

The *in vitro* PORE: An Improved Technique to Pull Out Regulatory Elements

Alpa Sidhu¹, Patrick J. Miller, and Andrew D. Hollenbach*

Department of Genetics, Louisiana State University Health Sciences Center,
533 Bolivar Street, CSRB 6th floor, New Orleans, LA 70112

*Corresponding author: Andrew D. Hollenbach, phone: (504) 568-2431, FAX: (504) 568-8500, e-mail: aholle@lsuhsc.edu

¹Present address: Detroit Medical Center, Children's Hospital of Michigan, Pediatric Education Office, 2901 Beaubien Street, Detroit, MI 48201, USA

The regulation of gene expression through the interaction of a specific DNA binding protein (DNA-BP) and its associated genomic regulatory element drives many biological processes. Alterations in this normal regulation, either through mutations in the regulatory element or within the DNA-BP itself, many times contribute to the development of the diseased state. However, despite the knowledge of the role that DNA-BPs play in normal biological processes and in the development of disease, more often than not the genomic regulatory elements and associated genes whose expression is altered by a DNA-BP are not known. **Therefore, the ability to perform a non-biased global genomic screen would greatly assist in the identification of regulatory elements that control the expression of genes thereby providing a better understanding of the role that a specific DNA-BP plays in various biological processes.**

At present, chromatin immunoprecipitation (ChIP) and its variants are the favored methods to identify genomic regulatory elements¹. Although these techniques are considered to be standard methods for examining protein-DNA interactions, they have several significant drawbacks. First, some DNA-BPs are not amenable to ChIP analysis due to the lack of a suitable antibody or because the epitope is inaccessible when complexed to DNA. Alternatively, for certain DNA-BPs, the ChIP process is technically difficult and not readily reproducible²⁻⁴. Second, ChIP

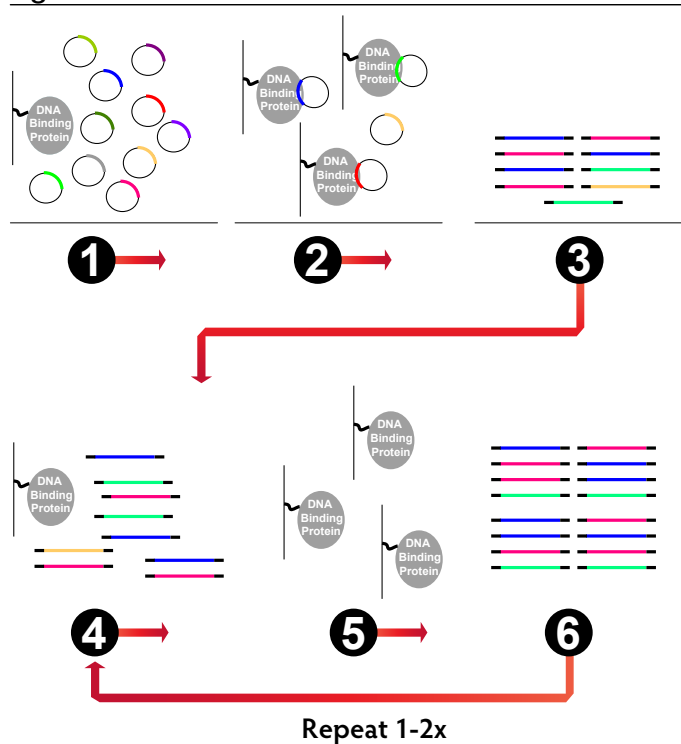
analysis requires extensive cellular manipulations that must be optimized for each individual DNA-BP, an optimization process that may sometimes take weeks or longer. Third, because ChIP is performed in a specific cell- or tissue-type, it is limited to identifying genes that are expressed only in a particular cellular context. This analysis would therefore limit the ability to identify novel regulatory elements for ubiquitous DNA-BPs or DNA-BPs with broad biological activities in multiple cell- or tissue-types. Finally, analysis of the DNA fragments using some of the ChIP variants (ChIP-on-Chip) requires expensive microarray chips, experienced personnel, and is limited by the regions of genomic DNA present on the microarray chip, making the process prohibitive for certain labs. A simple, cost-effective, rapid, global method that can circumvent these limitations would greatly facilitate the isolation of genomic regulatory elements.

