

Identification of DNA-associated

Proteins by Sequence-Specific Capture and Mass Spectrometry

Hector Guillen Ahlers¹, Shama Mirza¹, Sherry Zhang¹, Regina Cole¹, Milena Zelembaba¹, Michael Zickus¹, Marla Chesnik¹, Cheng-Hsien Wu², Siyuan Chen², Yuan Yuan², Gloria Kreitinger², Mark Scalf², Michael R. Shortreed², Lisa A. Cirillo¹, Lloyd M. Smith², Michael Olivier¹

Wisconsin Center of Excellence in Genomics Science, 1 Medical College of Wisconsin, 2 University of Wisconsin Madison



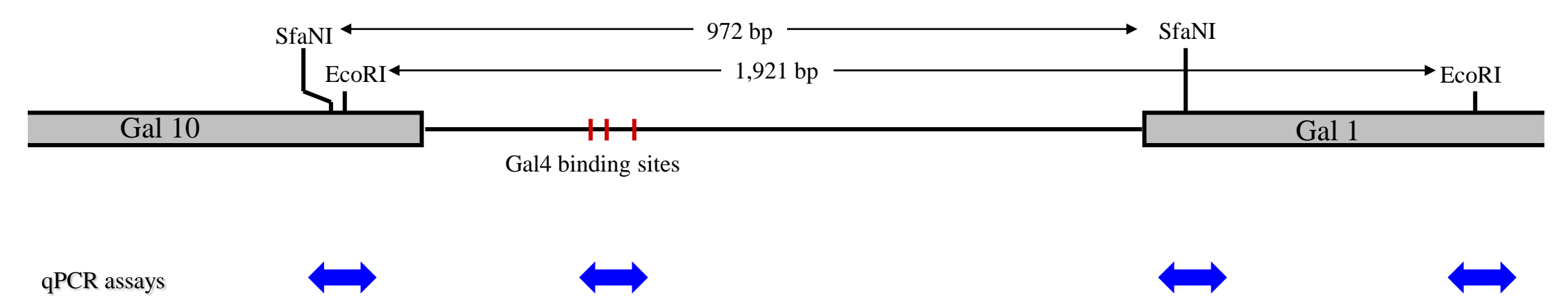
Introduction

Recent advances in genomics and proteomics have brought us closer to reaching a detailed and comprehensive understanding of our genome and how it is regulated. Numerous proteins mediate DNA stability, control its activity, and regulate transcription of the genetic information. However, currently no technologies exist that allow the dissection of these protein-DNA interactions in a comprehensive global manner, and examine alterations in disease. To overcome this challenge, we report on the development of an entirely novel technology. Unlike ChIP-chip methodology, where DNA sequences that interact with individual known proteins are characterized, the Wisconsin Center for Excellence in Genomics Science (CEGS) utilizes an oligonucleotide capture technology to isolate targets of interest in a sequence-specific manner in order to analyze protein complexes attached to these regions. Two model systems are used for this purpose:

