A uniform expression library for the exploration of FOX transcription factor biology

Lavanya Moparthi\textsuperscript{a,b,***}, Stefan Koch\textsuperscript{a,b,*}

\textsuperscript{a} Department of Biomedical and Clinical Sciences (BKV), Linköping University, Linköping, Sweden
\textsuperscript{b} Wallenberg Centre for Molecular Medicine (WCMM), Linköping University, Linköping, Sweden

\textbf{ARTICLE INFO}

\textbf{Keywords:}
- Forkhead box
- Luciferase reporter
- Transcription
- DNA repair
- Ku complex

\textbf{ABSTRACT}

Forkhead box (FOX) family transcription factors play essential roles in development, tissue homeostasis, and disease. Although the biology of several FOX proteins has been studied in depth, it is unclear to what extent these findings apply to even closely related family members, which frequently exert overlapping but non-redundant functions. To help address this question, we have generated a uniform, ready-to-use expression library of all 44 human FOX transcription factors with a convenient peptide tag for parallel screening assays. In addition, we have generated multiple universal forkhead box reporter plasmids, which can be used to monitor the transcriptional activity of most FOX proteins with high fidelity. As a proof-of-principle, we use our plasmid library to identify the DNA repair protein XRCC6/Ku70 as a selective FOX interaction partner and regulator of FOX transcriptional activity. We believe that these tools, which we make available via the Addgene plasmid repository, will considerably expedite the investigation of FOX protein biology.

\section{1. Introduction}

Forkhead box (FOX) domain-containing proteins comprise one of the largest transcription factor families. There are 44 known FOX genes in humans (not counting duplicated and presumed pseudo-genes), which are subdivided into 19 classes, A through S, based on structural similarity of their highly conserved, DNA-binding forkhead box domain (Golson and Kaestner, 2016; Hannenhalli and Kaestner, 2009; Jackson et al., 2010). FOX transcription factors play central roles in essentially all aspects of mammalian physiology, including embryonic development and tissue homeostasis. Consequently, mutations in FOX genes have been linked to numerous developmental defects, immune disorders, and other major diseases such as cancers (Golson and Kaestner, 2016; Myatt and Lam, 2007).

The biology of some FOX family members, such as the oncogene FOXM1 or the longevity-associated FOXO3, has been exceptionally well studied. In contrast, the physiological role of other FOX proteins, especially more recently discovered ones, remains poorly understood (Golson and Kaestner, 2016). Moreover, even though many FOX transcription factors appear to have overlapping functions in, for example, tumorigenesis and cell differentiation, it is unknown in most cases if these biological effects are mediated by shared molecular mechanisms. One of the reasons for this knowledge gap is the lack of convenient tools for the parallel investigation of FOX transcription factors, which means that even closely related FOX family members are often studied in isolation. This issue is compounded by the fact that available clone collections from non-for-profit (e.g., the human ORFeome library (M. G. C. Project Team et al., 2009)) as well as commercial sources are incomplete, prohibitively expensive for many researchers, and require time-consuming subcloning and validation.

To address these problems, we have generated a uniform, fully sequenced collection of expression plasmids covering the main 44 human FOX transcription factors, as well as several other FOX-related genes. In addition, we have cloned a set of forkhead box reporter plasmids that can be used to monitor the transcriptional activity of most FOX proteins. Finally, we demonstrate the usefulness of these tools by identifying Ku70 as a FOX interactor that controls FOX transcriptional activity.
2. Materials and methods

2.1. Plasmid design and construction

FOX genes were cloned by PCR amplification from DNA templates using high-fidelity Q5 polymerase (New England Biolabs, Ipswich, USA) and custom, restriction-site containing primers (see also supplemental Table S1, and (Moparthi et al., 2019)). Amplicons were inserted into a 1x N-Flag pCS2+ vector (a gift from Christof Niehrs, IMB Mainz) by standard restriction cloning. Custom forkhead box reporter plasmids were
generated by replacing the TCF binding motifs of the M50 Super 8x TOPFlash reporter (a gift from Randall Moon, Addgene plasmid # 12, 456) with pre-designed dsDNA fragments (GeneArt Strings, Thermo Fisher Scientific, Waltham, USA; or custom clones, GenScript, Piscataway, USA) containing consecutive forkhead box binding sites. Plasmid inserts were validated by full-coverage DNA sequencing (Eurofins Genomics, Ebersberg, Germany). All plasmids generated in this study and their accompanying sequencing data are available from Addgene (https://www.addgene.org/, plasmid IDs 153109-153161).