To this end, we have developed an improved, cost-effective, rapid *in vitro* method to Pull Out Regulatory Elements (*in vitro* PORE)⁵ that circumvents the significant drawbacks and limitations of the present ChIP technology. The *in vitro* PORE is a modification of the standard Systematic Evolution of Ligands by Exponential enrichment (SELEX)⁶. Standard SELEX uses equilibrium binding of a synthesized oligonucleotide library *in vitro* to partition DNA followed by PCR amplification of the bound species. Multiple rounds of binding and amplification

result in a population significantly enriched for oligonucleotides that contain DNA sequences specifically recognized by a known DNA-BP. We have taken the basic SELEX concept and expanded it for application to genomic DNA by first having Lucigen Corp. create a stable mouse genomic library in their transcription-free pSMART[®]-LC-Kan vector. This genomic library has >95% stability and contains genomic DNA inserts ranging in size from 0.65 – 2.5 kb resulting in a 2.5 over-representation of the entire mouse genome (see Sidhu, et al., for a complete analysis of the stable genomic library⁵). Use of our stable genomic library has several significant improvements over previous reports of similar techniques. First, unlike previous reports, our genomic library does not suffer from the inherent instability of genomic DNA fragments when cloned into standard vectors⁷. Second, the larger insert size of the library is, on average, nearly 10-fold larger than the small size of insert (100 – 300 bp) previously reported⁸⁻¹¹, allowing the use of a significantly smaller amount of starting material. Finally, the use of a stable genomic library negates the introduction of unnecessary levels of complexity and uncertainty through the inclusion of additional ligation steps or complicated PCR amplifications^{12,13}.

The *in vitro* PORE, as illustrated in Figure 1 and outlined

Figure 1



in Figure 2, consists of first bacterially expressing and purifying a tagged-DBP using standard techniques such as Lucigen's Expresso[®] T7 Cloning and Expression System (Lucigen, Corp., cat. no. 49000-1). Once immobilized on the resin the purified protein is incubated with a small aliquot ($\approx 1 - 10 \mu\text{g}$) of the stable genomic library. After washing away any unbound DNA species, the bound DNA is eluted from the resin and the genomic insert is PCR amplified using primers specific for the pSMART-LC-Kan vector, as previously described⁵. The amplified fragments are subsequently gel purified and used for two to three additional rounds of binding and amplification. The final amplified product is cloned into a standard PCR cloning vector, such as Lucigen's GC Cloning Kit (Lucigen, Corp., catalog #40731-1), after which DNA is isolated from individual colonies, sequenced, and analyzed using available genomic DNA bioinformatics databases.

We first demonstrated that we could use the *in vitro* PORE to selectively isolate a known genomic regulatory element when the element is present in the context of our stable genomic library (Figure 3)⁵. We utilized the promoter region of ceruloplasmin (Cp), a known transcriptional target of the DNA-BP FOXO1¹⁴, alone or diluted 1:10 and 1:1000 with our stable mouse genomic library along with bacterially expressed, purified, and immobilized GST-FOXO1 or GST alone (negative control) in the *in vitro* PORE. We observed a FOXO1-specific binding and amplification of DNA of approximately 0.75 – 2.5 kb, consistent with the insert sizes of the stable genomic library, confirming that our method efficiently and accurately amplifies a complex genomic library containing large inserts in a DNA-BP-dependent manner (Figure 3A). To confirm the presence of the Cp promoter in our selected and amplified products, aliquots of the gel extracted DNA were used to amplify the Cp promoter with primers specific for this DNA sequence. We observed an amplification of the Cp promoter in the presence of GST-FOXO1 with all samples, an amplification that was greatly reduced or non-existent in the GST negative control (Figure 3B). Therefore, this result demonstrates that a known DNA-BP is capable of isolating a known genomic regulatory element when present in the context of a stable genomic library.