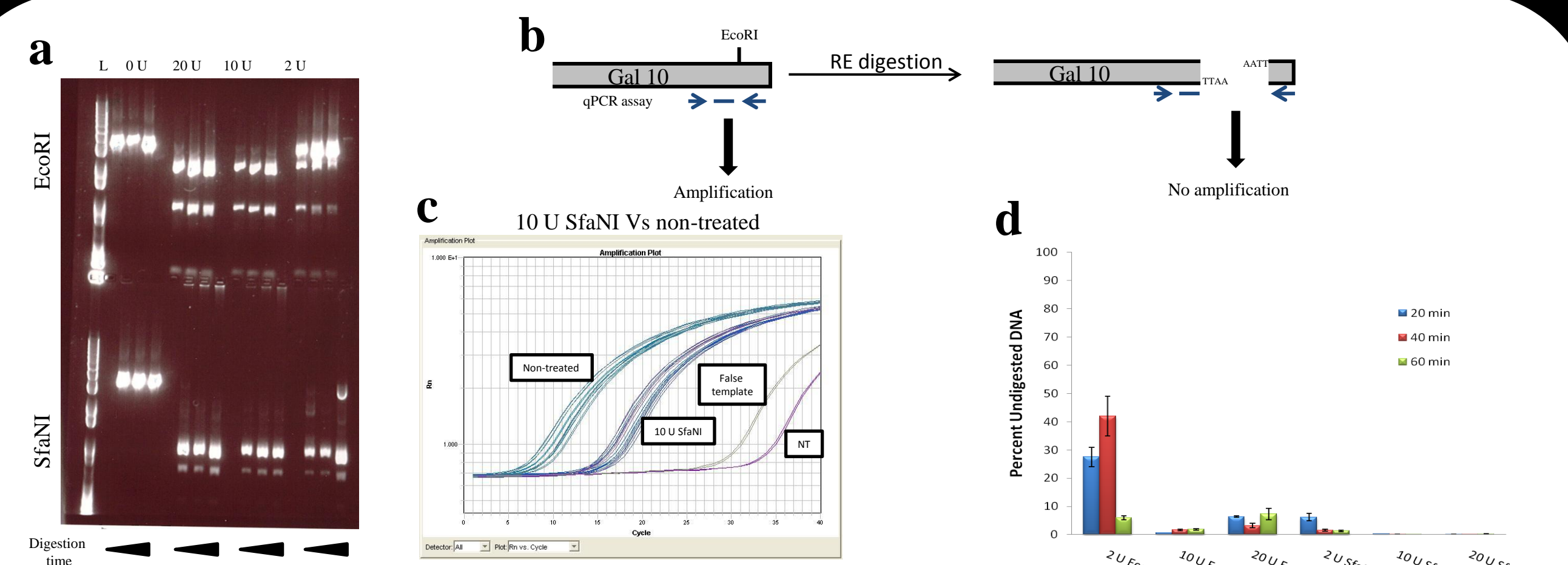
➤ The mouse insulin-like growth factor-binding protein 1 (IGFBP1) promoter region was used as an *in vitro* model system. Hybridization was optimized to sequester PCR-products containing an exposed single-stranded overhang. After on-chip protease digestion, FoxO1 binding to the DNA sequence was detected by tandem mass spectrometry using an LTQ XL mass spectrometer.

➤ The yeast upstream activator sequence for Gal (UAS_{GAL}) was used as an *in vivo* model. Conditions were developed to capture the region of interest directly from cell lysates.

