Breakthroughs in Biology

Biowire

Fall 2010

SIGMA Life Science Where bio begins™

A Look Into the Future of
Gene Editing

Biodelete.

Knockout genes in the way of your research.

CompoZr® Zinc Finger Nuclease (ZFN) technology gives you the ability to permanently knockout a gene of interest. Allowing the creation of cell lines and model organisms with precise and heritable genome modifications. CompoZr ZFN technology, a breakthrough that enables simple and efficient genome editing, gives you the power to *add, delete, swap your gene of interest.*

wherebiobegins.com/bioediting

SIGMA-ALDRICH®

CompoZr is a registered trademark of Sigma-Aldrich and Sigma-Aldrich Biotechnology, LP.

in this issue

Fall 2010

ARTICLE SPOTLIGHTS

David Smoller, Ph.D.
President, Research Biotech

GENOME BY DESIGN

In any area of scientific research, asking the right question has always been paramount to advancing knowledge.

Reprinted with kind permission from DDW Insights a supplement to Drug Discovery World Summer 2010 Volume 11 Issue 3.

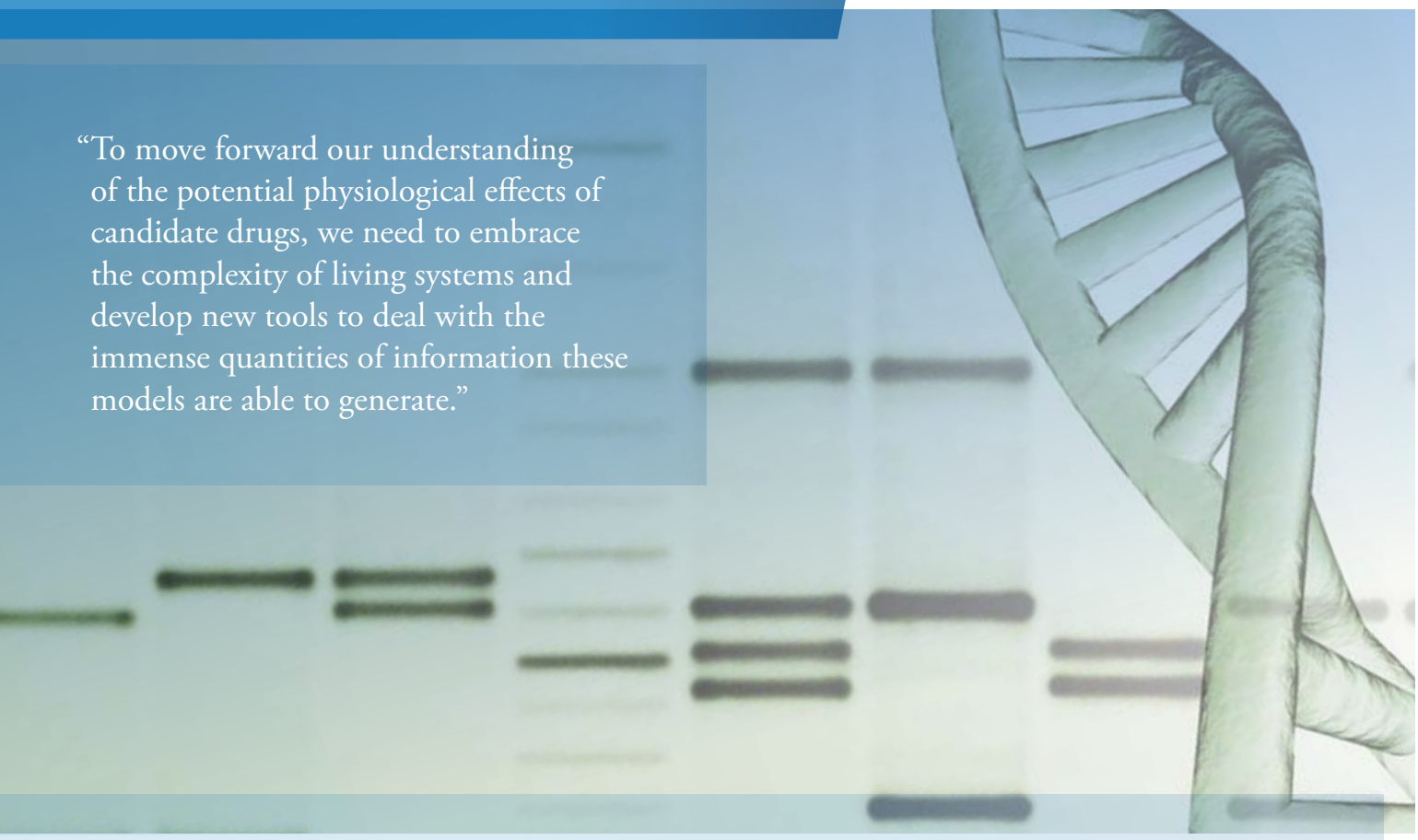
Without a robust understanding of what is being observed, no conclusions can be drawn. In biology, the multifaceted nature of even basic *in vitro* systems makes this a complex task, requiring not just stringent control of as many variables as possible, but also an appreciation of the likely impact of those variables that cannot be influenced.

This is especially true in drug discovery, where a candidate compound that appears highly effective in biochemical trials may prove ineffectual, or even harmful, at a cellular or physiological level due to a wide variety of factors. Adverse drug reactions (ADRs) in particular can be highly detrimental for pharmaceutical manufacturers, either in terms of lost revenue or negative publicity, and generally stem from the inherent flaws in the traditional models routinely used for investigation of candidate drugs. Despite the wealth of knowledge gained from studying immortalized cell lines, for example, these only provide an approximation of the biology of the target organism.

To improve the accuracy of cell- and animal-based models, Sigma® Life Science is focused on development of innovative technologies – such as zinc finger nucleases – to allow highly accurate genetic manipulation of model cell lines and organisms, enabling researchers to answer specific biological questions.

Zinc finger nucleases (ZFNs) offer a unique solution for mammalian genome editing, and can be designed to generate site-specific double-strand breaks (DSBs) at virtually any point in the genome. These breaks can then be repaired by either non-homologous end joining (NHEJ), leading to gene knockout, or homologous recombination (HR), achieving precise gene insertion, deletion or replacement. This technique allows genes to be altered virtually at will, and has already been accomplished at multiple loci using ZFNs in conjunction with donor DNA containing the desired mutation. ZFNs also enable delicate, specific mutagenesis of single-base sites in the native context of the gene, lending more biological relevance to the resulting data. The ability to make such mutations enables a more rigorous investigation of single nucleotide polymorphisms (SNPs) associated with disease states. Application of ZFN technology to existing immortalized cell lines potentially offers several advantages for current cell-based assays, both in terms of introducing new genetic variations and eliminating erroneous mutations, however the advent of induced pluripotent stem cells (iPS cells) raises the possibility of developing bespoke cell lines which truly reflect genetic make-up of a given disease.

To truly understand the impact of editing the genome, and the subsequent effect of a candidate compound on negating this genetic change, it is necessary to go beyond basic cellular systems and investigate animal models. In the last 20 years, the availability of knockout mouse technology has established the mouse as the most widely used model system in biomedical research. The key to the success of this technology lies in the ease with which mouse



"To move forward our understanding of the potential physiological effects of candidate drugs, we need to embrace the complexity of living systems and develop new tools to deal with the immense quantities of information these models are able to generate."

embryonic stem (mES) cells can be cultured, genetically modified, cloned, and implanted into blastocysts to derive chimeric animals.

Although this technology provides an excellent mechanism for the study of basic mammalian genetics, it cannot be used to generate clinically-relevant data for drug discovery. However, in 2009, the first targeted knockout rats exploiting ZFNs were created via microinjection of ZFN mRNA into single cell embryos. Screening of the resulting animals was surprisingly rapid and efficient, enabling the production of founders in as little as two months post-microinjection, and involving screening of less than 100 animals. Such speed of targeted mutations in animals is virtually unprecedented, and both Sigma Advanced Genetic Engineering (SAGE™) Labs and Medical College of Wisconsin have been able to rapidly establish libraries of knockout mice and rats for the research community. Most importantly, the ease of ZFN-mediated knockout rat generation indicates the suitability of the technique for targeted gene knockout in other species. Because the technique does not require ES cell lines for the species of interest, existing embryo isolation, injection and implantation protocols for creation of transgenic animals can be employed. As robust embryo manipulation methods already exist for a variety of species – including mice, rats, rabbits, chickens, sheep, goats, cows and pigs – this offers a straightforward route to efficiently delivering ZFNs and deriving knockout animals.

The advent of ZFN technology highlights the potential for such techniques to be used in development of new model systems to answer specific biological questions. To move forward our understanding of the potential physiological effects of candidate drugs, we need to embrace the complexity of living systems and develop new tools to deal with the immense quantities of information these models are able to generate. Greater throughput, coupled with the ability to collect more data per sample, will offer a more comprehensive look at how individual molecular interactions fit into the complex biochemical pathways of the living cell. The future is still dependent on scientists asking the right questions, but the ability to design and control new biological systems is an important step along the path to answering those questions. ■

ZFN TECHNOLOGY

Have your genomic work cut out for you

The genomes of several organisms, including humans, have been sequenced, yet we have a limited understanding of the functional implications of the identified genes.

Extensive analysis is required to understand the function of each gene and put that into the systemic context. This has led to an increased focus on functional genomics. Traditional approaches to understand gene function in mammalian systems have been limited to the use of RNAi or genetically modified mice, due to the lack of requisite tools for targeted genome editing in other systems. The discovery of the modularity of zinc finger proteins and subsequent development of Zinc Finger Nuclease (ZFN) technology has trampled these technological barriers, giving scientists the ability to manipulate the genome of any organism in a highly targeted manner. The applications of ZFNs are broad and far-reaching, enabling scientists to answer technologically challenging questions with unparalleled ease.

What are Zinc Finger Nucleases?

Zinc Finger Nucleases are a class of engineered proteins that create a highly targeted double-strand break (DSB) within the genome and enable the manipulation of the genome with unprecedented ease and precision. ZFNs consist of two domains, a zinc finger DNA binding domain and a catalytic nuclease domain (**Figure 1**). The zinc finger domains used here are derived from the Cys(2)-His(2) zinc finger proteins that each recognize 3 bases of DNA. These zinc finger proteins can be engineered to recognize most 3 base pair sequences, allowing targeting to almost any sequence in the genome. The modularity of each zinc finger permits us to link together 4-6 zinc finger proteins creating 12-18 base pair specificity within the genome.

Fewer fingers can also be used but specificity is reduced. To complete the engineered ZFN, we add a modular nuclease domain from *Fok*I, a Type IIs restriction endonuclease (which has a separate DNA-binding domain and nuclease domain), to the zinc finger domain. The nuclease domain from *Fok*I functions as a dimer, requiring that a second ZFN binds adjacent to the first one. The requirement of a pair of zinc finger nucleases to create a double-strand break gives the site 24-36 base pair specificity. The target sites of the two ZFNs must be separated by 5-7 base pairs to allow formation of the catalytically active *Fok*I dimer (**Figure 1**).

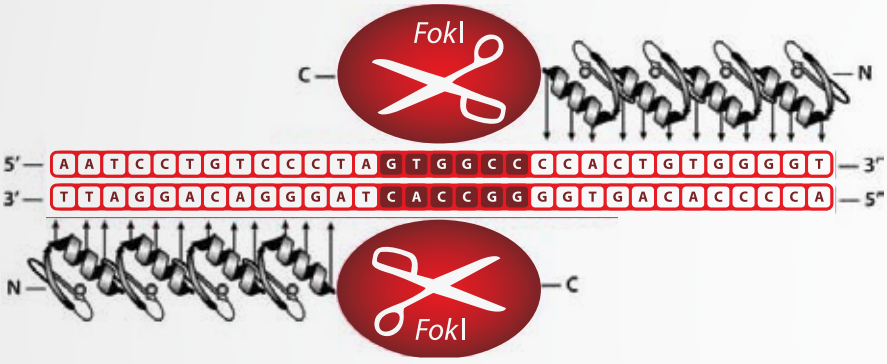


Figure 1.

Each Zinc Finger Nuclease (ZFN) consists of two functional domains: A DNA-binding domain comprised of a chain of zinc finger modules, each recognizing a unique triplet (3 bp) sequence of DNA. Four to six zinc finger modules are stitched together to form a Zinc Finger Protein (ZFP), with specificity of ≥12 bp. A DNA-cleaving domain comprised of the nuclease domain of *Fok*I is attached to the ZFPs. When the DNA-binding and DNA-cleaving domains are fused together, a highly specific pair of 'genomic scissors' is created that binds with 24-36 bp specificity of the ZFPs and cleaves the DNA.

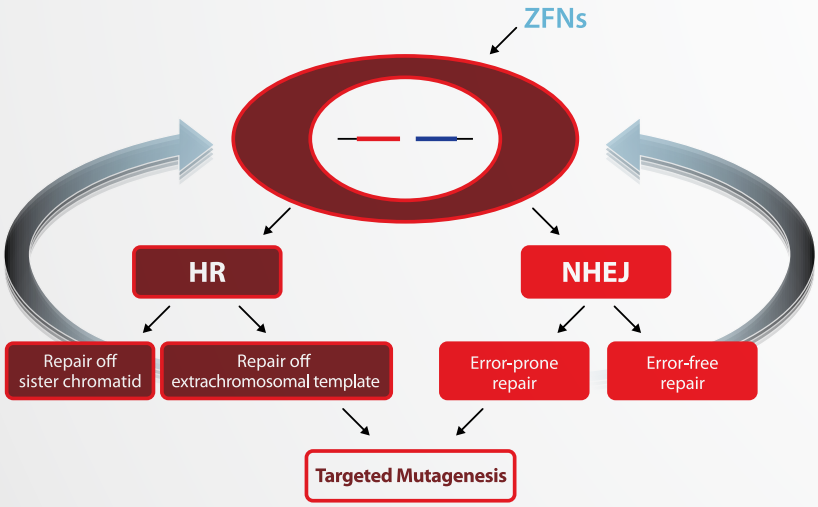


Figure 2.

The addition of zinc finger nucleases to the cell results in creation of a double-strand break at the target site. This double-strand break is repaired by one of two endogenous repair pathways, either the non-homologous end joining (NHEJ) or the homologous recombination (HR) pathway. NHEJ is used to create gene knockouts while HR is utilized for targeted integration.

How do Zinc Finger Nucleases work?