We then used the *in vitro* PORE process to successfully identify genomic DNA sequences and their associated genes regulated by the transcription factor FOXO1⁵ a protein that is relatively intractable to ChIP analysis. FOXO1, a member of the FOXO family of winged-helix transcription factors, is a ubiquitously expressed DNA-BP and as such has been demonstrated to be a key transcription factor in various cellular processes including the regulation of glucose homeostasis, apoptosis, cell cycle control, muscle differentiation, the oxidative stress response, and angiogenesis¹⁵. Due to the universal actions of FOXO1, standard ChIP analysis would limit identification of regulated genes to a cell-dependent function of FOXO1 rather than provide a global analysis of FOXO1-regulated genes. In contrast, our *in vitro* PORE screen identified genomic regulatory elements, isolated multiple times from three independent replicates of the entire *in vitro* PORE procedure, that are associated with genes whose functions are consistent with many

of the varied FOXO1 biological functions including apoptosis, cell cycle regulation, muscle differentiation, T-cell differentiation, erythroid differentiation, and oxidative stress (Tables 1 and 2)⁵. Providing even more support to the validity of the results, 31% of the isolated sequences lie in proximity to genes whose expression is regulated in a FOXO1-dependent manner (Table 1) with an additional 38% of the isolated sequences being present near genes whose functions are consistent with the known biological activities of FOXO1 (Table 2)⁵.

Further enhancing the power of this technique, the *in vitro* PORE has many benefits and advantages over ChIP and its associated analyses, as described in Table 3. First, the *in vitro* PORE requires no extensive optimizations or cellular manipulations. The screen can be initiated using standard

Figure 2: The *in vitro* PORE process and time frame.

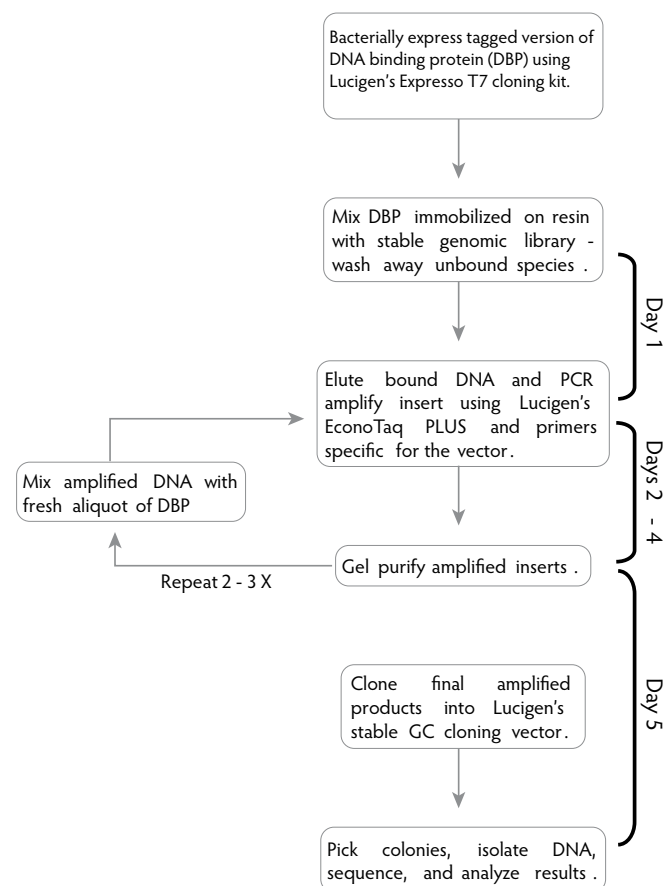
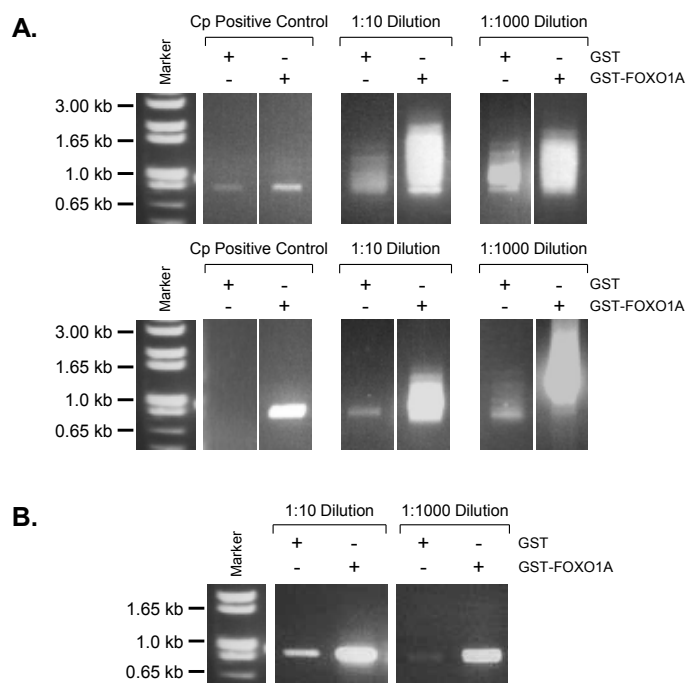