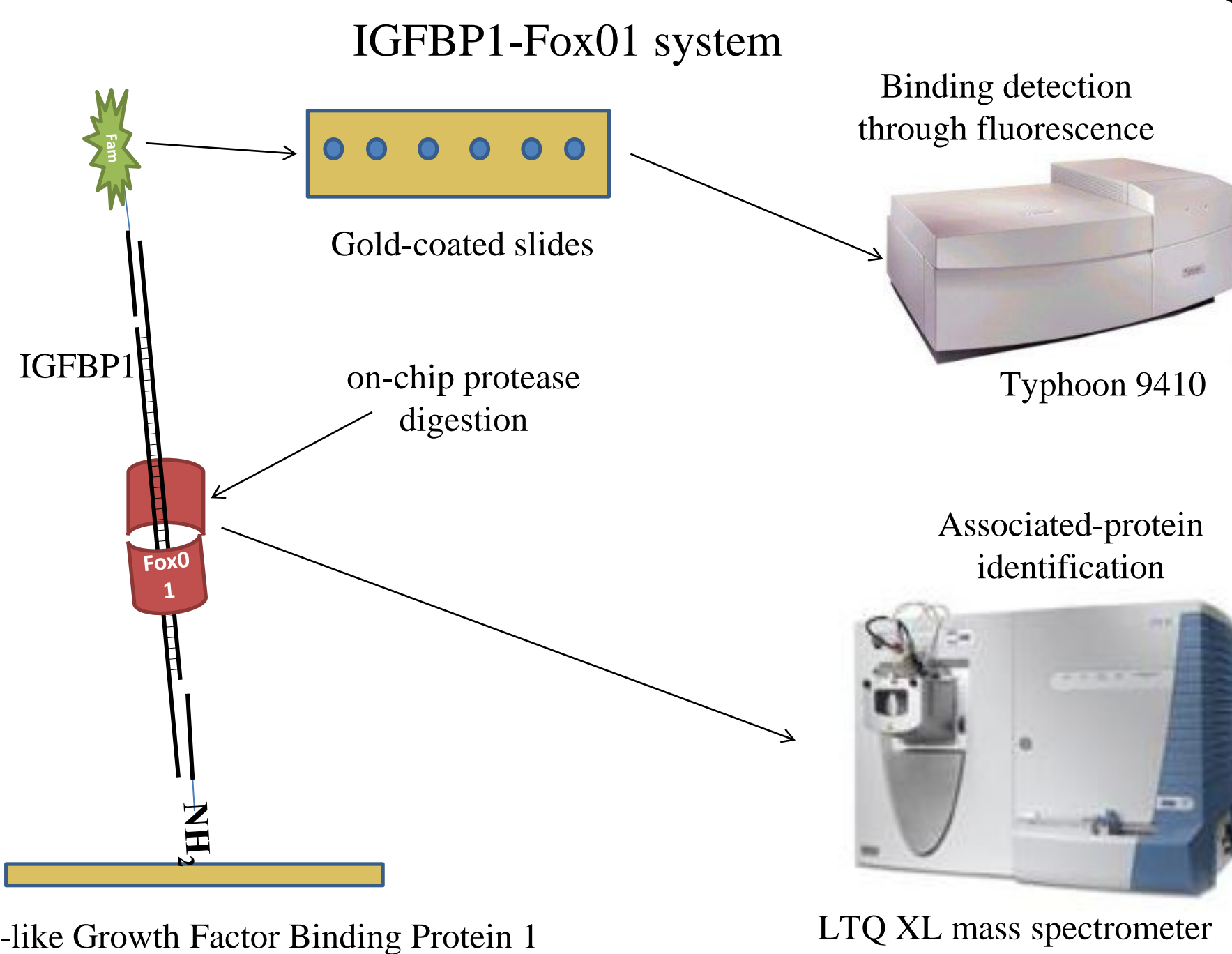
Genomic Strategy using Streptavidin Beads