Zinc Finger Nucleases can be designed to target any gene in any genome. These can then be delivered to the cell as DNA or RNA. The ZFN proteins are expressed, then translocate to the nucleus and bind their target sites with high specificity. Upon binding the target sites, the *Fok*I nuclease forms its catalytically active dimer and creates a single, specific double-strand break at the user-defined locus. Living cells have evolved several methods to repair double-strand breaks, and these well-studied, endogenous processes can be harnessed to create gene knockouts or knock-ins (**Figure 2**). Following the creation of a double-strand break, the cell chooses to repair it either via non-homologous end joining (NHEJ) or via homologous recombination (HR). NHEJ is an imperfect repair system, and can lead to insertions or deletions of base pairs at the site of the double-strand break. The mutation leads to the creation of a frameshift within the coding sequence of the gene, an in-frame deletion or exon skipping that may disrupt gene translation and knockout gene function. If the cell is provided with a donor plasmid that has a sequence homologous to regions immediately flanking the double-strand break, the cell may use the donor for homology directed repair or HR. The donor plasmid can be designed to include transgenes for targeted integration, create specific mutations at the endogenous locus, or tag endogenous genes. Here we describe an example of a ZFN-induced gene knockout in a human lung cancer cell line, a ZFN-induced targeted integration of a miRNA at the AAVS1 locus in the human genome, and a ZFN-induced tagging of human tubulin with green fluorescent protein.

Figure 3A.

```
wt   AGGAAGTCCAATGTCCAGCCCATGATGGTTCTGATCAGTTCCGGCACCTTGGTGCACAGGGCCTGTGG
-2   AGGAAGTCCAATGTCCAGCCCATGATGGT--TGATCAGTTCCGGCACCTTGGTGCACAGGGCCTGTGG
-4   AGGAAGTCCAATGTCCAGCCCATGATGGTT----TCAGTTCCGGCACCTTGGTGCACAGGGCCTGTGG
-7   AGGAAGTCCAATGTCCAGCCCATGATG-------TCAGTTCCGGCACCTTGGTGCACAGGGCCTGTGG
-19  AGGAAGTCCAATGTCCAG-----------------TCCGGCACCTTGGTGCACAGGGCCTGTGG
```

Figure 3B.

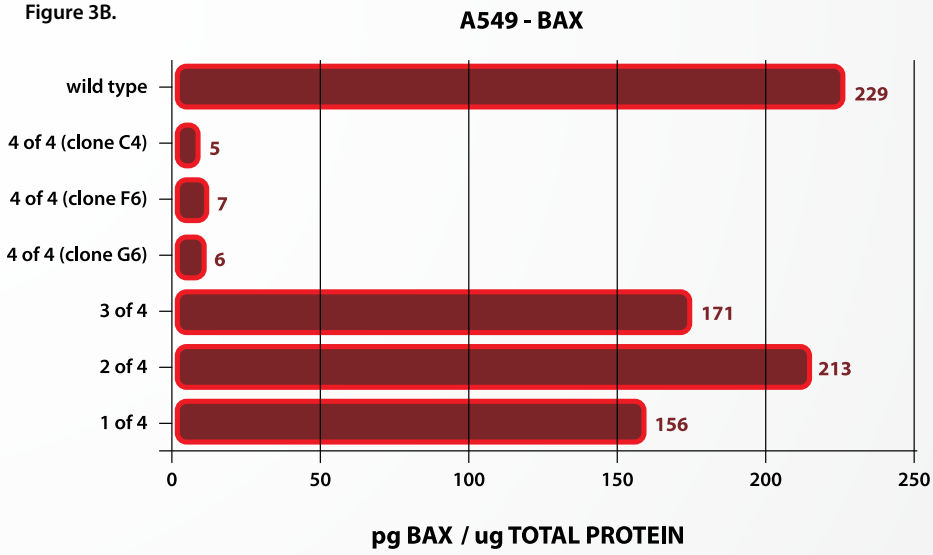


Figure 3.

Knockout of tetraploid BAX in A549 cells using ZFNs. **(A)** Following treatment with a ZFN specific for BAX, a clone was isolated that contained a unique disruption in all four alleles of BAX. The DNA sequences of the wildtype and four disrupted alleles are shown above. **(B)** BAX protein concentration was measured in wild type and knockout cell lines using an enzyme immunometric assay specific for BAX. Quantification of BAX protein levels was obtained through comparison of BAX protein levels in the cell lysate to a standard curve of recombinant BAX protein. Clones having 1, 2 or 3 out of 4 alleles disrupted produce less BAX protein. For three unique clones, each with all 4 alleles disrupted, the measurement was below the lower limit of detection demonstrating that no BAX protein is produced when 4 out of 4 alleles are disrupted. Clones with 1, 2 or 3 of 4 alleles disrupted produce less BAX protein (The linear detection limit for the assay is at 15 pg). Data provided by Suzanne Hibbs and Gregory Wemhoff, Ph.D., at Sigma Life Science.

Targeted Knockout of BAX in Human Lung Cancer

Previously, mammalian cell line engineering techniques to disrupt specific genes was limited to time consuming homologous recombination and screening approaches. The advent of ZFN technology now enables simple, rapid and permanent disruption of specific gene loci. Sequence specific nuclease cleavage, followed by imperfect DNA repair, gives rise to permanent gene mutations. ZFNs introduced against a particular gene create a single, specific double-strand break at the target genomic locus. This activates the NHEJ repair pathway that is referred to as "non-homologous" because the break ends are directly ligated without the need for a large homologous template. NHEJ usually repairs the break accurately, but sometimes the broken DNA is acted upon by nucleases and polymerases that can lead to the insertion and/or deletion of bases. This insertion or deletion of bases leads to disruption of the coding sequence and functional loss of the protein. It is this less frequent outcome of incorrect repair that enables the targeted DSBs from the ZFNs to result in the specific loss of gene product and function.

Here, we demonstrate an example of the rapid construction of genetically defined cell lines using ZFNs (The protocol used here is in the technical bulletin for CSTZFN). The pro-apoptotic Bcl-2–associated X protein, BAX, was targeted for deletion in three human cancer cell lines: A549, DLD-1, and SW48. In all cases, even in A549 cells where BAX is tetraploid, knockout lines were successfully generated. An immunometric assay was used to confirm that BAX protein is not produced in the complete knockout cell lines, and protein production was decreased in partial knockout cell lines. In addition to multi-allelic disruptions, cell lines with monoallelic disruptions were also isolated. DNA sequence alterations were verified by sequencing, and the absence of BAX protein was confirmed by enzyme immunometric assay (**Figure 3**). The straightforward approach coupled with the power of ZFN technology to address both diploid as well as polyploid targets, makes this an ideal tool in the generation of these cell lines. Targeted gene editing provides a critical tool that enables the study of genes involved in cancer and other processes.

Targeted Integration of a miRNA at the AAVS1 locus

The targeting power of the ZFN technology allows placement of transgenes precisely into user-specified sites in the genome with high-efficiency. Traditional use of random insertion to introduce a DNA construct into the human genome may lead to gene disruption, variable epigenetic effects on transgene expression and presence of an unknown number of integrants in each cell, making comparisons difficult between isolated single cell clones. ZFN-induced targeted integration of a transgene into a specific locus in the human genome enables stable transgene expression, regulation of copy number and site of integration, without the need to engineer cell lines with randomly integrated transgene landing pads, e.g., the insertion of recombinase sites within the genome. The specificity of targeting comes not only from the homology arms in the donor, but also from the site of the DSB created by site-specific ZFNs. Upon creation of this DSB the cell may use HR for repair using a homologous template, typically

Figure 4A.

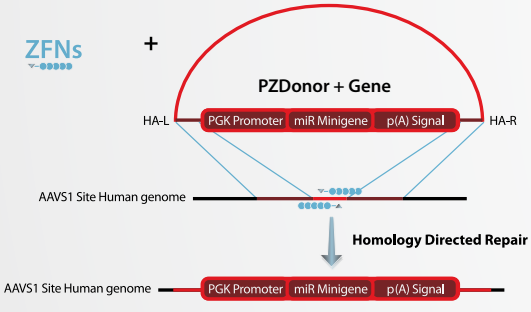


Figure 4C.

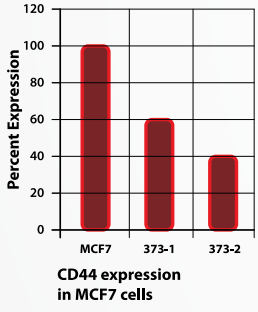


Figure 4B.

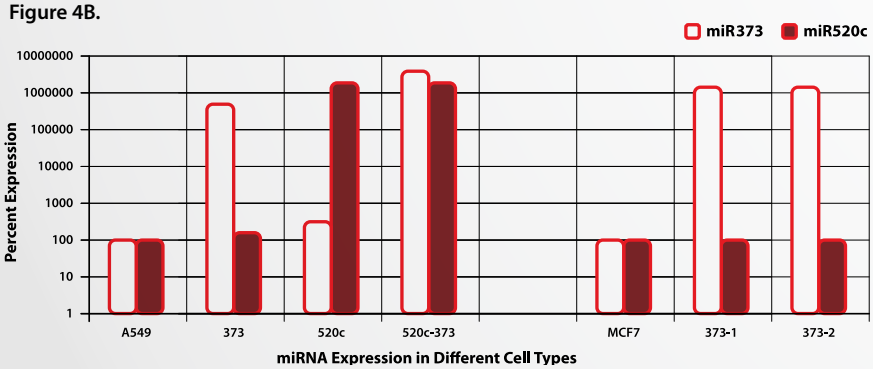


Figure 4.

(A) Schematic showing the construction of A549 and MCF7 cells with ZFNs and pZDonor plasmid containing a microRNA (miRNA) gene. The pZDonor plasmid containing the miRNA was co-nucleofected into A549 and MCF7 cells with mRNAs that express AAVS1-specific ZFNs. These ZFNs cut at the AAVS1 locus within the human genome (binding site in red). The miRNA gene is integrated into the genome through homologous recombination using the homology arms on the pZDonor plasmid. The miRNA is driven by a PGK (phosphoglycerate kinase) promoter and has a synthetic poly(A) signal. **(B)** The miRNA is expressed and processed correctly in A549 and MCF7 cells. Human miRNA-373, miRNA-520c and tandem (cluster) miRNA520c-373 were cloned into the pZDonor vector. All three combinations were used in A549 cells and only miR-373 was used in MCF7 cells. Values are expressed on a log scale as percent expression compared to the wild type. **(C)** Endogenous CD44 mRNA levels are reduced in MCF7 cells. CD44 is a known target for human miRNA-373. Endogenous levels of CD44 mRNA were reduced upon expression of miRNA-373. Values are expressed as percent expression compared to the wild type. Data provided by Kevin Forbes, Ph.D., and Carol Kreader, Ph.D., at Sigma Life Science.

the sister chromatid. We usurp this natural process by flooding the cell with a donor DNA template bearing homologous sequences to the site of the genetic lesion and any additional genetic information the user wants to insert at that locus between the homology arms (**Figure 4**). The cell uses the donor as the repair template and subsequently integrates the desired transgene into the locus. This form of repair can be thought of as a "copy and paste" mechanism having a higher fidelity than NHEJ, but requiring a longer region of homology.

Here we demonstrate an example of the use of a safe harbor site, a genomic locus that maintains transcriptional competence and has no known adverse effect on the cell from its disruption, for the expression of miRNA genes in metastatic cancer cells. Earlier studies have demonstrated the function of miRNA genes in metastatic cancer cells using lentiviral introduction of the miRNAs. The lentiviral introduction of the miRNAs will lead to random integration of the construct into the human genome with the risk of deleterious effects on the integration site. The use of a safe harbor locus (such as Rosa26 in mice) will allow expression of a construct from a known genetic and epigenetic locus. We used the AAVS1 site, which has been characterized as a safe harbor locus. We were able to express one or multiple miRNAs from a single locus and successfully demonstrate the knockdown of CD44, a known target of the expressed miRNA. This method demonstrates that the successful use of the AAVS1 safe harbor locus for the creation of transgenic human cell lines expressing miRNA genes as individuals or as clusters. Recent work has also demonstrated that CCR5 can also be used as an additional safe harbor locus to create transgenic human cells.

Tagging Endogenous Tubulin with GFP

As discussed earlier, ZFN technology is a fast and reliable way to manipulate the genome in a targeted fashion. In the majority of cases ZFNs have been used to create gene knockouts utilizing non-homologous end joining (NHEJ) – one of the main pathways of double-strand break (DSB) repair in somatic cells. Here we relied on the second DSB repair pathway – homologous recombination (HR) to tag cytoskeletal genes by integrating a fluorescent reporter sequence into the desired location in the genome. The integration resulted in endogenous expression of the corresponding cytoskeletal protein fused to a fluorescent protein. These fusion proteins have the advantage of easy and quick detection at their native expression level and characteristic localization pattern in the cell. Three gene loci were tagged: LMNB1 (lamin B1, nuclear envelope), TUBA1B (α-tubulin 1b, microtubules: **Figure 5**), & ACTβ (β-actin, actin fibers). Green and red fluorescent proteins were used as reporters. Post transfection into U2OS cells, fluorescent cells were isolated using fluorescence activated cell sorting (FACS) to identify single cells with gene integration at the endogenous locus. These cells were further characterized for gene expression via southern blotting and western blotting to determine integration specificity and protein expression levels. Moreover, we were also able to isolate cells that had received dual treatment with ZFNs and donor, and were double integrated for fluorescent protein fusions of Lamin B1 and β-Actin. This demonstrates the ability to stack different traits in a single cell line. The details of the donor design method and protocols associated with the creation, isolation and characterization of the cell lines can be found on page 39.

Figure 5A.

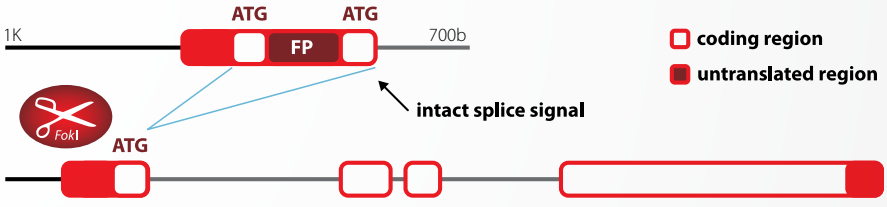


Figure 5B.