Figure3



Binding and amplification of a known FOXO1 regulatory sequence from a genomic library pool. **(A)** Agarose gel electrophoresis demonstrating two rounds (top panel – round #1; bottom panel – round #2) of binding and amplification using the Ceruloplasmin (Cp) promoter region alone (positive control), and with the Cp promoter diluted 1:10 and 1:1000 with the mouse genomic library. Binding reactions with bacterially expressed GST were carried out in parallel as a negative control. **(B)** To confirm the presence of the Cp promoter region in the second round of selection, 2 μ l of the gel purified eluate was PCR amplified using primers specific for the Cp promoter region and visualized on a 1% agarose gel.

DNA-binding buffers as soon as the tagged-DBP is expressed, purified, and immobilized on a chromatography resin. Second, PORE uses basic techniques, such as equilibrium binding and PCR amplification, which are standard procedure in many labs, removing the necessity of introducing more complex cellular and molecular biological techniques. Third, because the expression and immobilization step is the rate-limiting step and the PORE uses basic laboratory techniques, the *in vitro* PORE is rapid. The researcher can send isolated DNA from the final round of PORE for sequencing within a week of obtaining immobilized protein (Figure 2), a time period in which cells would still be expanding in order to obtain the large amounts of material needed for ChIP. Fourth, *in vitro* PORE uses a whole-genome library. Therefore, the screen is not limited to a single cell- or tissue-type thereby enhancing the unbiased nature of the screen. Finally, the analysis of the final round of PORE simply involves sequencing of isolated DNA and searching of standard bioinformatics databases thereby removing the necessity of purchasing expensive microarray chips or requiring the assistance of experienced personnel.

In conclusion, the *in vitro* PORE provides a rapid, inexpensive, and highly accessible technique for the non-biased and global isolation of genomic regulatory elements directly bound by a DBP. Although originally engineered and developed to identify genomic DNA sequences bound by “difficult-to-ChIP” DNA-BPs, this technique can be used with any DNA-BP of interest. Finally, except in cases where a DNA-BP is intractable to ChIP analysis, the *in vitro* PORE is not meant to replace the standard ChIP methodology. Instead, the *in vitro* PORE is a powerful, non-biased and global technique that like the yeast two-hybrid, yeast one-hybrid, and other techniques used for global screening, acts as the initial step for the identification of regulatory elements and genes whose expression is controlled by a DBP or can serve as an independent confirmation of other regulatory element isolation and identification techniques.