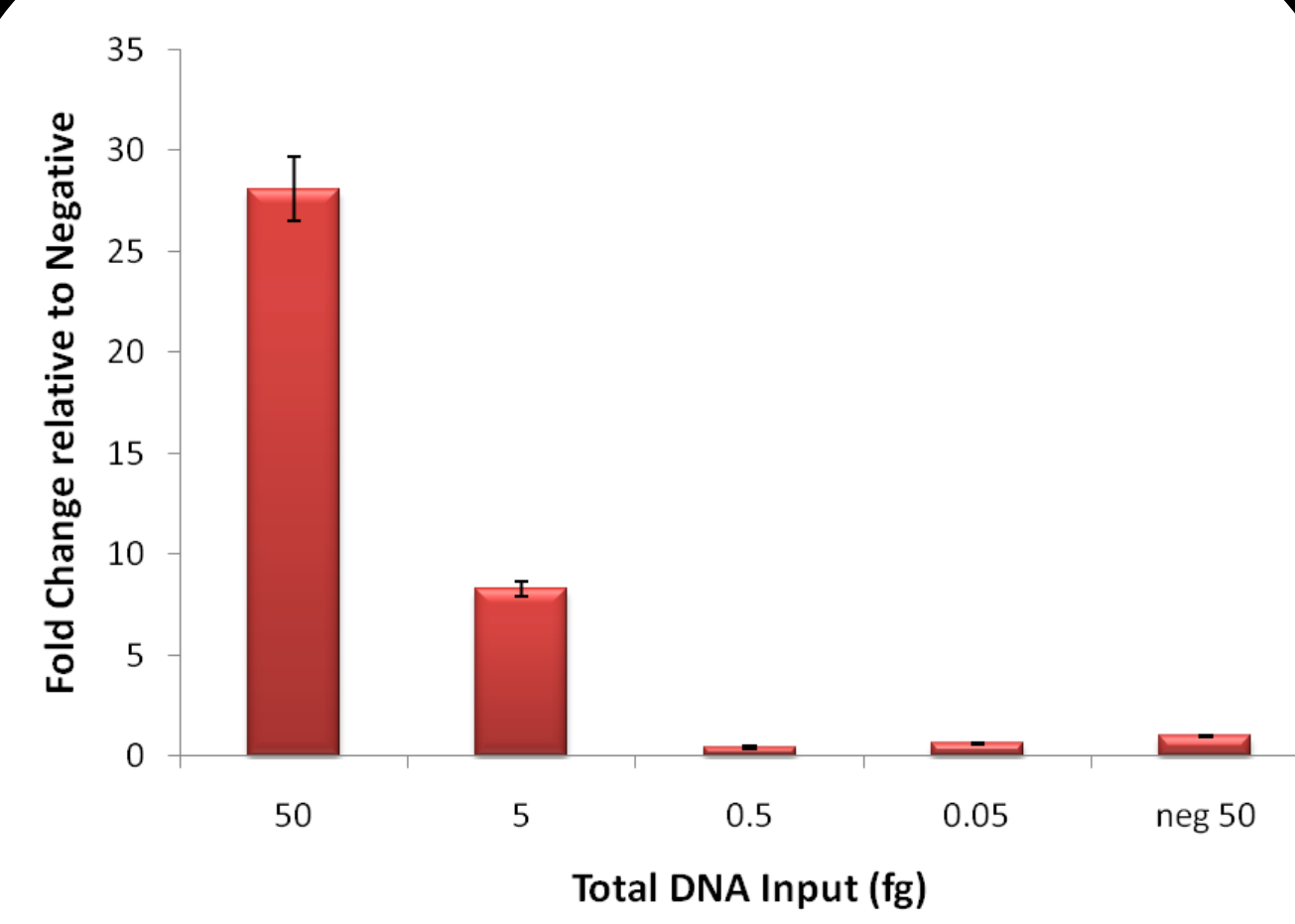
The UAS_{GAL} Region. A 3k bp PCR product was generated, which featured the upstream activation sequence for Gal (UAS_{GAL}) region with its three Gal4 binding sites, and restriction sites for EcoRI and SfaNI on the Gal10 and Gal1 genes. Initial restriction enzyme digestions and hybridization were performed on PCR products. Further captures were then carried out on genomic DNA and cell lysates.