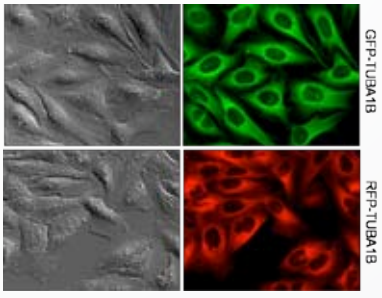


Figure 5C.

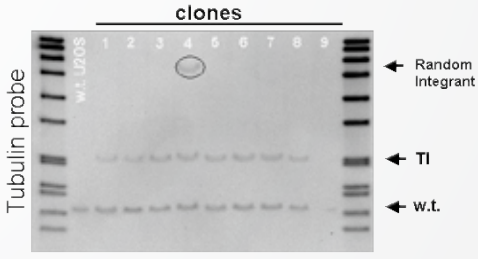


Figure 5.

U2OS cell lines with TUBA1B tagged with green fluorescent protein (GFP) and red fluorescent protein (RFP) at the endogenous locus. **(A)** Schematic of the TUBA1B loci and donor including the length of the homology arms, the position of the fluorescent protein at the genomic locus and the preservation of the splice site on the first exon of the gene. **(B)** DIC & fluorescence microscopy images of endogenously labeled TUBA1B with GFP and RFP in U2OS cells. **(C)** Southern hybridizations were performed on DNA isolated from wild-type U2OS and nine single cell clones positive for red tubulin fluorescence. Using the tubulin probe, the ~1900bp band represents the RFP-fused tubulin genomic locus while the ~1200bp band represents wildtype (wt) tubulin genomic locus. Targeted integration (TI) in these cells occurred in a heterozygous manner as seen by the presence of both bands in the lanes 1-9. Lane 4 also has a random integration event aside from the targeted integration. Data provided by Hongyi Zhang and Nathan Zenser, Ph.D., and Dmitry Malkov, Ph.D., at Sigma Life Science.

Professor of Biochemistry Dana Carroll, Ph.D., one of the originators of Zinc Finger Nuclease technology for genome editing at the University of Utah

An Interview with Dr. Dana Carroll

University of Utah

In each issue of Biowire, we will bring you a spotlight of a thought leader who has paved the way for technologies that benefit scientists everywhere.

CompoZr® Zinc Finger Nucleases (ZFNs) are the newest, most efficient, method for genome editing, and are now available from Sigma® Life Science. One of the researchers who made this technology possible is Dr. Dana Carroll of the University of Utah. His lab demonstrated that ZFNs have the capability to make targeted double-strand breaks on chromosomes within living cells.

Strangely enough, Dr. Carroll's early science studies were focused on chemistry, physics and mathematics. It was not until he started graduate school that he began to attend seminars in biology.

Dana Carroll:

> (At first) I didn't know the vocabulary. After a while I learned to speak the language, and I decided I was more interested in biology and the applications of biochemistry than I was in straight chemistry.

From the time Dr. Carroll started in biology, site-specific DNA breaks were of high interest to him. In the early to mid-1990s, his lab was working with triplex forming oligos, which bind in the major groove of duplex DNA with high specificity and affinity. Therefore, when the lab of Dr. Srinivasan Chandrasegaran demonstrated that site-specific DNA cleavage could be achieved by attaching *Fok*I to novel DNA binding proteins, Dana saw the potential for collaboration.

Dana Carroll:

> In the early 1990s we were already trying to develop reagents that could make site specific breaks in DNA. We knew that double-strand breaks could be recombinogenic, and so we were playing around with triplex forming oligos. When Dr. Chandrasegaran's paper appeared in 1996, we were already primed to understand the potential of these hybrid enzymes. I called Chandra right after his paper came out. I had never met him before, but I called him and suggested we collaborate to study whether these hybrid proteins could stimulate recombination. And that's how we got started. We already knew about double-strand breaks in DNA stimulating recombination, and we knew there was potential to manipulate the DNA binding modules of zinc fingers. So, when that paper came out, we were ready.

Dr. Carroll is quick to give credit to those who have come before him in the field, making his accomplishments possible. We asked him who some of his inspirations were. The list is long and impressive, and shows how several scientists have contributed to illuminating the complexities of biology.

> "Technologies which are now commonplace in laboratories across the world had to be discovered, then optimized for all to use. This is true for DNA isolation and purification, PCR, etc." *– Dr. Carroll*

Dr. Dana Carroll

Dana Carroll:

> I have a huge amount of admiration for a lot of people in science. One is Don Brown, who I did a postdoc with back in the early 1970s. He was somebody who worked in the lab throughout his career, and was a very serious scientist, not interested in promoting himself. He was the one who characterized the first eukaryotic transcription factor, TF3A, and this is actually where zinc fingers were first identified. Aaron Klug's group found that these repetitive modules bind zinc and are involved in DNA recognition.
>
> And another person (I admire) is Joe Gall. He's the same generation as Don Brown, and if you look back over his career, he developed *in situ* hybridization and was the first to isolate and sequence telomeres. He knows so much about biology that, when he has a particular problem, he can choose the organism that he thinks is best suited to solving that problem.

Not able to stop at just two people, Dr. Carroll added Mario Capecchi, 2007 Nobel Prize winner in Physiology or Medicine for developing gene targeting in mouse ES cells.

Dana Carroll:

> I have a huge amount of admiration for Mario. He doesn't let himself get distracted by irrelevant things. He went after the problem of doing recombination in mammalian cells, despite people thinking that they would never work. He developed the technique that people are using all over the world now.

In talking with Dr. Carroll, and looking at his website (www.biochem.utah.edu/carroll/), it's clear that collaborations play a large part in his research. We asked him to comment on the role that collaboration plays in his work.

Dana Carroll:

> Collaboration is a way to enhance a single lab's capabilities. In my recent experience, working with others has allowed us to try the ZFN gene targeting technology in quite a number of different organisms. My scientific outlook has been broadened enormously, as I learn about the particular features of these various systems. In some cases, these collaborations have led to publications with people who were already friends, and in others they have led to the development of valuable and enjoyable new friendships. I am particularly grateful to the many people who have been willing to share their knowledge, and to devote their time to projects I find interesting.

The discovery that ZFNs can be used to make such specific mutations has a huge impact for the overall genetics research community. We asked Dr. Carroll how ZFNs are affecting the world of science now, and implications for the future.

Dana Carroll:

> People can now contemplate making very specific mutations in their genes of interest, which was only previously possible in yeast and some simple organisms, and in mice. With ZFN technology, this targeting capability is available for lots of different organisms. So that's been a big change for geneticists, but the field is still expanding. The number of applications is still increasing as we learn more about how we can use them. And so it's still in the growth phase, which is very exciting.

As we discussed the possible applications for zinc fingers, Dr. Carroll told us that he talks to ZFN users across the world, and commented on some of the work that has been done in the different organisms.

"People can now contemplate making very specific mutations in their genes of interest, which was only previously possible in yeast and some simple organisms, and in mice."

– Dr. Carroll

Dana Carroll:

People are starting to work on larger agricultural animals like pigs and cows, to study disease and to make genetic modifications. For example, pigs are used as a source of organs for human transplantation, and work is now being done to engineer transplant organs to make them more compatible with humans. These organs have cell surface products that are immunogenic in humans, and so (the) knockout of these markers could reduce the frequency of organ rejection.

Gene therapy is another area where ZFNs are undergoing initial testing. Right now, the only practical way to use the ZFNs for human therapy is to treat cells outside the body, then put them back. Sangamo Biosciences Inc. has a clinical trial in the wings for treatment of T-cell precursors that can be reintroduced to the body after they've had the CCR5 gene knocked out, making them resistant to infection by HIV. Until such clinical trials have really been done extensively, we just don't know what the unexpected effects really are. I'm sure that the people at Sangamo are holding their breath.

Sigma® Life Science acquired a license from Sangamo Biosciences to develop and distribute ZFNs in 2007. We had to ask … what does Dr. Carroll feel is Sigma's role in the use of the ZFN technology?

Dana Carroll:

I recently spoke at a Sigma Life Science-sponsored Targeted Genome Editing workshop, and I was impressed with a few things, including the number of people at the workshop, and the way that Sigma gave people the chance to try things out in the lab, doing real experiments.

Sigma's commitment to ZFN technology has also amazed me; the quality of the scientists who were working on it, and the way they have adapted the Sangamo technology by keeping in constant contact and finding the best way to make zinc finger sets. I was also impressed with their ambition to stay ahead of the curve … making cell lines and animals that people will want. By the time the ability to make a knockout rat using ZFNs had been published, Sigma had already created a facility to create rats for people who want them (SAGE™ Labs).

Until Sigma got into the game, the only way to get access to Sangamo's databases and assays was to collaborate with them. As a result, they were overwhelmed with requests, and had to be selective in whom they collaborate with. Sigma Life Science making ZFNs available to everybody really is a great contribution to the field. ■

To listen to Dr. Dana Carroll and his lab speak about their work, **visit wherebiobegins.com/biounity.**

YOUR FAVORITE GENE:

making connections in bio logical

favorite gene

Your Favorite Gene powered by Ingenuity has always allowed you to make connections to your biology through dynamic interaction networks and pathways. Based on your feedback, Your Favorite Gene now includes a biologically relevant literature search, new gene regulation viewers, expression study results, clinical trials, and identifies biochemical compounds related to your gene.

Save time searching biological content

Find products and information of interest via user-friendly queries on gene names or IDs, disease, function, species, tissue, pathways, or other contextual information. Use our unique gene regulation viewer to see promoter, conservation, CpG island, methylation, transcription factor, protein-protein interaction, and miRNA data all in one spot. With our latest update, you can access comprehensive data from across the Web in one location.

Effectively design your experiments

Use the interactive, up-to-date pathways and biological and chemical network diagrams to understand upstream and downstream relationships to optimize your experimental design. You can quickly understand alternative product and experimental design options without having to search multiple sites.

Dynamic biological pathways and interaction networks

Gain insight on your gene's interactions by viewing detailed biological pathways and molecular networks curated from the scientific literature occurring between molecules belonging to different biochemical families such as proteins, nucleic acids, lipids, and carbohydrates.

- Search for your gene of interest (i.e., EGFR) and view its interactors.
- View relationships between chemicals or drugs, groups or complexes, along with growth factors, ion channels, kinases, and more.
- Identify different relationships based on interaction types, bindings, inhibitors, and many more.
- Identify products relevant to your gene of interest by simply clicking on one of the many associations.

Filtering by biological relevance

In our latest update you can apply ontological filters to search literature, clinical trials and expression results by relevancy to your own unique set of terms. The tag clouds provide a view of how your set of terms connects to terms in other published work and allow you to uncover associations to other work.

Your product in context to your gene of interest

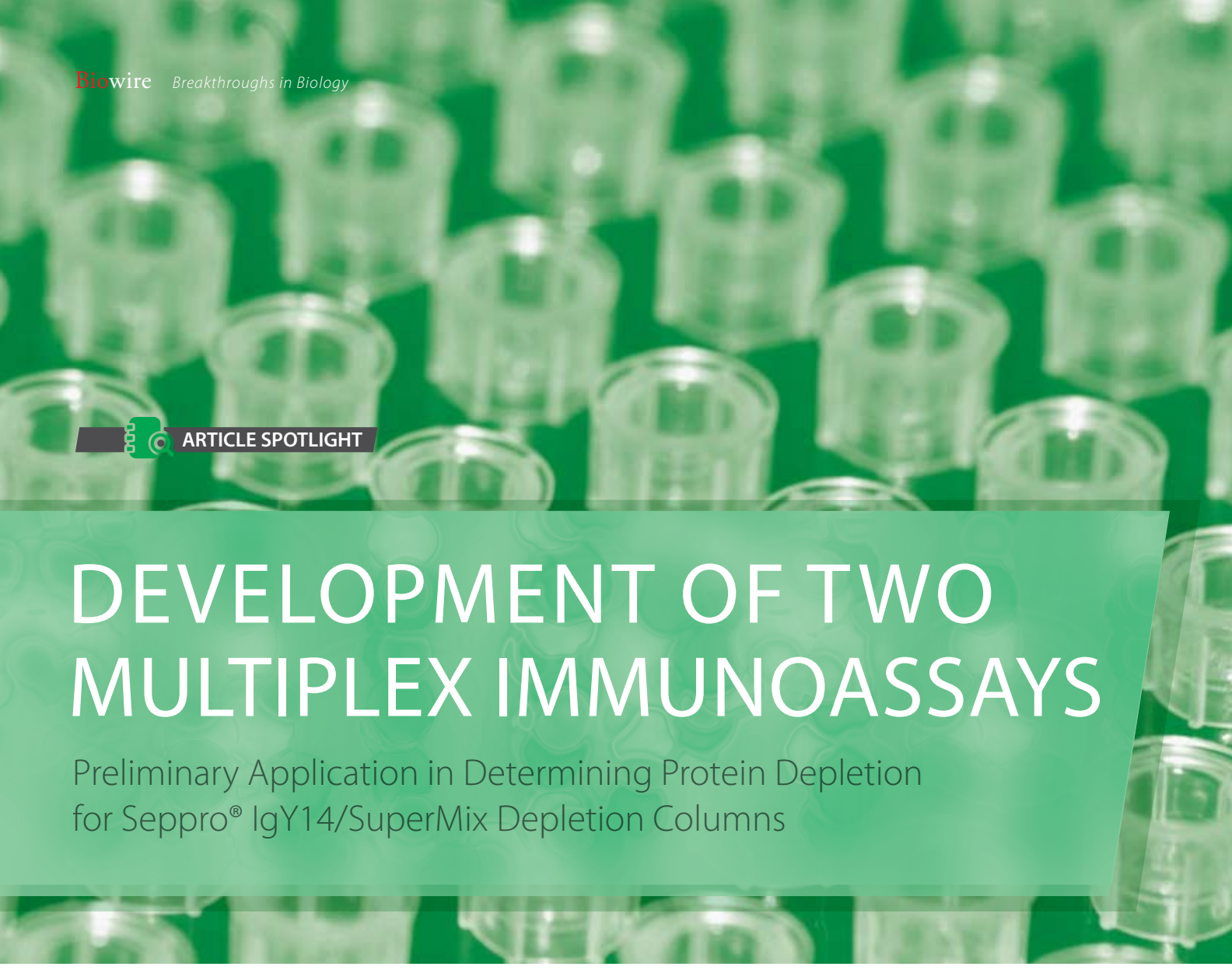
The Product List tab organizes all of the products available to study your gene in one spot for quick identification. Easily find what you need to realize your research. You can also use the filters embedded within both the interaction networks and biological pathways to find products in the context of the relationship you are investigating. The Literature tab also lists products cited in materials and methods.