Table 1: Known FOXO1 regulated genes within proximity of the isolated genomic regulatory sequences⁵.

| Gene ^a | FOXO1-related function | Replicate | | | Total |
|-------------------|------------------------------------|-----------------|----|----|-----------------------|
| | | #1 ^b | #2 | #3 | |
| Tnfrsf5 | Pro-apoptotic | 42 | 9 | 2 | 53 (20%) ^c |
| Cyclin D1 | Cell cycle regulation | 20 | 7 | | 27 (10%) |
| BAMBI | TGF- β signaling/Development | 1 | | | 1 |
| Zwint-1 | Cell cycle regulation | | 1 | | 1 |

^a Tnfrsf5 – tumor necrosis factor receptor super family 5; BAMBI – BMP and activin membrane bound inhibitor; Zwint-1 – ZW10 interacting protein 1.

^b The numbers for each replicate indicate the frequency with which each gene was isolated from 118 (replicate #1), 94 (replicate #2), and 50 (replicate #3) individual clones.

^c The percentage of the total number of clones from all three replicates that were comprised of each isolated gene.

Table 2: Putative FOXO1 regulated genes within proximity of the isolated genomic regulatory sequences⁵.

| Gene ^a | FOXO1-related function | Replicate #1 ^b | Replicate #2 | Replicate #3 | Total |
|-------------------|--|---------------------------|--------------|--------------|-----------------------|
| IGFBP9 | Muscle differentiation | 20 | 2 | 7 | 29 (11%) ^c |
| Col4a2 | Apoptosis, Inhibition of Akt | | 14 | 6 | 20 (8%) |
| RACK1 | TNF- α signaling/ apoptosis, Cell cycle regulation | 12 | 4 | 2 | 18 (7%) |
| MKL1 | Muscle differentiation | 1 | 1 | 3 | 5 (2%) |
| T3JAM | TNF- α signaling/ apoptosis | | 2 | 3 | 5 (2%) |
| Akap6 | Muscle differentiation | 2 | 1 | | 3 (1%) |
| Dss1 | Cellular proliferation | | 2 | 1 | 3 (1%) |
| Phlpp1 | Apoptosis, Inhibition of Akt | | 2 | | 2 |
| epb4.1l3 | Erythroid differentiation | | 2 | | 2 |
| NMS | Oxidative stress/circadian clock | 1 | | 1 | 2 |
| Map3K7IP1 | Regulation of FOXO1 activity | | 2 | | 2 |
| Kir3.1 | Muscle differentiation/ metabolism | | | 2 | 2 |
| Ctla2a | T-cell homeostasis | 1 | | | 1 |
| EGR-3 | TNF- α signaling/ apoptosis | | | 1 | 1 |
| MAGI-1 | Apoptosis | | | 1 | 1 |
| Yap1 | Apoptosis | | | 1 | 1 |
| Nfyb | Cell cycle regulation | | | 1 | 1 |

^a IGFBP9 – insulin like growth factor binding protein 9; Col4a2 – procollagen type IV alpha 2; RACK1 – receptor for activated kinase; MKL1 – megakaryoblastic leukemia 1; T3JAM – TRAF3-interacting JNK-activating modulator; Akap6 – A kinase (PKA) anchor protein 6; Dss1 – split hand/foot deleted gene 1; Phlpp1 – PH domain and leucine rich repeat protein phosphatase 1; epb4.1l3 – erythrocyte protein band 4.1-like 3; NMS – neuromedin S; Map3K7IP1 – MAPKKK7 interacting protein 1; Kir3.1 – potassium inwardly rectifying channel J3; Ctla2a – cytotoxic T-lymphocyte associated protein 2 alpha; EGR-3 – early growth response 3; MAGI-1 – membrane associated guanylate kinase; Yap1 – yes associated protein 1; Nfyb – nuclear transcription factor Y beta.