2.2. Cell culture and transfection

Authenticated 293T and HCT116 cells were obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany), and maintained in DMEM with 10% fetal calf serum, 2 mM glutamine, and 1% v/v penicillin/streptomycin at 37 °C, 5% CO2. All experiments were performed using low-passage cells from confirmed mycoplasma-free frozen stocks, as determined by analytical qPCR (Eurofins Genomics). Cells were transfected using jetOPTIMUS (Polyplus Transfection, Illkirch, France), according to the supplier’s recommendations.

2.3. Dual luciferase assay

The dual luciferase assay was performed as described previously, with minor modifications (Hampf and Gossen, 2006). In brief, cells were transfected with the Firefly luciferase reporter of interest and a Renilla luciferase control plasmid (pRL-SI, a gift from David Bartel (Addgene plasmid # 12,179)) at a 10:1 ratio, and incubated overnight. The following day, cells were lysed in passive lysis buffer (25 mM Tris, 2 mM DTT, 2 mM EDTA, 10% v/v glycerol, 1% v/v Triton X-100, pH 7.8), and shaken for 10 min. Lysates were transferred to a flat-bottomed, 96-well luminescence assay plate. After addition of Firefly substrate (200 μM D-luciferin in 200 mM Tris-HCl, 15 mM MgSO4, 100 μM EDTA, 1 mM ATP, 25 mM DTT, pH 8.0), the plate was incubated for 2 min at room temperature, and luminescence was recorded on a Spark 10M plate reader (Tecan Life Sciences, Männedorf, Switzerland). Then, Renilla substrate was added to the plate (4 μM coelenterazine-h in 500 mM NaCl, 500 mM Na2SO4, 10 mM NaOAc, 15 mM EDTA, 25 mM sodium pyrophosphate, 50 μM phenyl-benzothiazole, pH 5.0), and luminescence was measured immediately. All Firefly values were divided by Renilla luminescence in the same sample for normalization.

2.4. Immunoblotting

Cells were lysed in ice-cold RIPA lysis buffer with 1x protease inhibitor cocktail (Thermo Fisher), cleared by brief centrifugation, and boiled in Laemmli sample buffer with 50 mM DTT. Samples were separated on 10% polyacrylamide gels (Bio-Rad, Hercules, USA), transferred to nitrocellulose membranes, and blocked in TBS-based blocking buffer. Then, membranes were incubated with anti-Flag M2 (F3165, Sigma-Aldrich, St. Louis, USA) and anti-HSP70 antibodies (4873, Cell Signaling Technology, Danvers, USA), washed, and incubated with near infrared fluorophore-labeled secondary antibodies. Images were acquired on an Odyssey CLx scanner using Image Studio software (LI-COR, Lincoln, USA). Consumables and secondary antibodies were obtained from LI-COR.

2.5. Immunocytochemistry

Cells were fixed with 4% w/v paraformaldehyde, permeabilized with 0.25% v/v Triton X-100, and blocked with 5% w/v bovine serum albumin and 10% v/v donkey serum (all in PBS). Then, cells were incubated with anti-Flag M2 antibody, washed, and incubated with a DyLight 488-labeled secondary antibody (Thermo Fisher). Samples were mounted in Prolong Glass Antifade with NucBlue (Thermo Fisher). Images were acquired on an LSM700 confocal laser scanning microscope controlled through ZEN software (Carl Zeiss, Jena, Germany), and post-processed using ImageJ 1.52p (NIH, Bethesda, USA).