Your Favorite Gene powered by Ingenuity provides easy access to comprehensive information and thousands of research products associated with your gene of interest.

Spend less time searching multiple sites and find it all on Your Favorite Gene, and best of all, it's free.

Research your favorite gene at
wherebiobegins.com/yfg

Your Favorite Gene SEARCH
Powered By INGENUITY



ARTICLE SPOTLIGHT

DEVELOPMENT OF TWO MULTIPLEX IMMUNOASSAYS

Preliminary Application in Determining Protein Depletion for Seppro® IgY14/SuperMix Depletion Columns

Jiaxin Dong, Chris Melm, Holly Chapman, Farrah Fan, Dian Er Chen and Henry Duewel
Protein Technology and Assay, Research Biotechnology, Sigma-Aldrich Corporation

Abstract

Depletion of Highly Abundant Protein (HAP) and Moderately Abundant Protein (MAP) from plasma or serum samples can be achieved through Sigma Seppro® Tandem IgY14/ SuperMix columns.

A range of 77–129 MAPs has been identified from the bound fraction of Seppro SuperMix resin by LC/MS-MS method. However, the efficiency and lot-to-lot consistency of MAP depletion using SuperMix resin has not been determined and requires a robust method to monitor. By taking advantage of the Luminex® technology, we have developed two multiplex immunoassays to monitor protein targets of depletion or retention. The MAP-10PLEX assay can simultaneously quantify 10 MAPs (Ceruloplasmin, C4, Plasminogen, IgD, C1q, Antithrombin III, Hemopexin, α-Antichymotrypsin, CRP and Prealbumin) from plasma or serum samples depleted utilizing IgY14/ SuperMix columns. The LAP4PLEX assay, which is used to track loss of low abundant proteins (LAP) during depletion, is able to simultaneously quantify Adiponectin, soluble L-selectin, soluble ICAM-1 and TIMP-1. The preliminary assay data from the R & D, Operations and QC departments on Seppro IgY-SuperMix depleted samples has suggested the two assays to be accurate, fast and cost-saving methods to monitor depletion consistency from lot-to-lot of the SuperMix resin.

Introduction

An avian IgY based immunodepletion technology is used for Sigma Seppro depletion products. The Human IgY14 Columns can remove 14 highly abundant proteins (HAP) and the unique Human SuperMix System is used to remove moderately abundant proteins (MAP). The combination of IgY14/SuperMix columns can deplete approximately 96–99% of total protein mass from human serum or plasma (**Figure 1**).

While the depletion of 14 HAPs has been well established, the further removal of MAPs is less well characterized. Previously, a range of 77-129 MAPs has been identified from the bound fraction of Seppro SuperMix resin by LC/MS-MS. However, the degree of specific depletion of MAPs as well as the non-specific impact on low abundant proteins (LAP) is still unclear. This also makes tracking the lot-to-lot depletion consistency of SuperMix products a challenge.

Regular ELISA methods are time-consuming and commercial ELISA kits are costly. Therefore, finding a fast and economical alternative analysis has appeared to be important. Luminex xMAP technology is a widely used multiplex platform to quantify multiple analytes from biological fluid samples. It combines the accuracy of ELISA and multiplicity of antibody arrays through protein-antibody interaction on the surface of microspheres and is characterized by speed and minimal consumption of samples (**Figure 2**). In order to evaluate Seppro SuperMix resins in the R & D department and to monitor the lot-to-lot consistency of materials produced by Operations in the QC department, we developed two Luminex immunoassays to simultaneously quantify 10 representative MAPs (Ceruloplasmin, C4, Plasminogen, IgD, C1q, Antithrombin III, Hemopexin, α-Antichymotrypsin, CRP and Prealbumin) and 4 representative LAPs (Adiponectin, soluble L-selectin, soluble ICAM-1 and TIMP-1) from depleted plasma via Seppro SuperMix columns.

Materials and Methods

1. Instrument: Luminex® 100 (**Figure 2**), Luminex Corp., Austin, Texas, USA
2. Vacuum pump, 96-well microtiter filter plates and filtration manifold (Millipore Corp.)
3. Capture antibodies were immobilized on Luminex® beads, detection antibodies were biotinylated before use, streptavidin-phycoerythrin conjugate (Sigma Cat. No. E4011)
4. Data were calculated using STATLIA® software with a 5-parameter logistic curve-fitting method (Brendan Scientific, CA).
5. Depletion procedure: 50 or 100 µL human plasma (Sigma P9523) was loaded to Seppro IgY14-LC5/SuperMix-LC2 columns with tandem connection (**Figure 3**). Depletion was run in Waters 2695 HPLC System (Alliance). After collection of flowthrough from combined columns (SuperMix depleted sample), bound proteins were eluted from IgY14-LC5 and SuperMix LC2 column separately.
6. Multiplex assay protocol: 50 µL standards or samples are mixed with 25 µL antibody conjugated bead mixture in 96-well filter plate. Shake 2 hr at RT or overnight at 4ºC and wash plate twice. Add 50 µL/well detection mixture, shake 1hr at RT and wash plate once. Add 50 µL/well SAPE, shake 30 min at RT and wash plate twice. Add 100 µL/well sheath fluid and shake 2–5 min at RT. Read plate on Luminex Instrumentation.

Figure 1. Seppro SuperMix Depletion System

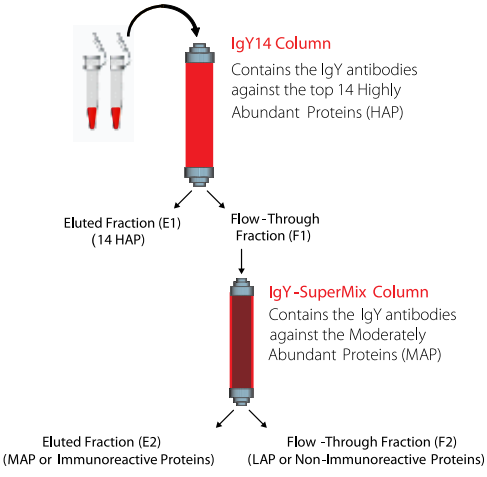
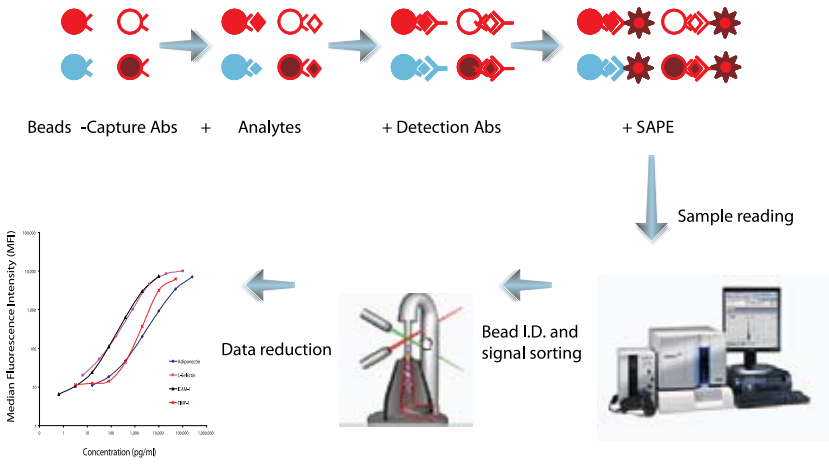


Figure 2. Principle of Luminex Multiplexing Assay



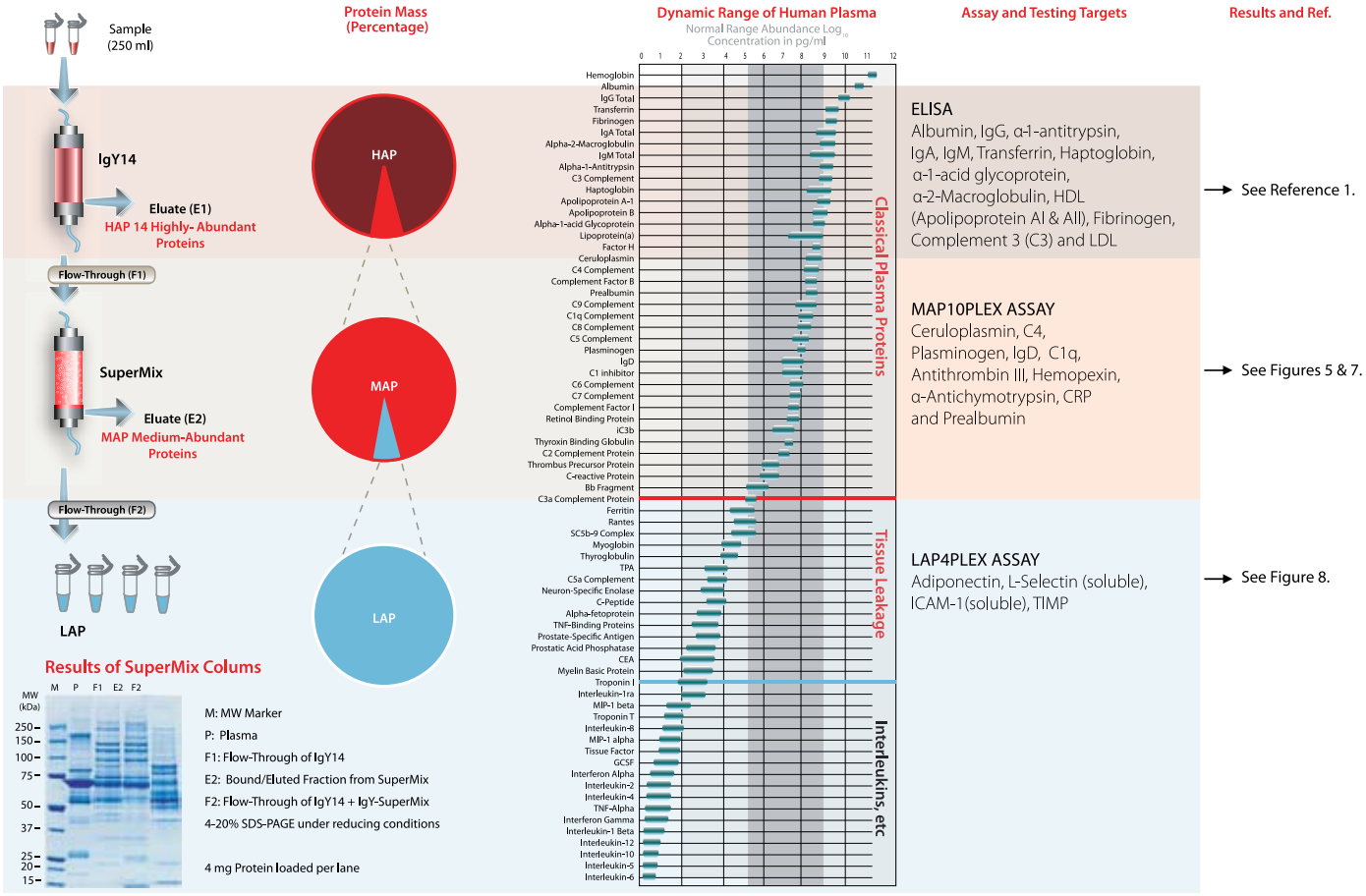


Figure 3. Flowchart of Seppro® SuperMix Depletion Process and the Targets to Test

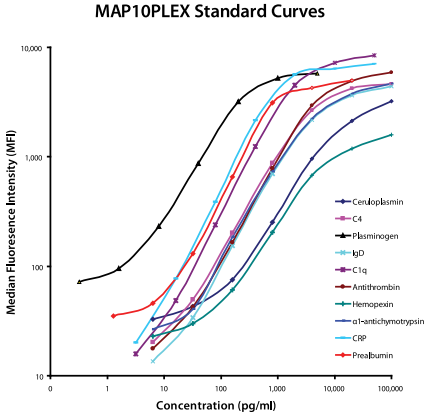
The top 14 HAPs (plasma level ranges in 0.5-50 mg/ml) compose 95% of total plasma proteins. The 10 targets to test for SuperMix depleted plasma cover a wide range concentration (2.3 µg/ml for CRP, 0.5 mg/ml for Ceruloplasmin) of MAPs. The 4 LAP targets in normal human plasma ranges in 0.2-10 µg/ml.

MAP10PLEX										
Final dilution of plasma (folds)	Ceruloplasmin	C4	Plasminogen	IgD	C1q	Antithrombin III	Hemopexin	α1-antichymo-trypsin	CRP	Prealbumin
512000	103.73	74.21	153.77	112.98	125.17	127.80	104.88	115.24	86.50	95.99
256000	105.03	69.44	137.89	99.57	11.08	109.75	97.32	99.54	102.85	89.93
128000	105.01	68.59	120.09	114.09	103.57	101.52	104.10	93.74	110.09	92.18
64000 (start)	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
AVG	104.59	70.74	137.25	108.88	113.27	113.02	102.10	102.84	99.81	92.70

LAP4PLEX				
Final dilution of plasma (folds)	Adiponectin	L-Selectin	ICAM-1	TIMP-1
2000	140.59	98.95	107.32	87.11
1000	127.67	96.97	118.63	96.45
500	114.96	93.20	108.45	99.48
250 (start)	100.00	100.00	100.00	100.00
AVG	127.74	96.37	111.47	94.35

Table 1. Assay Accuracy—Dilution Linearity for Plasma Samples

Dilution linearity is commonly used to evaluate assay accuracy that is one of the most important requirements for immunoassays. A serial dilution of a pooled plasma was tested with MAP10PLEX and LAP4PLEX assay. The numbers in the table represent percentage of concentrations at each dilution point against start point. Both assays perform well. For normal human plasma samples, dilutions in the range from 5×10^4 to 5×10^5 for MAP10PLEX assay or 250 to 2000 for LAP4PLEX assay are suggested.



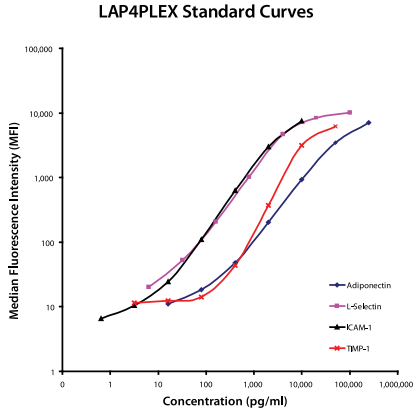


Figure 4. MAP10PLEX and LAP4PLEX assays

Performed with protocol described above and standard curves are plotted for each assay.

