mori, N., 2003. Forkhead transcription factor FOXO1 (FKHR)-dependent induction of PDK4 gene expression in skeletal muscle during energy deprivation. Biochem. J. 375, 365–371.

Furuyama, T., Nakazawa, T., Nakano, I., Mori, N., 2000. Identification of the differential distribution patterns of mRNAs and consensus binding sequences for mouse DAF-16 homologues. Biochem. J. 349, 629–634.

Hall, R.K., Yamasaki, T., Kucera, T., Waltner-Law, M., O'Brien, R., Granner, D.K., 2000. Regulation of phosphoenolpyruvate carboxykinase and insulin-like growth factor-binding protein-1 gene expression by insulin. The role of winged helix/Forkhead proteins. J. Biol. Chem. 275, 30169–30175.

Higgins, D.G., Thompson, J.D., Gibson, T.J., 1996. Using CLUSTAL for multiple sequence alignments. Meth. Enzymol. 266, 383–402.

Holzenberger, M., Dupont, J., Ducos, B., Leneuve, P., Geloen, A., Even, P.C., Cervera, P., Le Bouc, Y., 2003. IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. Nature 421, 182–187.

Kenyon, C., Chang, J., Gensch, E., Rudner, A., Tabtiang, R., 1993. A *C. elegans* mutant that lives twice as long as wild type. Nature 366, 461–464.

Kops, G.J., de Ruiter, N.D., De Vries-Smits, A.M., Powell, D.R., Bos, J.L., Burgering, B.M., 1999. Direct control of the Forkhead transcription factor AFX by protein kinase B. Nature 398, 630–634.

Korf, I., Flicek, P., Duan, D., Brent, M.R., 2001. Integrating genomic homology into gene structure prediction. Bioinformatics 17 (Suppl. 1), 140–148.

Lawrence, C.E., Altschul, S.F., Bogouski, M.S., Liu, J.S., Neuwald, A.F., Wooten, J.C., 1993. Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. Science 262, 208–214.

Lee, S.S., Kennedy, S., Tolonen, A.C., Ruvkun, G., 2003. DAF-16 target genes that control *C. elegans* life-span and metabolism. Science 300, 644–647.

Matys, V., Fricke, E., Geffers, R., Gossling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A.E., Kel-Margoulis, O.V., et al., 2003. TRANSFAC: transcriptional regulation, from patterns to profiles. Nucl. Acid Res. 31, 374–378.

Medema, R.H., Kops, G.J.P.L., Bos, J.L., Burgering, B.M.T., 2000. AFX-like Forkhead transcription factors mediate cell-cycle regulation by Ras and PKB through $p27^{kip1}$. Nature 404, 782.

Nadal, A., Marrero, P.F., Haro, D., 2002. Down-regulation of the mitochondrial 3-hydroxy-3-methylglutaryl-CoA synthase gene by insulin: the role of the Forkhead transcription factor FKHRL1. Biochem. J. 366, 289–297.

Nakae, J., Kitamura, T., Silver, D.L., Accili, D., 2001. The Forkhead transcription factor Foxo1 (Fkhr) confers insulin sensitivity onto glucose-6-phosphatase expression. J. Clin. Invest. 108, 1359–1367.

Nasrin, N., Ogg, S., Cahill, C.M., Biggs, W., Nui, S., et al., 2000. DAF-16 recruits the CREB-binding protein coactivator complex to the insulin-like growth factor binding protein 1 promoter in HepG2 cells. Proc. Natl. Acad. Sci. U.S.A. 97, 10412.

O'Brien, R.M., Granner, D.K., 1996. Regulation of gene expression by insulin. Physiol. Rev. 76, 1109–1161.

Pearson, W.R., 2000. Flexible sequence similarity searching with the FASTA3 program package. Meth. Mol. Biol. 132, 185–219.

Pierrou, S., Hellqvist, M., Samuelsson, L., Enerback, S., Carlsson, P., 1994. Cloning and characterization of seven human Forkhead proteins: binding site specificity and DNA bending. EMBO J. 13, 5002–5012.

Puig, O., Marr, M.T., Ruhf, M.L., Tjian, R., 2003. Control of cell number by *Drosophila* FOXO: downstream and feedback regulation of the insulin receptor pathway. Genes Dev. 17, 2006–2020 (e-publication: July 13, 2003).

Ramaswamy, S., Nakamura, N., Sansal, I., Bergeron, L., Sellers, W.R., 2002. A novel mechanism of gene regulation and tumor suppression by the transcription factor FKHR. Cancer Cell 2, 81–91.

Streeper, R.S., Svitek, C.A., Chapman, S., Greenbaum, L.E., Taub, R., O'Brien, R.M., 1997. A multicomponent insulin response sequence mediates a strong repression of mouse glucose-6-phosphatase gene transcription by insulin. J. Biol. Chem. 272, 11698–11701.

Suzuki, Y., Yamashita, R., Nakai, K., Sugano, S., 2002. DBTSS: database of human transcriptional start sites and full-length cDNAs. Nucl. Acid Res. 30, 328–331.

Tatar, M., Kopelman, A., Epstein, D., Tu, M.P., Yin, C.M., Garofalo, R.S., 2001. A mutant Drosophila insulin receptor homolog that extends life-span and impairs neuroendocrine function. Science 292, 107–110.

The Gene Ontology Consortium, 2000. Gene ontology: tool for the unification of biology. Nat. Genet. 25, 25–29.

Tiainen, M., Vaahtomeri, K., Ylikorkala, A., Makela, T.P., 2002. Growth arrest by the LKB1 tumor suppressor: induction of p21(WAF1/CIP1). Hum. Mol. Genet. 11, 1497–1504.

von Groote-Bidlingmaier, F., Schmoll, D., Orth, H.M., Joost, H.G., Becker, W., Barthel, A., 2003. DYRK1 is a co-activator of FKHR (FOXO1a)-dependent glucose-6-phosphatase gene expression. Biochem. Biophys. Res. Commun. 300, 764–769.

Zhang, M.Q., 2003. In: Proceedings of the CSHL Quantitative Biology Symposium 68 on Prediction Annotation and Analysis of Human Promoters, CSHL Press, pp. 217–225.