2.6. Co-immunoprecipitation

Cells were co-transfected with either FOX transcription factor plasmids or empty vector control and Ku70 plasmid (pCE-HA-Ku70-siR-WT, a gift from Patrick Calos, Addgene plasmid # 82,329). Whole-cell lysates were prepared by lysis cells with ice-cold 1% NP-40 in TBS pH 7.4 with 1x protease inhibitor cocktail. Lysates were first pre-cleared with 5 μl of protein A/G PLUS agarose (sc-2003, Santa Cruz Biotechnology, Dallas, USA) and then incubated with 10 μl of Flag M2 Affinity gel (A2220, Sigma-Aldrich, St. Louis, USA) overnight at 4°C. The protein complexes were eluted with 4x Laemmli buffer and resolved by SDS-PAGE for immunoblotting with anti-Flag and anti-HA antibodies (NB600-363, Novus Biologicals, Centennial, USA).

2.7. Data analysis and presentation

Data were plotted in Prism 8 (Graphpad Software, San Diego, USA), and are shown as mean ± standard deviation of three biological replicates. Statistical analyses of the luciferase reporter data were done using Dunnett’s post-hoc test following ANOVA, and results can be found in supplemental Table S2.

3. Results

3.1. Generation and validation of the human FOX collection

To obtain a set of expression constructs covering the core 44 human FOX transcription factors, we isolated full-length cDNAs from in-house epithelial cell cDNA pools, DNASU and Addgene plasmid repository clones (Seiler et al., 2014; Seoane et al., 2004), and commercial vectors by PCR amplification (supplemental Table S1). The amplicons were inserted into a pCS2+ plasmid vector with a single N-terminal Flag/-DYKDDDDK tag by restriction cloning, omitting the starting methionine (Fig. 1 A and B). In this manner, we obtained ready-to-use expression plasmids for the main 44 FOX transcription factors, as well as the presumed pseudo-gene FOXO3B and the FOXD4-related genes D4L1, D4L3, D4L5, and D4L6. Full coverage sequencing confirmed greater than 98% amino acid identity of all constructs with the reference protein sequence, and in most cases, deviations from the reference were already present in the template source (supplemental Table S1). We selected pCS2+ as a backbone because of its widespread use as a high-expression vector, the option to perform in vitro transcription/translation of the insert, and the compatibility of its polylinkers with other commonly used vectors, including pcDNA3 family plasmids. Similarly, we chose the 1x Flag tag to strike a balance between maximal convenience in subsequent applications, and minimal interference with protein folding and activity. Moreover, the peptide tag can be removed by enterokinase cleavage, if required. Nonetheless, any peptide tag can potentially interfere with protein function, and we advise investigators to confirm results obtained with these constructs using the native protein or other means.

We initially validated the human FOX collection by immunoblotting after overexpression in 293T cells, which showed that all constructs were expressed and could be detected at or around their calculated molecular weight (Fig. 1C). Of note, protein levels of some FOX family members such as FOXM1 and FOXN2 were low in these assays, possibly due to constitutive proteasomal degradation (Chen et al., 2016; Ma et al., 2018). Next, we assessed the subcellular localization of all constructs by immunofluorescence microscopy (Fig. 1D). As expected, most FOX proteins predominantly localized to the nucleus. Cytosolic staining was observed for class O transcription factors (FOXO1, 3(B), 4, and 6), which are known to shuttle between the cytoplasm and nucleus in a...
stimulus-dependent manner (Greer and Brunet, 2005; Li et al., 2015). Of note, the subcellular localization of FOX3 and FOXR2 was ambiguous, since the proteins were detected both in the cytosol and nucleus (Fig. 1E). Whether this observation indicates active shuttling of these proteins is unclear at this point. Collectively, these results demonstrate the successful generation of a uniform expression library comprising all 44 human FOX transcription factors, as well as additional FOX genes.