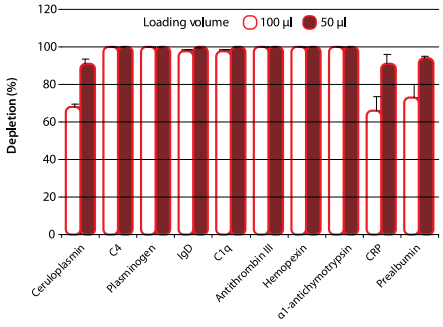


Figure 5. Depletion Efficiency of Seppro SuperMix Columns Tested by Luminex MAP10PLEX Assay

Depletion efficiency of 10 representative MAPs is related to loading volume of plasma samples. >90% of all 10 MAPs can be depleted from 50 µL human plasma.

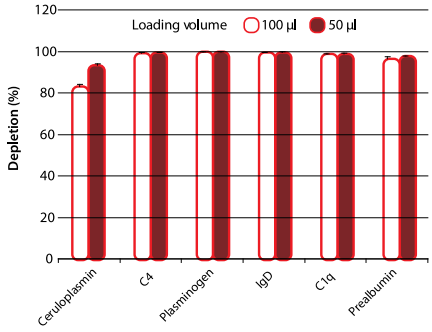


Figure 6. Depletion Efficiency of Seppro SuperMix Columns Tested by Indirect ELISA (for 6 targets only)

The same set of SuperMix depleted samples were tested by indirect ELISA and displayed a comparable result as Luminex MAP10PLEX assay.

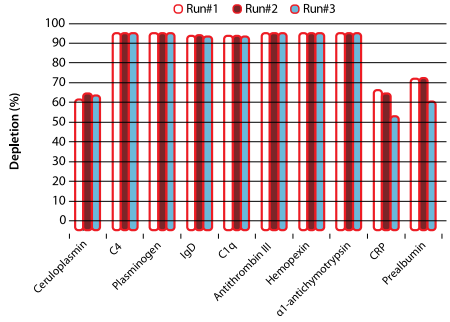


Figure 7. Run-to-Run Depletion Consistency of SuperMix Columns

Three depletion runs were performed with 100 µL plasma utilizing Seppro IgY14 LC5/SuperMix LC2 columns and the depletion efficiency was determined by Luminex MAP10PLEX assay. Depletion of all 10 targets is consistent (CV%<15%)

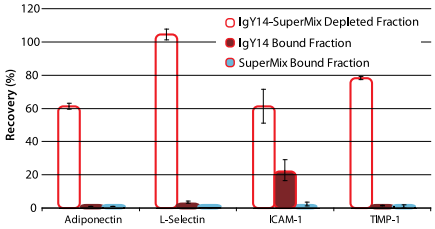


Figure 8. Impact of SuperMix Depletion Process on LAPs

Four representative LAPs were tested by Luminex LAP4PLEX assay to monitor the loss of LAPs during SuperMix depletion process. Depending on different biological properties of the individual target, approximately 1-40% loss has been observed from SuperMix depleted plasma (Note: some proteins in bound fraction are unstable).

Conclusion

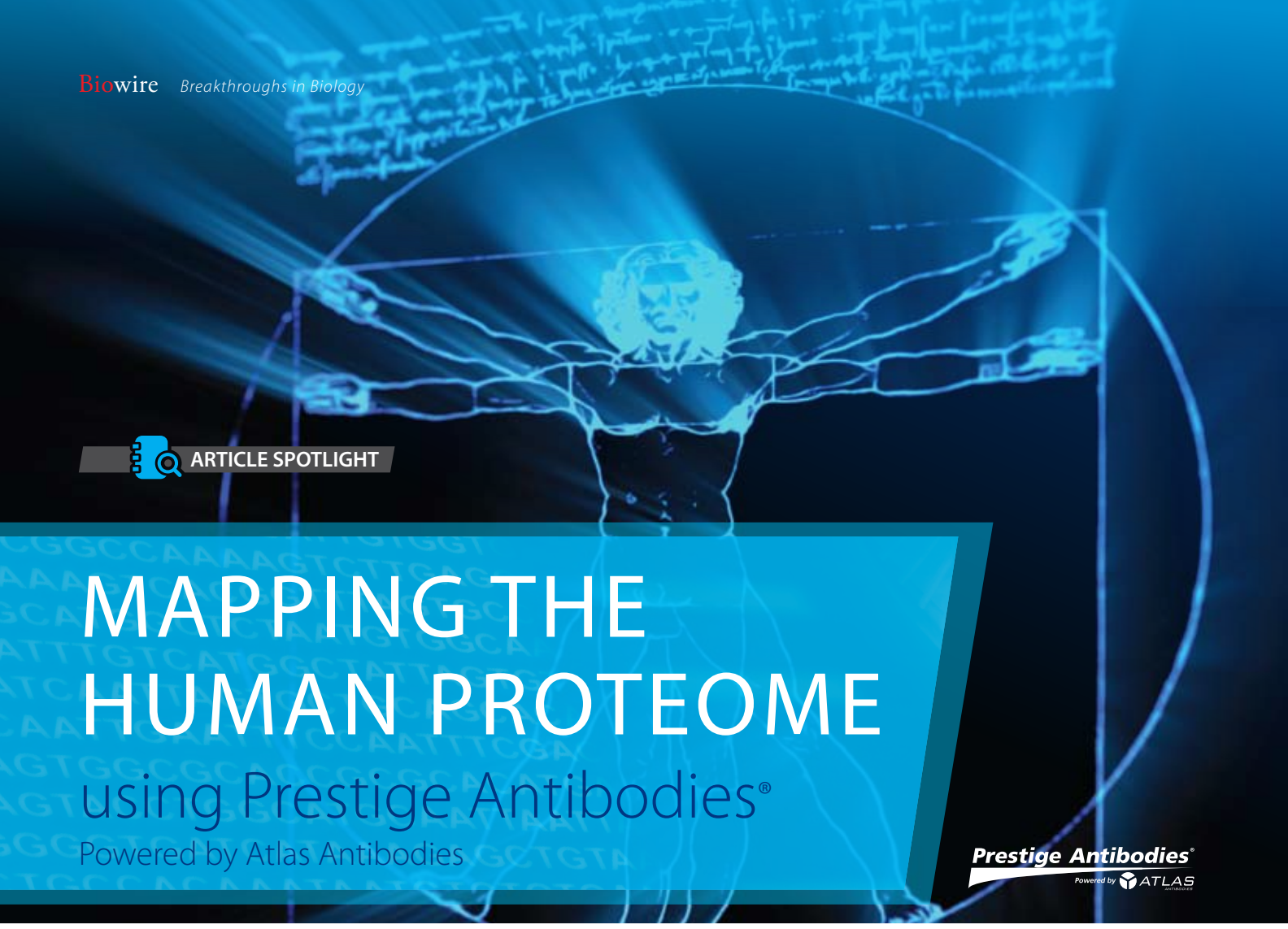
1. The MAP10PLEX and LAP4PLEX assays are able to simultaneously measure 10 MAPs or 4 LAPs from human plasma or serum samples within 4 hours.
2. These two multiplex assays can be used to replace regular ELISA methods to determine MAP depletion and LAP retention from Seppro® IgY-14/SuperMix depleted samples.
3. The preliminary test indicated that depletion of 10 MAPs from human plasma through the tandem IgY14/SuperMix columns were reproducible and is consistent with Dr. Smith's result based on spectral count data (Ref. 3).
4. In order to remove > 90% of MAPs from plasma, columns with equal volume of IgY14 and SuperMix resins (IgY14/SuperMix ratio = 1:1) should be combined.

Summary

The newly developed MAP10PLEX and LAP-4PLEX assays using Luminex technology have been proven to be both time-and cost-saving solutions to monitor the robustness of protein depletions from lot-to-lot and sample-to-sample.

References

(1) Seppro® Depletion Product Info, sigmaaldrich.com
(2) Technical Bulletin, Luminex Corp. luminexcorp.com, (2001).
(3) Qian W.J., et al., *Molecular & Cellular Proteomics*, **10**, 1963-75 (2008).



ARTICLE SPOTLIGHT

MAPPING THE HUMAN PROTEOME

using Prestige Antibodies®

Powered by Atlas Antibodies

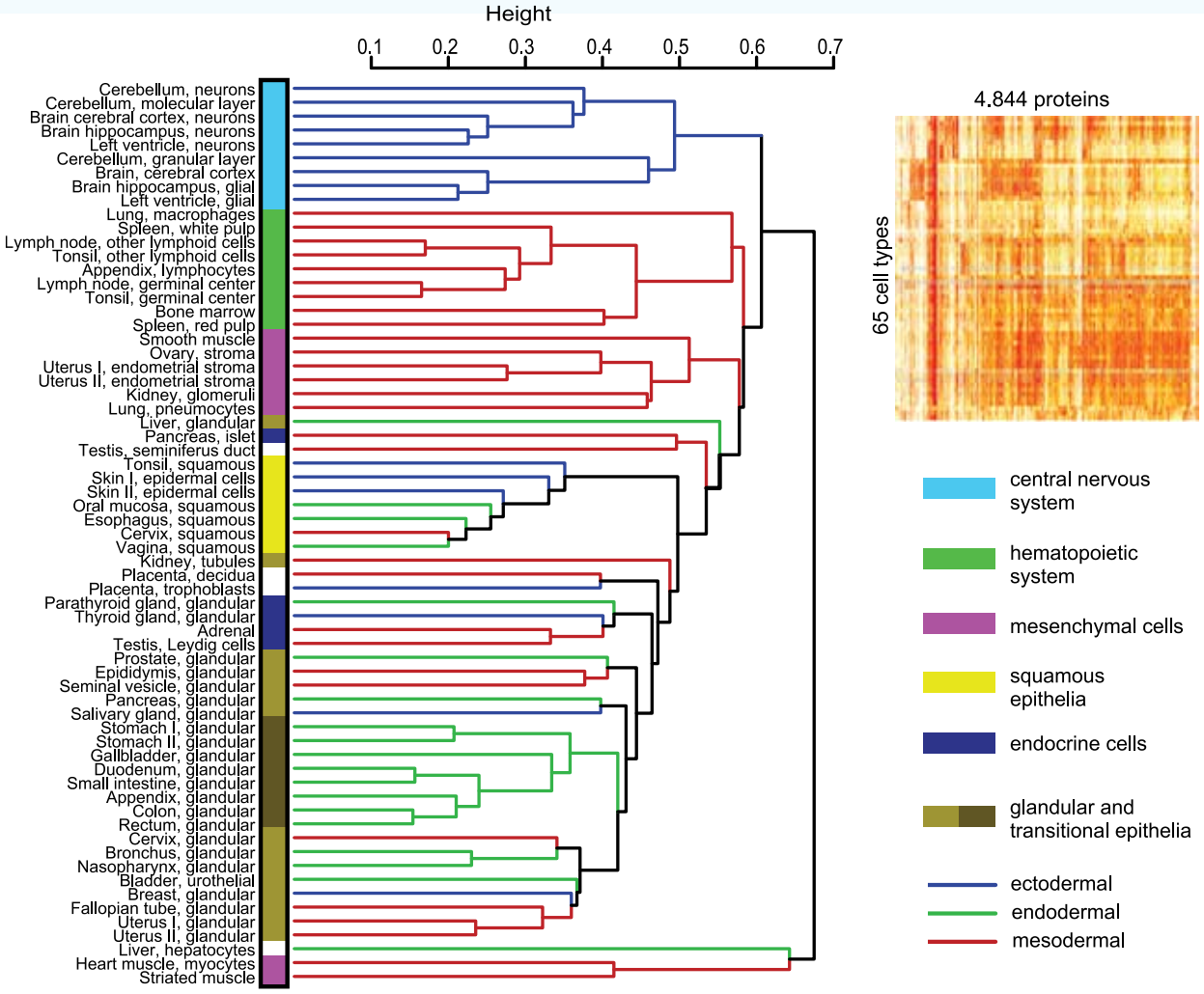
Prestige Antibodies® Powered by ATLAS ANTIBODIES

Prestige Antibodies and the Human Protein Atlas

For the first time ever, a comprehensive localization analysis has been performed on 8,000+ human proteins in a standardized set-up manner[1,2]. The analysis has been performed in several levels using different methodologies. All results are presented on www.proteinatlas.org, a portal run by the "Human Protein Atlas" (HPA) project. The HPA project is an international 10-year research effort that aims to map the full human proteome and present the data on a public web portal to researchers all over the world. The Human Protein Atlas portal contains expression profiles for proteins corresponding to over 40% of the human proteome. Each year characterization data of ~2,500 new proteins are added. By the end of 2014, a first draft of the localization of the full human proteome will be available.

The HPA project involves over 100 researchers working full time on developing and characterizing HPA antibodies for protein expression and localization studies[1,2]. Late 2010, the project will have mapped 50% of all human proteins by the use of HPA antibodies. About 20 new HPA antibodies are generated each day, and successfully validated antibodies are made available to the research community under the brand name Prestige Antibodies.

The results presented on the HPA portal can be used to study and compare tissue distribution of proteins from immunohistochemistry (IHC) images in human normal, as well as diseased, tissues and cells. In addition, subcellular localization information from immunofluorescence (IF) analysis of 3,500+ proteins can be viewed and compared. For example, IHC images from over 8,000 proteins revealed a highly ubiquitous expression of most proteins with a minority of the proteins detected in a single cell type[3]. These cell- or tissue-specific proteins may be interesting starting points for further studies, facilitated by the publicly available annotation results and the underlying original images on the Human Protein Atlas.



By studying protein expression profiles of tissues and organs, the unique characteristics of the various cell types in the human body can be elucidated.

Protein expression profiles for 4,842 proteins in 48 human tissues and 45 human cell lines have recently been defined and investigated[3]. More than 2 million high resolution IHC images have manually been annotated by certified pathologists who have determined the relative level of expression for each specific protein.

A hierarchical clustering analysis was carried out based on the protein levels in 65 normal cell types. The analysis showed that cells cluster into groups that could be expected on the basis of traditional embryology, histology and anatomy, and most of the cells cluster into the six major groups shown in **Figure 1**.

Figure 1.