^b The numbers for each replicate indicate the frequency with which each gene was isolated from 118 (replicate #1), 94 (replicate #2), and 50 (replicate #3) individual clones.

^c The percentage of the total number of clones from all three replicates that were comprised of each isolated gene.

Table 3: Advantages of the *in vitro* PORE process.

| | <i>in vitro</i> PORE | ChIP and Variants |
|------------------------------------|----------------------|-------------------|
| No extensive optimization | ✓ | |
| No complex cellular manipulations | ✓ | |
| Uses basic, standard techniques | ✓ | |
| Rapid (\approx 1 week) | ✓ | |
| Not limited to cell or tissue type | ✓ | |
| Inexpensive | ✓ | |

References

1. Hudson, M.E. & Snyder, M. High-throughput methods of regulatory element discovery. *Biotechniques* **41**, 673, 675, 677 passim (2006).
2. Matsuzaki, H. et al. Acetylation of Foxo1 alters its DNA-binding ability and sensitivity to phosphorylation. *Proc Natl Acad Sci U S A* **102**, 11278-83 (2005).
3. Hall, R.K., Wang, X.L., George, L., Koch, S.R. & Granner, D.K. Insulin represses phosphoenolpyruvate carboxykinase gene transcription by causing the rapid disruption of an active transcription complex: a potential epigenetic effect. *Mol Endocrinol* **21**, 550-63 (2007).
4. Motta, M.C. et al. Mammalian SIRT1 represses forkhead transcription factors. *Cell* **116**, 551-63 (2004).
5. Sidhu, A., Miller, P.J. & Hollenbach, A.D. Isolation of putative FOXO1 genomic elements using an improved *in vitro* technique to isolate genomic regulatory sequences. *Gene* **458**, 45-53 (2010).
6. Bouvet, P. Determination of nucleic acid recognition sequences by SELEX. *Methods Mol Biol* **148**, 603-10 (2001).
7. Godiska, R., Patterson, M., Schoenfeld, T. & Mead, D.A. Beyond pUC: Vectors for cloning Unstable DNA. 1 - 3 (Jones and Bartlett Publishers, 2005).
8. Singer, B.S., Shtatland, T., Brown, D. & Gold, L. Libraries for genomic SELEX. *Nucleic Acids Res* **25**, 781-6 (1997).
9. Ferrieres, L., Francez-Charlot, A., Gouzy, J., Rouille, S. & Kahn, D. FixJ-regulated genes evolved through promoter duplication in *Sinorhizobium meliloti*. *Microbiology* **150**, 2335-45 (2004).
10. Wen, J.D. & Gray, D.M. Selection of genomic sequences that bind tightly to Ff gene 5 protein: primer-free genomic SELEX. *Nucleic Acids Res* **32**, e182 (2004).
11. Itou, H. et al. The CGL2612 protein from *Corynebacterium glutamicum* is a drug resistance-related transcriptional repressor: structural and functional analysis of a newly identified transcription factor from genomic DNA analysis. *J Biol Chem* **280**, 38711-9 (2005).
12. Shostak, Y., Van Gilst, M.R., Antebi, A. & Yamamoto, K.R. Identification of *C. elegans* DAF-12-binding sites, response elements, and target genes. *Genes Dev* **18**, 2529-44 (2004).
13. Williams, R. et al. Amplification of complex gene libraries by emulsion PCR. *Nat Methods* **3**, 545-50 (2006).
14. Sidhu, A., Miller, P.J. & Hollenbach, A.D. FOXO1 stimulates ceruloplasmin promoter activity in human hepatoma cells treated with IL-6. *Biochem Biophys Res Commun* **404**, 963-7 (2011).
15. Salih, D.A. & Brunet, A. FoxO transcription factors in the maintenance of cellular homeostasis during aging. *Curr Opin Cell Biol* **20**, 126-36 (2008).