### 3.2. Functional validation of the human FOX collection using reporter constructs

FOX proteins act as transcriptional activators or repressors, and accordingly, the activity of some FOX transcription factors has been studied in vitro using forkhead box reporter plasmids (Wang et al., 2018; Zanella et al., 2009). To validate the functional activity of our FOX proteins, we therefore tested their ability to regulate luciferase-based
reporter constructs in 293T cells (Fig. 2A). We first tested an incomplete, earlier build of our collection (Moparthi et al., 2019) using the FOXO3 reporter FHRE-luc (Brunet et al., 1999), which contains a short fragment of the FAS ligand promoter with three forkhead box binding sites. This reporter was activated by FOXO3 as well as some other tested FOX proteins, including FOXC1 and FOXF1 (Fig. 2B). In contrast, none of these transcription factors activated the YAP1 reporter 8x GTIIC-luc (Dupont et al., 2011), confirming the specificity of the response. However, luciferase expression was generally low, and the results were in part inconsistent with earlier reports on, for example, FOXA1 and FOXL1 (Nakada et al., 2006; Wang et al., 2018).

We therefore generated three additional reporters in a common pGL3 backbone: 6x DBE-luc, based on an earlier construct containing six consecutive repeats of the daf-16 family protein-binding element (Zanella et al., 2009); 10x UFR-luc, a universal forkhead box reporter with ten motifs covering known FOX transcription factor binding matrices curated in the JASPAR 2020 database (Fornes et al., 2019); and 33x μUFR-luc, with a mix of minimal forkhead box binding sites separated by short spacer sequences. We tested these plasmids using FOX clones representing all 19 classes, and observed strong positive or negative regulation of at least one of the reporters by all included transcription factors except FOXS1 (Fig. 2C). With few exceptions the results were generally comparable across reporters, and consistent with earlier studies.

Because 10x UFR-luc showed the most robust response in the aforementioned assays, we used this reporter to test the activity of all FOX transcription factors in our collection. We observed that 28 of the 44 core FOX proteins regulated this reporter at least two-fold compared to the empty vector control, with most of them acting as transcriptional activators except for FOXL1 and FOXQ1 (Fig. 2D). The transcriptional response was mostly consistent within the different FOX subfamilies, with notable exceptions, namely classes B, L, and N. FOXO3B, which was recently identified as a protein-coding gene (Santo and Paik, 2018), also displayed approximately two-fold activation in this assay, whereas the FOXD4L clones in our collection had minimal transcriptional activity (supplemental Figure 1). Collectively, these data suggest that most clones in our human FOX collection are transcriptionally active, and that their activity can be monitored using universal forkhead box reporter plasmids.

3.3. XRCC6/Ku70 is a selective FOX interactor

We next assessed the usefulness of our plasmid library for the parallel investigation of FOX transcription factor biology. Jin et al. recently reported that FOXL2 interacts with the Ku complex proteins Ku70/Ku80 to control DNA double-strand break repair (Jin et al., 2020). Based on the observation that FOXL2 interacts with Ku70/80 via its conserved forkhead domain, the authors proposed that regulation of Ku complex activity and thus DNA repair may be a common function of FOX proteins. We therefore tested the association of Ku70 with multiple putative FOX
family interactors (Li et al., 2015) by co-immunoprecipitation (Fig. 3A and B). Whereas Ku70 readily co-precipitated with FOX1L and FOXL2, we observed no interaction with FOXA1 and FOXA2, and only marginal interaction with FOXC1 and FOXC2 (Fig. 3B). Additionally, using the 10x UFR-luc reporter, we observed that overexpression of Ku70 resulted in strongly reduced FOX transcriptional activity (Fig. 3C). However, concomitant expression of FOXL1 and FOXL2, and to a lesser extent FOXC2, de-repressed luciferase reporter activity. Collectively, our results suggest that Ku complex proteins interact only with select FOX proteins, and that Ku70 sequestration by these specific interactors may increase the transcriptional activity of other FOX family members.