Global protein expression profiling in 65 normal human cell types using 5,934 antibodies corresponding to 4,842 proteins. The dendrogram was constructed by hierarchical clustering; the inset shows the original heatmap. Dendrogram bars are colored according to embryonic origin, ectoderm (blue), mesoderm (red) and endoderm (green). The cell types were classified into six categories according to the color code in the figure.

Tissue specific protein expression

The results from an analysis of the fraction of cell types expressing each protein is shown in **Figure 2A**. The analysis indicated that a high fraction of the proteome is expressed in most cell types and tissues and very few proteins are expressed in a single cell type. **Figure 2B** shows the fraction of proteins that are expressed in a specific cell type. An average of 68% of all proteins expressed were present in any given cell type. **Figure 2C** shows an example of a previously non-characterized cell-type specific protein, angiomoitin-like protein 1 (AMOTL1). Immunohistochemical staining using anti-AMOTL1 (Prestige Antibody HPA001196) in placental tissue shows distinct expression in basal cytotrophoblasts and in the brush-border membrane overlying syncytiotrophoblasts.

Global protein expression analysis using confocal microscopy

Subcellular analysis on 466 proteins in three human cell lines has been performed using Prestige Antibodies® and confocal microscopy[4]. Similarly as for the IHC analysis, the results showed a ubiquitous expression pattern of the proteins with more than 70% of the analyzed proteins detected in the three cell lines. **Figure 3A, B, C** show the distribution of localization in three major compartments for the three cell lines and a more detailed overview of the subcellular localizations. In **Figure 3D**, only antibodies specific to one organelle are included. Out of the 466 proteins investigated, about 200 showed a strict localization, 200 were present in multiple cellular compartments, and about 70 were considered unspecific (or negative).

Discover more at

wherebiobegins.com/prestige

Figure 2A.

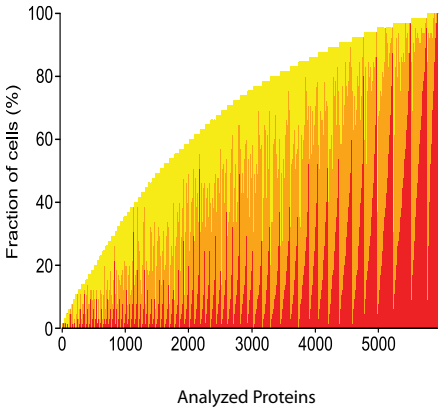


Figure 2B.

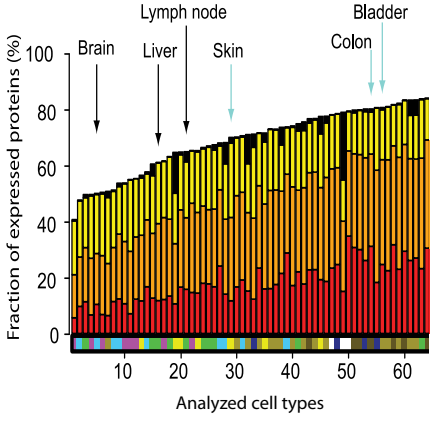
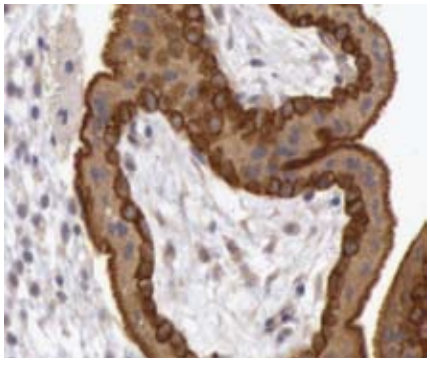


Figure 2C.



Immunohistochemistry

Figure 2.

Tissue specific protein expression. Expression intensity is color-coded as follows: red=strong, orange=medium, yellow=weak. **(A)** The fraction (%) of cells in which a particular protein was expressed. Cell-type specific proteins are arranged to the left and "housekeeping proteins" to the right. **(B)** The fraction (%) of proteins expressed in a specific cell type. The bar at the bottom display the six major categories of cell types described in Figure 1 (black represents missing data). **(C)** Immunohistochemistry staining of angiomoitin-like protein 1 (anti-AMOTL1, HPA001196) in human placental tissue.

Summary

- The analysis revealed a highly ubiquitous expression of the majority of the proteins with very few proteins detected in a single cell type.
- These studies suggest that tissue and cell specificity is achieved by precise regulation of protein levels in space and time and not by whether proteins are expressed or not.
- The Human Protein Atlas portal can be used for global protein expression analyses.
- All Prestige Antibody annotation results and the underlying original images are available as a public resource in the Human Protein Atlas portal at www.proteinatlas.org.
- The Human Protein Atlas portal has an advanced search function to allow complex queries, including combined searches based on protein classes, chromosomal location and/or tissue profiles.

References:

(1) A gene-centric human protein atlas for expression profiles based on antibodies. Berglund, L., et al., *Molecular & Cellular Proteomic*, **7**, 2019-2027, (2008).

(2) A human protein atlas for normal and cancer tissues based on antibody proteomics. Uhlén, M., et. al., *Molecular & Cellular Proteomics*, **12**, 1920-1932 (2005).

(3) A global view of protein expression in human cells, tissues, and organs. Pontén, F., et al., *Molecular System Biology*, **5**, 337 (2009).

(4) Toward a confocal subcellular atlas of the human proteome. Barbe, L., et. al., *Molecular & Cellular Proteomics*, **3**, 499-508 (2008).

Figure 3A.

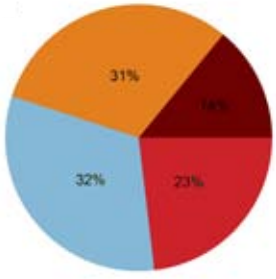


Figure 3B.

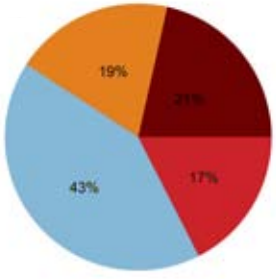


Figure 3C.

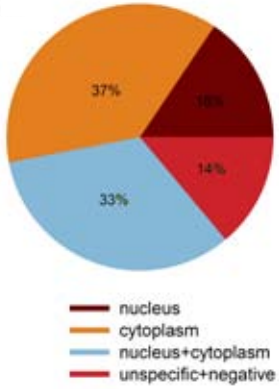


Figure 3D.

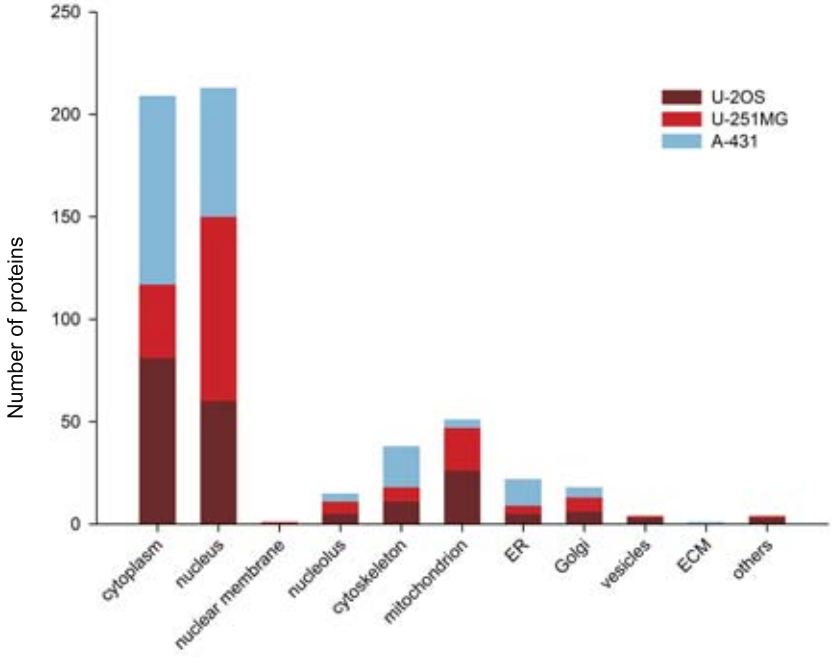


Figure 3.

Localization distribution into broad compartments for U-2OS **(A)**, U-251MG **(B)** and A-431 **(C)** cell lines. **(D)** More detailed subcellular distribution in U-2OS, U-251MG and A-431. (ER, endoplasmic reticulum; ECM, extracellular matrix).

ARTICLE SPOTLIGHT

PHOSPHATASE INHIBITOR COCKTAIL

A key additive in protein phosphorylation research

Ian Gleiser, Pnina Yaish, Efrat Barnea-Gedalyahu, Dorit Zharhary
Research & Development Dept., Sigma-Aldrich Israel

Introduction

Phosphorylation is a reversible protein post-translational modification. Protein kinases catalyze the transfer of the negatively charged γ-phosphate group from ATP to a hydroxyl side chain of serine, threonine or tyrosine residues of the protein, changing the protein's conformation and activity. The reverse reaction, phosphate removal from the protein, or dephosphorylation, is performed by protein phosphatases. Three to four percent of mammalian proteins are kinases and phosphatases,[1,2] some being specific for a few target proteins while others act broadly on many proteins.

Protein phosphorylation is a key regulatory mechanism that controls protein activity, interactions, localization and degradation. It controls processes of signal transduction, cell cycle, apoptosis, metabolism and others. About one-third of the proteins present in a typical mammalian cell are phosphorylated, many of them on multiple sites. Serine phosphorylation accounts for 86% of the phosphoproteome, while threonine and tyrosine phosphorylation are 12% and 2% respectively. Numerous human diseases are associated with abnormal phosphorylation of cellular proteins.[3-5]

There are many advantages for a cell to utilize phosphorylation/dephosphorylation of a protein as a control mechanism. It is rapid, does not require the synthesis of new proteins or degradation of proteins, and is a reversible reaction.

When studying cellular processes involving protein phosphorylation such as signaling cascades and protein-protein interactions, or analysis of phosphorylated sites on a protein and phosphoprotein purification, it is essential to inhibit cellular protein phosphatases by phosphatase inhibitors. This allows freezing of the phosphorylation state of the target protein/s at an elected time point. The phosphorylation database www.phosphosite.org reports 80,000 phospho-sites. The detection of most of them would not have been possible without the use of phosphatase inhibitors.

Protein Phosphatases

Protein phosphatases are divided to subcategories based on their substrate specificity (see **Table 1**).

1. Alkaline phosphatase[6-8] – A family of non-specific phosphatases, responsible for removing phosphate groups from many types of molecules, including proteins, nucleotides and alkaloids. Mammalian alkaline phosphatase isoenzymes are inhibited by homoarginine and levamisole analogs. However, the intestinal and placental isoenzymes are not inhibited by levamisole but are inhibited by imidazole.

2. Protein Serine/Threonine phosphatase [9,10] – This class of phosphoprotein phosphatases accounts for the majority of Ser/Thr phosphatase activity *in vivo*. It includes two major subclasses, PP1 and PP2. The latter is further subdivided based on metal-ion requirement: PP2A that does not require metal ion; PP2B which is calcium stimulated; and PP2C that is Mg^{2+} dependent. Ser/Thr phosphatases are inhibited by many small molecules from marine sponges, soil streptomyces and others. The best known are okadaic acid, calyculin A, microcystin-LR, tautomycin, fostriecin and cantharidin. The activities of these compounds have different specificities towards various Ser/Thr phosphatases.[11]

3. Protein Tyrosine phosphatase[12] – A group of enzymes that remove phosphate groups from phosphorylated tyrosine residues on proteins, using a cysteinyl-phosphate enzyme intermediate. These enzymes are key regulatory components in signal transduction pathways. Tyrosine phosphatases are inhibited by orthovanadate and related compounds and by sodium fluoride.

4. Dual specificity (Tyr and Ser/Thr) phosphatases[3] – Dual-Specificity phosphatases are a subclass of protein tyrosine phosphatases that are also able to dephosphorylate both serine and threonine residues. They are involved in the regulation of key cell signaling pathways. Specific examples of these inhibitors can be found in reference.[13]

Table 1. Types of phosphatases present in animal tissues and cell cultures

Phosphatase	pH spectrum	Inhibitors
Alkaline phosphatases[6-8]	8-11	Bromotetramisole,[14] levamisole, vanadate,[15] imidazole[16]
Ser/Thr phosphatases[9,10] (i.e., PP1 and PP2A)	7-8	Okadaic acid,[17] Microcystin LR Cantharidin, Inhibitor-2
Ser/Thr phosphatases[9,10] - Ion dependent (i.e., PP2B and PP2C)	7-8	EDTA, Okadaic acid, cyclosporin, FK-506[18]
Tyrosine phosphatases[12]	7-8	Vanadate,[18] peroxovanadium compounds[19]
Dual specificity phosphatases[3]		Vanadate[18]

"When studying cellular processes involving protein phosphorylation such as signaling cascades and protein-protein interactions, or analysis of phosphorylated sites on a protein and phosphoprotein purification, it is essential to inhibit cellular protein phosphatases by phosphatase inhibitors."

Phosphatase Inhibitor Cocktails

The use of phosphatase inhibitors is critical for all types of phosphorylation studies. Sigma® Life Science provides a line of phosphatase inhibitor cocktails that cover the widest range of phosphatases to ensure protection of proteins from dephosphorylation. The line contains 2 cocktails: Phosphatase Inhibitor Cocktail 2 (Cat. No. P5726) and Phosphatase Inhibitor Cocktail 3 (Cat. No. P0044). Phosphatase Inhibitor Cocktail 3, Sigma's new phosphatase inhibitor cocktail, is a mixture of phosphatase inhibitors directed towards Ser/Thr phosphatases. Phosphatase Inhibitor Cocktail 3 replaces Phosphatase Inhibitor Cocktail 1 (Cat. No. P2850). The difference between the two cocktails is that the new Phosphatase Inhibitor Cocktail 3 contains Calyculin A instead of microcystin-LR in Phosphatase Inhibitor Cocktail 1. **Table 2** provides the list of inhibitors in each cocktail.