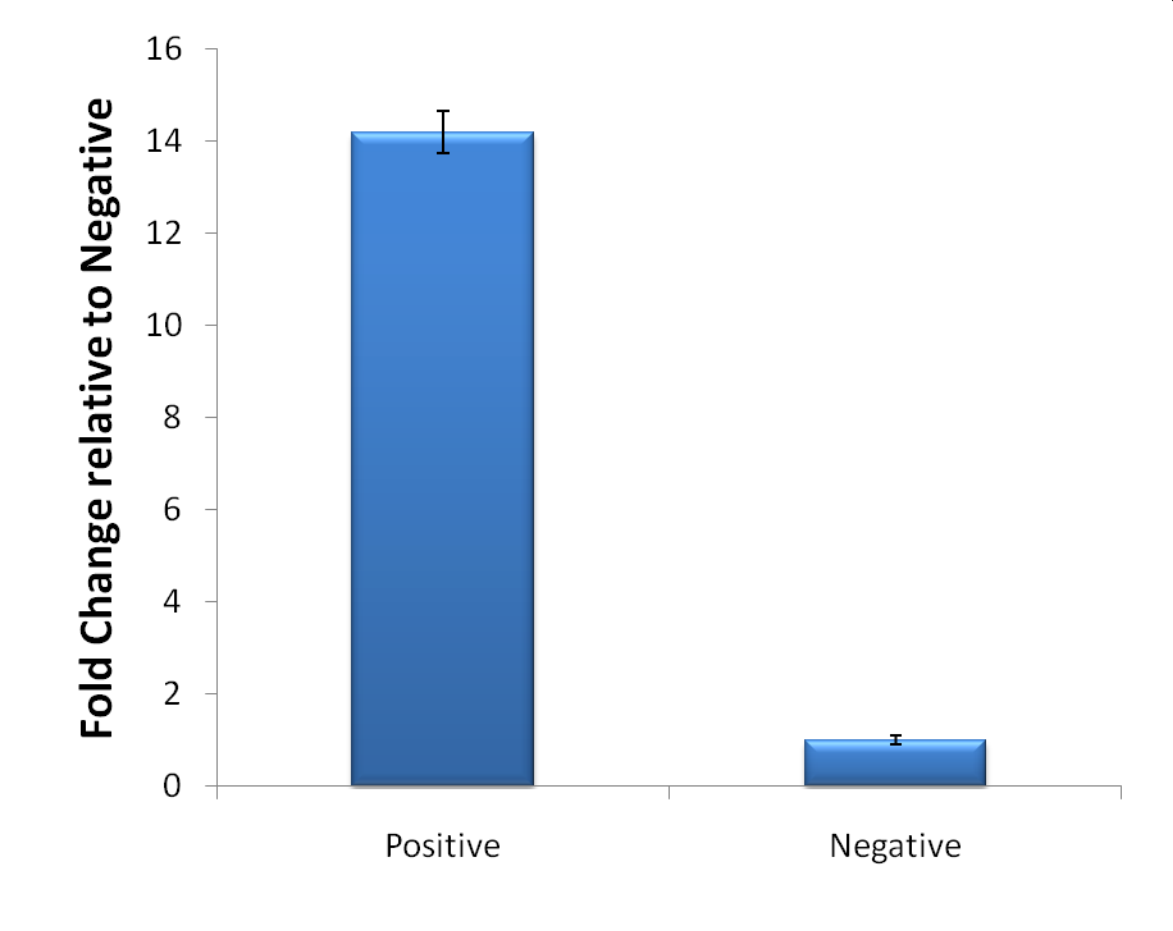
Restriction enzyme digestion. DNA samples were treated with varying concentration of either EcoRI or SfaNI during different incubation times and run in an agarose gel (a). Taqman assays were designed across restriction sites (b) so that untreated samples would show good amplification, while digested ones would show low amplification (c). Quantification of the qPCR experiments showed what concentrations work best (d).