4. Discussion

Here we report the generation of a ready-to-use plasmid library for the study of FOX transcription factor biology. In conjunction with our universal forkhead box reporter plasmids, this collection can be used to rapidly explore the function of multiple FOX proteins in parallel.

FOX transcription factors are involved in essentially all aspects of mammalian physiology, but in many cases their specific function and mode of action are poorly understood (Benayoun et al., 2011; Golson and Kaestner, 2016). Thus, the parallel investigation of multiple FOX proteins offers an opportunity to uncover unique as well as shared features of individual FOX family members. In an earlier study, Li and colleagues used large-scale proteomics to identify binding partners of 37 FOX proteins (Li et al., 2015). These authors observed that FOX transcription factors recruit distinct sets of interactors, with limited overlap even between FOX proteins within the same class. The study demonstrated the benefits of investigating multiple transcription factors simultaneously. At the same time, however, it cautions against extrapolating the function of FOX genes from related or orthologous family members. Indeed, even though some FOX proteins such as FOX1/2 and FOXC1/2 appear to be functionally redundant (Ormestad et al., 2006; Nakagawa, S., Gisselbrecht, S.S., Rogers, J.M., et al., 2013. DNA-binding specificity differences between FOX transcription factors in the same class, including FOXA1/3, FOXB1/2, and FOXO1/3 (Matsukawa et al., 2009; Moparthi et al., 2019; Motallebipour et al., 2009).

Most FOX family members act as transcription activators, with notable exceptions such as the known repressors FOXC1 and FOXO4 (Yao et al., 2001), FOXL1 (Nakada et al., 2006), and FOXP2 (Shu et al., 2001). However, the relative activity of FOX transcription factors is unknown. Given that FOX proteins share a highly conserved DNA binding domain, and thus homologous binding motifs (Nakagawa et al., 2013), we reasoned that universal luciferase reporters can be used to directly compare their activity. Indeed, we identified numerous FOX family members whose transcriptional strength appears to be much higher than FOX proteins even within the same class, including FOXA3, FOXL1, and FOXN1/4. Moreover, the data suggest that additional FOX proteins, namely FOXD1 and FOXI1, act as transcriptional repressors. The function of individual FOX transcription factors is shaped in large part by their specific protein interactors (Golson and Kaestner, 2016; Myatt and Lam, 2007). Thus, further exploration of the FOX family-associated proteome will likely reveal common mechanisms of transcriptional regulation.

With regard to the recently reported link between FOXL2 and DNA damage repair (Jin et al., 2020), our proof-of-concept data argue against a general role of FOX proteins in Ku complex regulation, contrary to what has been suggested. However, our data provide several interesting starting points for further investigation. Firstly, we show that FOXL1 also binds Ku70 with high affinity, suggesting that FOXL1/2 may act redundantly in double-strand break repair. Secondly, the fact that FOXA1/2 and FOXC1/2 interact with Ku70 with considerably lower affinity despite their highly homologous forkhead boxes may provide important structural clues for the basis of FOXL2/Ku70 interaction.

Lastly, our data suggest that Ku70 inhibits FOX transcriptional activity, consistent with earlier observations (Brenkman et al., 2010). It is thus conceivable that sequestration of Ku complex proteins by select FOX proteins creates a permissive environment for gene transcription in the absence of DNA damage. We are confident that the toolset we describe here will be helpful in addressing these and other questions in FOX transcription factor biology.

Declaration of competing interest

The authors declare no conflicts of interest.

Acknowledgements

We would like to thank all researchers who have made resources and data available to the public. The authors thank Dr. Sergio Acerboni (COS, Heidelberg) for advice on the dual luciferase assay. SK is a Wallenberg Molecular Medicine Fellow, and receives financial support from the Knut and Alice Wallenberg Foundation. LM is supported by project grants from the Rotary Club Borgholm, and the Lions Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.diff.2020.08.002.

References