Table 2. Phosphatase Inhibitor Cocktail components

Phosphatase Inhibitor Cocktail 1 Cat. No. P2850	Phosphatase Inhibitor Cocktail 2 Cat. No. P5726	Phosphatase Inhibitor Cocktail 3 Cat. No. P0044
Cantharidin	Sodium orthovanadate	Cantharidin
(-)-p-Bromotetramisole	Sodium molybdate	(-)-p-Bromotetramisole
Microcystin LR	Sodium tartrate	Calyculin A
	Imidazole	

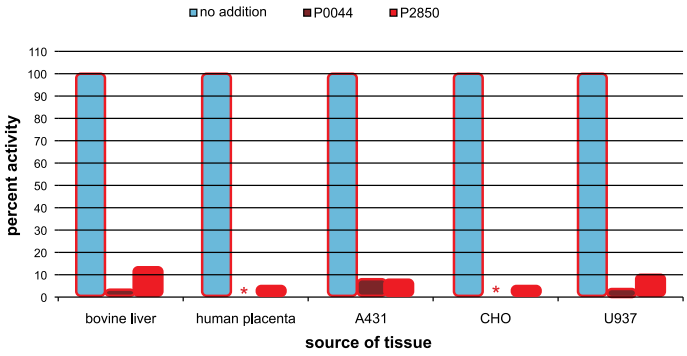


Figure 1. Inhibition of PP1α-like activity in cell and tissue extracts

Endogenous PP1α-like activity in various cell and tissue extracts was measured using the radioactive substrate ^{32}P-Ser phosphorylase A at pH 7.5, 30°C.[20] Phosphatase Inhibitor Cocktail 1 (Cat. No. P2850) or Phosphatase Inhibitor Cocktail 3 (Cat. No. P0044) were added to the extracts at a final concentration of 1% at 30°C, 5 minutes prior to assaying PP1α-like activity. The activities shown in **Figure 1** are relative and not absolute. The smaller the activity remaining, the greater the degree of inhibition by the cocktail that was placed in the reaction.

* The value observed for these samples is too low to be seen in the diagram and is very close to zero.

Experimental results

The following experimental data compares the inhibitory effect of the new Phosphatase Inhibitor Cocktail 3 and Phosphatase Inhibitor Cocktail 1 (Cat. No. P2850) which it replaces. The data shows that the activity of the two cocktails is equivalent.

It is important to mention that in order to inhibit a wide range of phosphatases and to effectively reduce total phosphatase activity, it is recommended to use both of Sigma's phosphatase inhibitor cocktails, Phosphatase Inhibitor Cocktail 2 (Cat. No. P5726) and Phosphatase Inhibitor Cocktail 3 (Cat. No. P0044).

Inhibition of Protein Phosphatase 1α-like activity

Phosphatase Inhibitor Cocktail 3 (Cat. No. P0044) is an effective inhibitor of Protein phosphatase 1α-like activity in cells and tissues. The following experiments demonstrate its inhibition efficiency and compare it to that of the former Phosphatase Inhibitor Cocktail 1 (Cat. No. P2850).

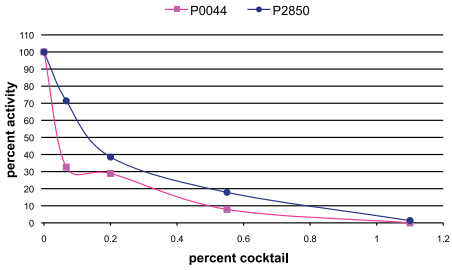


Figure 2. Dose dependent inhibition of PP1α-like activity in human placenta

Human placenta shows 10-20 folds more PP1α-like activity per mg protein than other tissue extracts. Therefore, it was chosen to demonstrate the inhibition efficiency of Phosphatase Inhibitor Cocktails.

Endogenous PP1α-like activity in human placenta extract was measured using the radioactive substrate ^{32}P-Ser phosphorylase A at pH 7.5, 30°C.[20] Different amounts of Phosphatase Inhibitor Cocktail 1 (Cat. No. P2850) or Phosphatase Inhibitor Cocktail 3 (Cat. No. P0044) were added to the extract at 30°C, 5 minutes prior to assaying PP1α-like activity.

Figure 3A

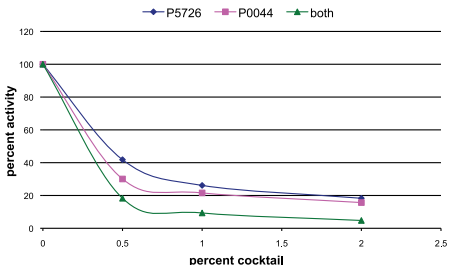
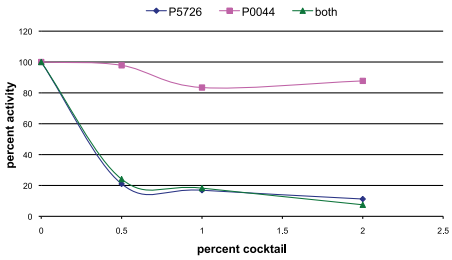


Figure 3B



Figures 3A & B. Dose dependent Inhibition of Alkaline Phosphatase (AP) like activity in bovine liver extract

Endogenous AP-like activity in bovine liver extract was measured in a colorimetric assay using pNPP as a substrate at pH 10.4, 37°C. Inhibition was performed by incubating the extract with Phosphatase Inhibitor Cocktail 3 (Cat. No. P0044) alone or in tandem with Phosphatase Inhibitor Cocktail 2 (Cat. No. P5726), for 3 minutes at 37°C prior to assaying AP-like activity.

Inhibition of Alkaline Phosphatase (AP) like activity

As mentioned above, different alkaline phosphatase isoenzymes are inhibited by different inhibitors. Phosphatase Inhibitor Cocktail 3 (Cat. No. P0044) strongly inhibits the L-isoforms of alkaline phosphatase present in extracts from bovine liver, and shows a synergistic effect when applied with Phosphatase Inhibitor Cocktail 2 (Cat. No. P5726) (see **Figure 3A**).

Phosphatase Inhibitor Cocktail 3 (Cat. No. P0044) has little effect on the inhibition of the P-isoforms of alkaline phosphatase activity, which are abundant in human placental extracts, and thus inhibition is achieved by the use of Phosphatase Inhibitor Cocktail 2 (Cat. No. P5726 (see **Figure 3B**).

References

(1) Johnson, L.N., *Biochemical Society Transactions*, **37**, 627-641 (2009)
(2) Moorhead, G.B.G., et al., *Biochemistry Journal*., **417**, 401-409 (2009).
(3) Shchemelinin, I., et al., *Folia Biologica (Praha)*, **52**, 81-101 (2006).
(4) Olsen, J.V., et al., *Cell*, **127**, 635-648 (2006).
(5) Graves, J.D., and Krebs, E.G., *Pharmacology and Therapeutics*, **82**, 111-121 (1999).
(6) Moss, D.W., *Clinical Chemistry*, **38**, 2486-2492 (1992).
(7) Onsgard-Meyer, M., et al., Proceedings of the Society for Experimental Biology and Medicine, **213**, 193-195 (1996).
(8) Crofton, P.M., *Critical Reviews in Clinical Laboratory Sciences*, **16**, 161-194 (1982).
(9) Barford, D., *Trends in Biochemical Sciences*, **21**, 407-412 (1996).
(10) Wera, S., and Hemmings, B.S., *Biochemistry Journal*, **311**, 17-29 (1995).
(11) Swingle, M., et al., *Methods in Molecular Biology*, **365**, 23-38 (2007).
(12) Wang, W-Q., et al., *Current Topics in Medicinal Chemistry*, **3**, 739-748 (2003).
(13) Pestell, K.E., et al., *Oncogene*, **19**, 6607-6612 (2000).
(14) Van Belle, H., et al., *Clinical Chemistry*, **23**, 454-459 (1977).
(15) Chakrabartty, A., and Stinson, R.A., *Biochimica et Biophysica Acta*, **839**, 174-180 (1983).
(16) Brunel, C., and Cathala, G., *Biochimica et Biophysica Acta*, **268**, 415-421 (1972).
(17) Takai, A., and Mieskes, G., *Biochemistry Journal*, **275**, 233-239 (1991).
(18) Lum, H., et al., *American Journal of Physiology - Lung Cellular and Molecular Physiology*, **281**, L546-L555 (2001).
(19) Posner, B.I., et al., *Journal of Biological Chemistry*, **269**, 4596-604 (1994).
(20) Cohen, P., et al., *Methods in Enzymology*., **159**, 390-408 (1988).

ARTICLE SPOTLIGHT

TRANSPLEX® WHOLE TRANSCRIPTOME AMPLIFICATION

TransPlex® Whole Transcriptome Amplification of RNA from Low Cell-Number Stem Cell Samples

Ken Heuermann and Brian Ward
Sigma-Aldrich Corporation

Abstract

The efficacy of amplification of small quantities of total RNA with the TransPlex® Complete Whole Transcriptome Amplification Kit (WTA2) was examined in this study. Total RNA extracted from decreasing numbers of FACS-isolated bone marrow stem cells (10-, 100-, and 1000-cells samples) was amplified with the TransPlex WTA2 kit. A call rate of 58.8% of unique biological array features was observed for the 10-cell vs. 100-cell microarray analysis, with a similar call rate of 61.46% for 10-cells vs. 1000-cells. Greater than 90% commonality existed between the intersecting data sets for the two analyses.

After a more stringent screening ($p = 0.0001$), the 10-cell vs. 100-cell comparison revealed 5568 intersecting features. The comparable analysis for 10 cells vs. 1000 cells resulted in 4977 features, with 3862 features common to both comparisons. In addition, the effect of decreasing RNA input on amplification efficiency was examined. Results indicated that an adjustment to the library synthesis primer concentration allowed for maintenance of linear amplification and representative qPCR at low RNA input quantities. This adjustment proved to be critical for downstream qPCR applications, but not for the case where the amplification product was used as microarray target. This study confirms that the TransPlex Complete Whole Transcriptome Amplification Kit can effectively amplify low input quantities of RNA, approaching the single cell level.

Introduction

The TransPlex Complete WTA2 kit was initially developed for enhanced amplication of damaged RNA, described at sigma.com/wta2, "Detection of Prostate Cancer Biomarkers by Expression Microarray Analysis of TransPlex® WTA2-Amplified FFPE Tissue RNA".

Moreover, efficient amplification of small reaction input quantities of RNA, approaching the single cell level, has been demonstrated using the TransPlex Complete WTA2 kit. RNA isolated from as few as 10 fluorescence-activated sorted cells (FACS) is amplified, maintaining expression patterns observed for 10- and 100-fold larger cell samples. Challenges unique to exponential amplification of small RNA input quantities are discussed and effectively addressed for downstream qPCR and microarray applications.

Materials and Methods

FACS-isolated bone marrow stem cells were provided by Albert Donnenberg, University of Pittsburgh. Cells were sorted directly into GenElute™ Mammalian RNA Extraction Kit Lysis Buffer (Cat. No. RTN10, component No. L8265), frozen and stored at -70°C until extraction. Total RNA was extracted from 10-, 100-, and 1000-cell samples using the GenElute Mammalian RNA Isolation kit. Sample RNA was eluted at a reduced volume of 40 µl with nuclease-free water, and subjected to a scaled-up (5X) RNase-free DNase digestion (Cat. No. AMP-D1), for a final volume of 97 µl. Each RNA sample was amplified in its entirety using the TransPlex Complete WTA2 kit: the entire heat-attenuated DNase digestion was directly added to the recommended bulk amplification reaction volume of 375 µl, compensating by subtraction of water volume. As per kit instructions, the 375 µl volume was split into five 75 µl reactions for amplification. Each reaction was carried through two cycles beyond stationary amplification ("plateau"). Unincorporated primers and other reaction components were removed from the amplification product using the GenElute PCR Cleanup kit (Cat. No. NA1020). Quantitative PCR was performed using SYBR® Green JumpStart™ Taq ReadyMix™ (Cat. No. S4438), with 250 nM primer concentration. Primer pairs, with the exception of the 18S rRNA set, are specific for cDNA detection.

Amplified RNA (double-stranded cDNA) was labeled with cy3 or cy5 following the Agilent Genomic DNA Labeling Kit Plus procedure (Cat. No. 5188-5309). Two amplified, differentially-labeled cDNA targets were combined, heat-denatured, applied to Agilent Whole Genome™ arrays (G4112F), and incubated at 65°C for 40 hours. Hybridization and subsequent wash procedures for the Agilent Oligonucleotide Array-Based CGH for Genomic DNA Analysis were followed, except for the omission of COT-1 DNA block during hybridization. Array were scanned (Agilent scanner model G2505B) and features extracted. Intensities representing non-uniform features or associated with local non-uniform background, in addition to "absent" and saturated array feature intensities, were removed using Microsoft Excel™. All intensities less than 100 were removed. Intensities were then subjected to median normalization, Welch's t-test, Benjamini-Hochberg false discovery rate correction, and a +1.5-fold differential threshold using Genesifter pair-wise comparison, except where only median normalization was performed, as indicated. Normalized-only intensities and intensities subjected to Genesifter™ (Geospiza) pair-wise analysis were further analyzed using the Genesifter Intersector software and Spotfire™ (TIBCO).

For the study of the effect of reduced library synthesis primer concentration on amplification of decreasing quantities of RNA input, mRNA levels for high- and medium-copy transcripts in the amplification products, plus a "rare" transcript, were evaluated by real-time qPCR. Preparation of the lowered concentrations of library synthesis primers was accomplished by dilution of the kit Library Synthesis Solution (component# L9293) with 5 mM dNTPs (a 2-fold dilution of 10 mM dNTPs, Cat. No. D7295).

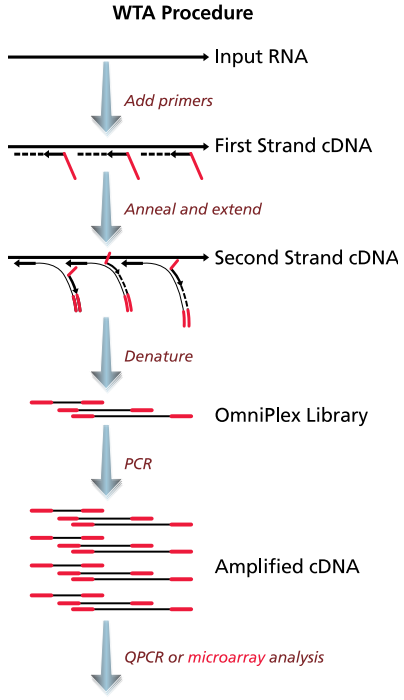


Figure 1. TransPlex WTA Methodology.[1]

Total RNA is combined with Library Synthesis Solution and heat denatured. This is followed with the addition of Reaction Buffer and the strand-displacing Library Synthesis Enzyme, for "single-tube" reverse-transcription and second-strand cDNA synthesis (OmniPlex™ Library). The 3' end of the Library Synthesis Primer is "quasi-random" and substantially non-self-complementary, while the 5' end is a single, constant non-self-complementary sequence that serves as the annealing site for the universal amplification primer.