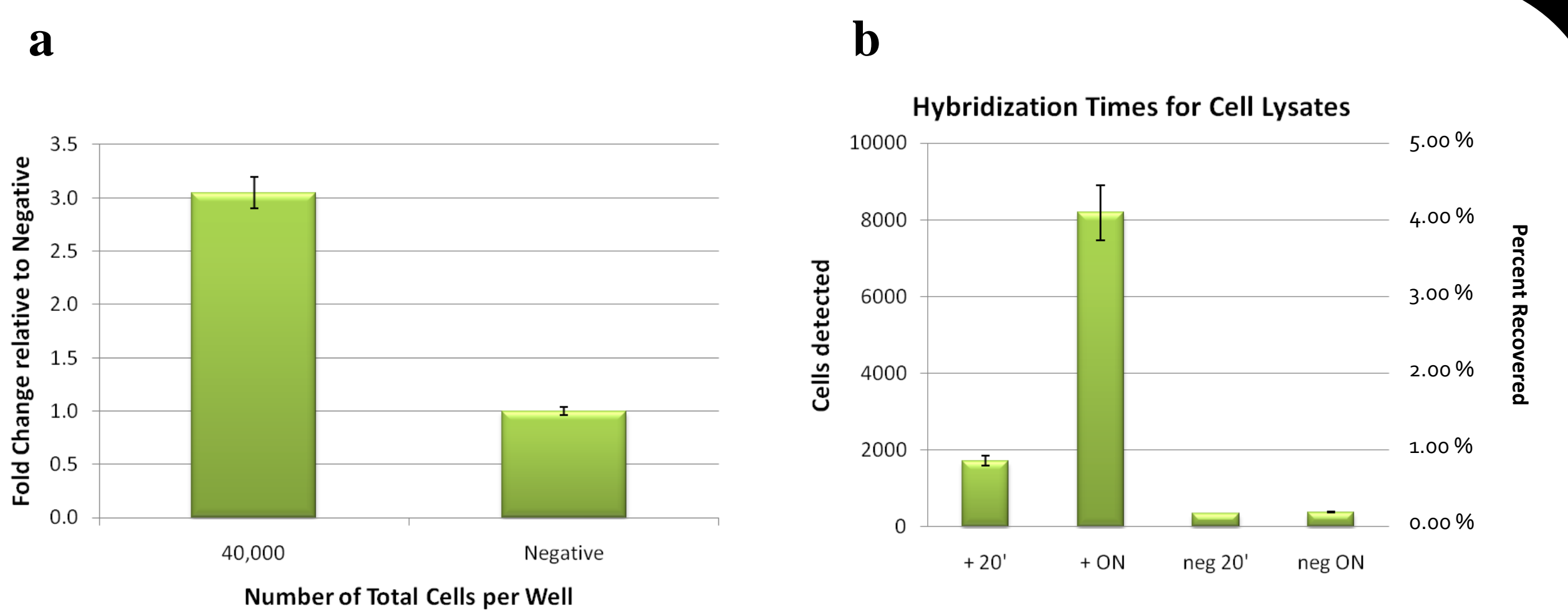
Workflow diagram for gold covered slides. Capture oligos specific to the IGFBP1 promoter region were designed and attached to gold-covered slides by linking them to succinimidyl ester-terminated alkyldisulfides, which are amine-reactive. PCR products were then bound to recombinant FoxO1 and single strand overhangs generated. Bound DNA was detected through a fluorescent labeled oligo specific for the opposite end. Bound FoxO1 was detected via tandem mass spectrometry.



Capture of PCR product. Digested PCR products were captured at different concentrations using streptavidin-coated beads. Bound DNA was detected and measured with qPCR. Fold change increase relative to the negative control is shown.



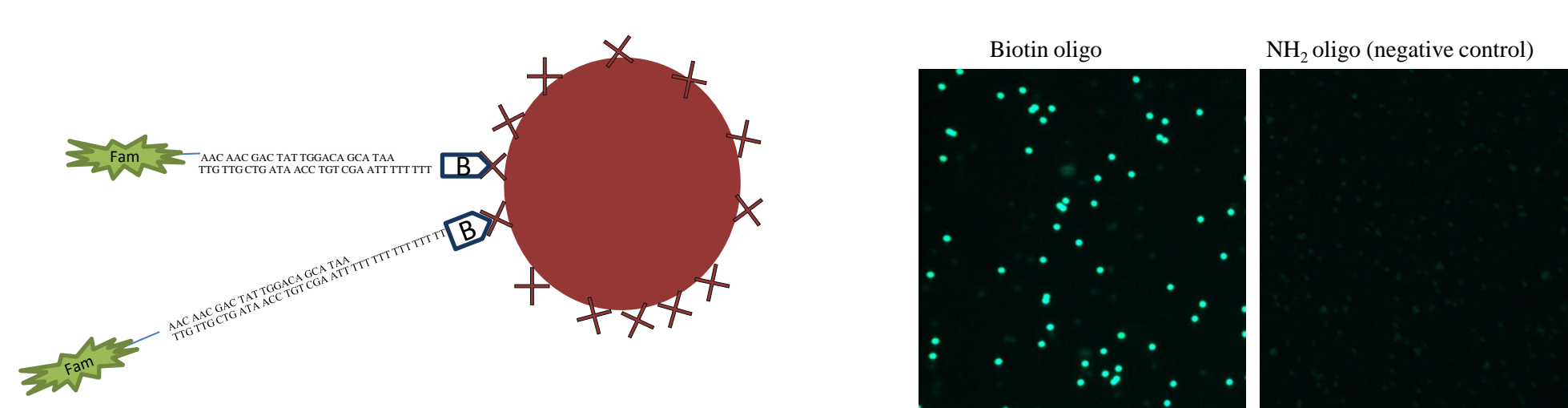
Capture of genomic DNA. Zymolase was used to lyse yeast cells and isolate purified genomic DNA. Samples were hybridized over night and DNA was detected and measured with qPCR. Fold change increase relative to the negative control sample is shown.



Direct capture from cell lysates. Yeast cells were lysed using a french press. After digestion and overnight hybridization, qPCR was used to detect and measure captured target. A 3 fold increase relative to the negative control was observed on samples that had at least the equivalent of 40,000 cells (a). Times of incubation were compared by hybridizing cell lysate samples for either 20 minutes or 16 hours (b). A clear increase in capture was observed for samples that underwent over night incubation. Out of an input of 200,000 cells, the equivalent of 8,000 cells was detected, which represents a 4% efficiency.

Rank	Name	Accession	Description	Protein PtcScore	Peptide Cg	Scan	Count
1	FOXO1_MOUSE	Q9R1E0	Forkhead box protein O1A	1	161	662	26
2	TRY1_RAT	P00762	Tryptophan 1 precursor	1	8	6768	16
3	ENO4_RAT	P04764	Alpha-enolase (EC 4.2.1.11)	0.999879	0.5193	1	1
4	ENO3_RAT	P15429	Beta-enolase (EC 4.2.1.11)	0.999879	0.5193	1	1
5	ENO2_RAT	P07323	Gamma-enolase (EC 4.2.1.1)	0.999879	0.5193	1	1
6	GSRP_RAT	Q09028	Gamma-aminobutyric acid 1	0.864876	0.28781	1	1
7	TAAT3_RAT	Q923Y2	Trace amine-associated rec	0.864532	0.31095	1	1
8	RIN3_MOUSE	P59729	Ras and Rab interactor 3 (R)	0.839074	0.43574	1	2
9	SPAST_MOUSE	Q9QYV8	Spastin	0.837104	0.19259	1	1
10	CRY1_RAT	Q32086	Cryptochrome-1	0.737867	0.27289	1	1

Fluorescent scan and data acquisition. The image on the left shows the fluorescence detected on the gold coated slide after hybridization. The negative control (arrow) does not show any detectable fluorescence. FoxO1 peptides were the only ones observed with any significance in the MS analysis (probability =1 and >2 peptides). These peptides represent a coverage of 34.97% of the entire protein sequence.



Capture on streptavidin beads. Biotin labeled oligos were generated specific to the overhangs generated after restriction enzyme treatment and digestion. Hybridization was tested with complementary Fam labeled oligos.

Summary

IGFBP1 promoter region (in vitro system)

- Gold coated slides allowed the capture of the IGFBP1 promoter region.
- Mass spectrometry clearly identified DNA-bound FoxO1 with good overall coverage.

UAS_{GAL} region (genomic DNA)

- Streptavidin beads offer a reliable and flexible alternative for capturing the upstream activation sequence (UAS_{GAL}) region in a sequence-specific manner.
- Taqman assays have been developed to detect and measure the capture of the UAS_{GAL} region. Taqman assays have also been developed to measure digestion efficiencies for EcoRI and SfaNI.
- Successful direct capture of genomic DNA fragments from whole cell lysates.