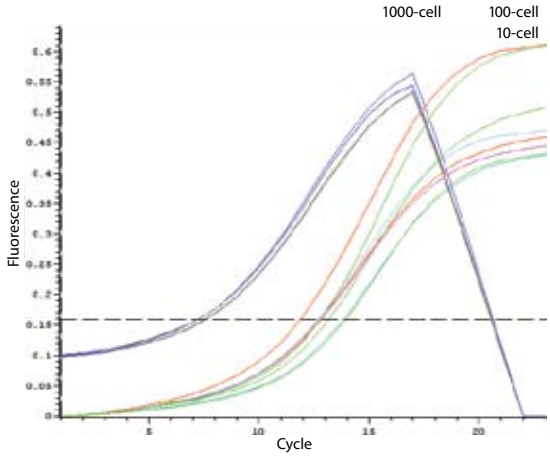


Figure 2. RNA Amplification Profiles.

Total RNA was extracted from cell samples as described in Materials and Methods. RNA amplification was allowed to proceed through 2 cycles of stationary amplification (observed by real time PCR using SYBR Green dye). One thousand-cell reactions were stopped at 19 cycles; 100- and 10-cell samples at 23 cycles.

Figure 3A.

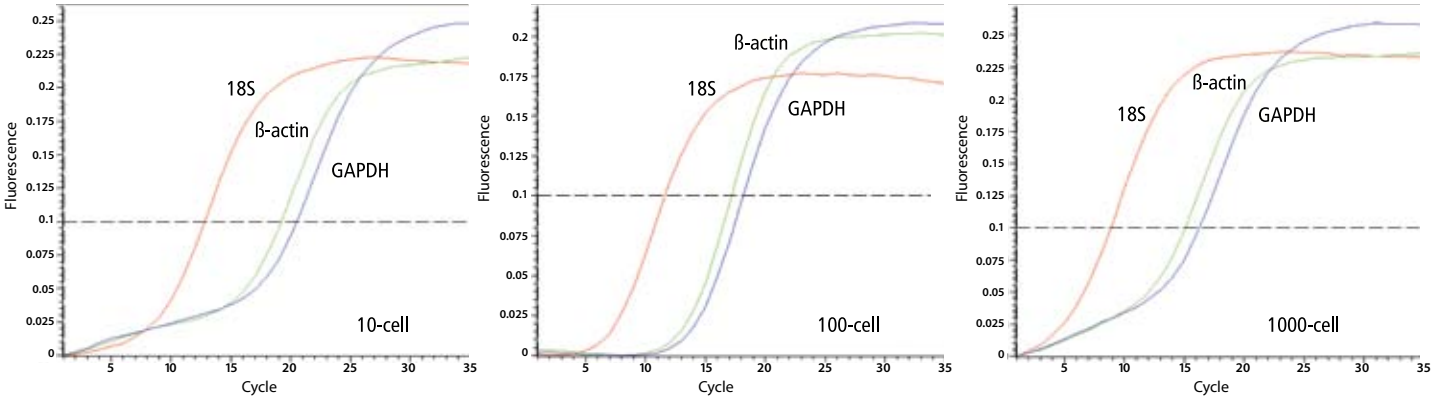


Figure 3B .

	18S C(t), 0.1	β-actin C(t), 0.1	GAPDH C(t), 0.1
10-cell	12.82	19.23	20.5
100-cell	11.63	17.2	18.16
1000-cell	8.89	15.1	16.26

1) 10-cell
2) 100-cell
3) 1000-cell

Figure 3C.

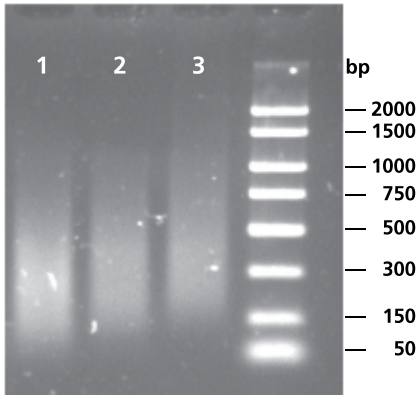


Figure 3. Amplification Qualitative Test.

Prior to microarray analysis, amplification reaction products were evaluated by real-time qPCR using primer pairs for 18S, GAPDH, and β-actin transcripts **(A, B)**. qPCR amplicons were further checked by 1% agarose electrophoresis **(C)**.

Results and Discussion

Amplification of Cellular RNA.

Total RNA was extracted from FACS-isolated human bone marrow stem cells as described in the Materials and Methods. Amplification for all samples was allowed to proceed through 2 cycles of "plateau", or stationary amplification (**Figure 1**). Amplification product was pre-evaluated by qPCR and agarose gel electrophoresis prior to microarray analysis (**Figure 2**). An intense low molecular weight band is observed for the "No RNA" control reaction. This banding pattern is occasionally observed for very low input quantities of RNA, though not in the case presented here. This fragment has not been sequence-characterized, though it is likely to be a multimer of the TransPlex WTA2 library synthesis primer sequence. The 3' region of TransPlex WTA "quasi-random" primers are non-self-complementary, providing a vast improvement in the efficiency of amplification over poly-N priming. However, the primer for the Transplex WTA2 kit has been re-engineered for amplification of fragmented RNA. In so doing, concatemerization of the primer could occur in the absence of template. An alternative explanation will be discussed below.

Microarray Analysis of Whole Cell Samples

A strong concordance between the three stem cell samples is observed for detection of microarray features, in this case, normalized intensities (> 100) (**Figure 3**). Greater than 85% of features detected in the 10-cell sample are also detected for the 100-cell sample, while > 90% of 10-cell features are also detected in the 1000-cell sample. Considering the total number of unique biological probes (41,000) on the array, detected features for the 10-cell and 100-cell comparison account for 59% of total features, with 48% of features shared between the data sets. Similarly, the 10-cell and 1000-cell sample comparison, 61% of total array features are detected, with 52% of features shared. It is noteworthy that exclusion of the data less than 100 denotes a conservative measure, since the Agilent feature extraction software removes calculated background.

Application of statistical tests provides additional confidence to the normalized data (**Figure 4**). An excess of 55,000 features were common for the 10-cell and 100-cell samples at a $p = 0.0001$. (This stringency reflects the analysis of an $n = 4$ for each data point). Moreover,

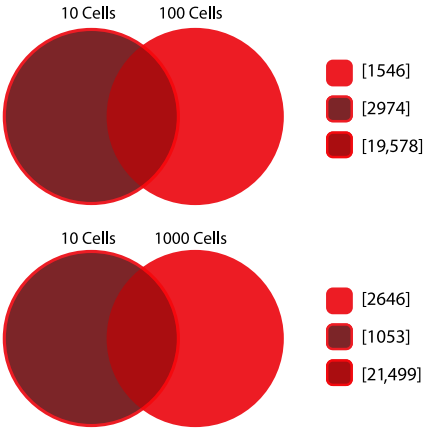


Figure 4.

Microarray analysis revealed a strong concordance between the cell samples. Overlapping regions represent median-normalized intensities (>100). Greater than 85% of detected features represented in the 10-cell sample are detected for the 100-cell sample, while >90% of 10-cell features are also detected in the 1000-cell sample. In addition, the overlapping regions are highly concordant, with greater than 98% of 10-cell/100-cell overlapping intensities detected in the 10-cell/1000-cell intersection.

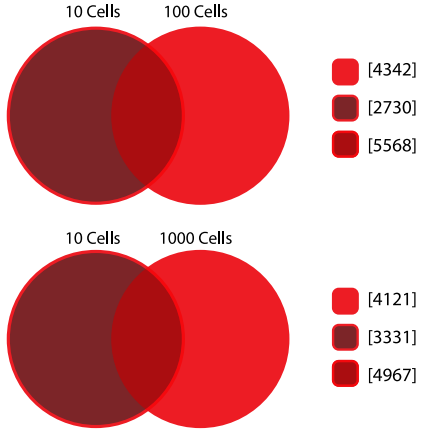


Figure 5.

A substantial number of statistically significant microarray features (>5500) are detected, with concordance between the two overlapping regions of the cell sample data sets maintained following stringent statistical analysis and corrections of the data presented in **Figure 3** (see Materials and Methods). Seventy percent of 10-cell/100-cell overlapping intensities are detected in the 10-cell/1000-cell intersection.

70% of features common to the two data sets (the 10-cell/100-cell comparison) were also detected in the 10-cell/1000-cell intersection. Similar percent values of commonality are observed in other studies, such as comparing dilutions of intact RNA, matched FFPE and frozen tissue sample RNA, or amplified versus unamplified RNA (not shown). Testing is underway to determine whether amplified transcripts represented in a "unique" data set are present in the unamplified sample. We hypothesize that the TransPlex WTA2 amplification method is sufficiently sensitive to detect basal expression. Such highly stochastic signal should, however, be able to eliminate with sufficient replication and stringent statistical testing.

Amplification of Low RNA Input: Effect of Banding Pattern on qPCR Analyses.

The low molecular weight banding pattern, observed for the No RNA Control in **Figure 2**, has been observed in some 10-cell amplifications, and occasionally in the 100-cell amplification (not shown). It was suggested above that the phenomenon might be reasonably explained by concatemerization of the library synthesis (LS) primer. An alternative hypothesis is that the relatively high concentration of the LS primer results in over-priming, and ultimately, a diminutive amplification product. This is possible since library synthesis (first and second strand cDNA synthesis) occurs in a single tube, driven by an enzyme with strand-displacement capability. To test this idea, decreasing human prostate RNA input quantities were matched with similarly decreasing LS primer concentration during the library synthesis step. Two-fold serial dilutions of each component allowed for evaluation of twelve RNA input values ranging from 25 ng to 12.2 pg, and LS concentration starting at 1X kit concentration, to 0.0078X, on a 96-well plate. A marked loss in linearity of amplification efficiency is observed when using the 1X LS primer concentration, at RNA input quantities less than ~800 pg (**Figure 6A**). Linearity returns for lower RNA input quantities as the concentration of the 1X LS primer concentration is reduced by 2-fold increments (data not shown), corresponding to the reduction of the low molecular weight banding pattern (**Figure 5**). Quantitative PCR results using primer sets for high- (18S rRNA), medium- (beta-actin, GAPDH), and low-copy TRCC2 transcripts corroborate these observations (**Figures 7A-D**). Interestingly, detection of TRCC2, presumably a low copy transcript in prostate,

Figure 6A.

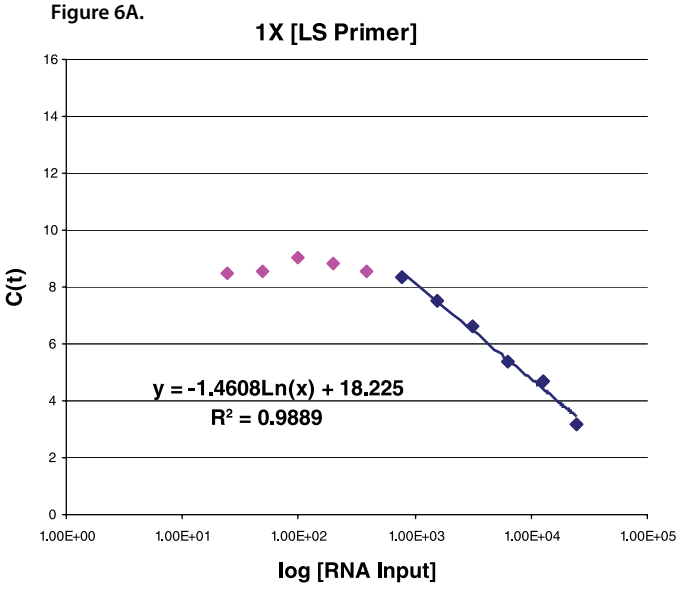


Figure 6B.

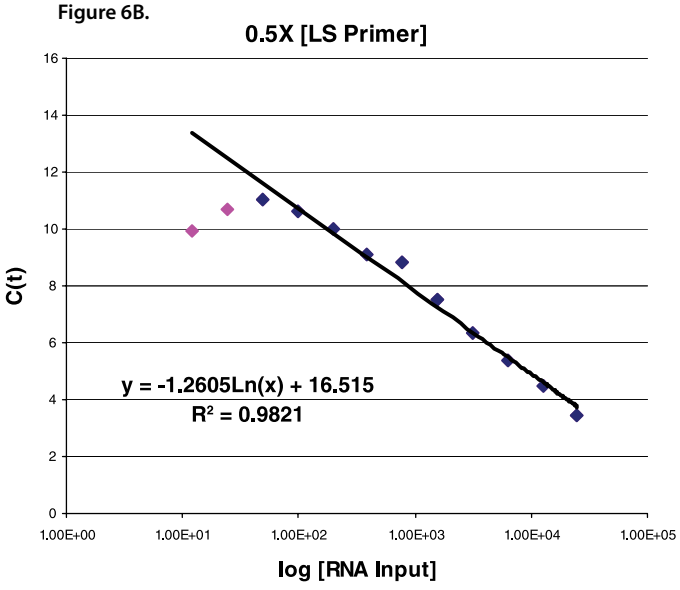


Figure 6C.

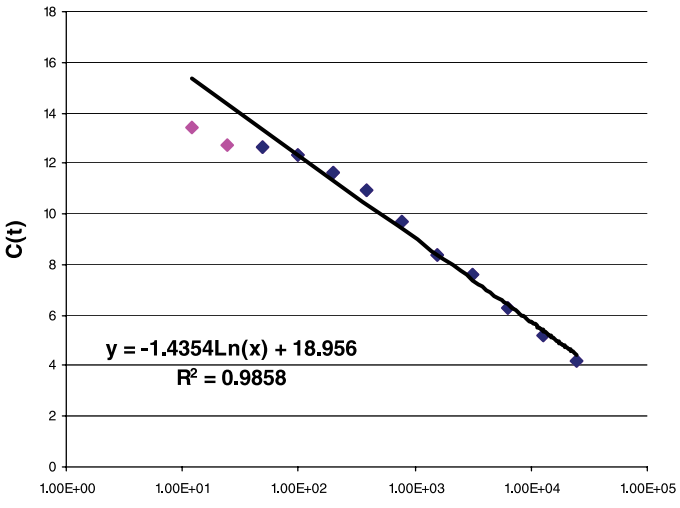


Figure 6D.

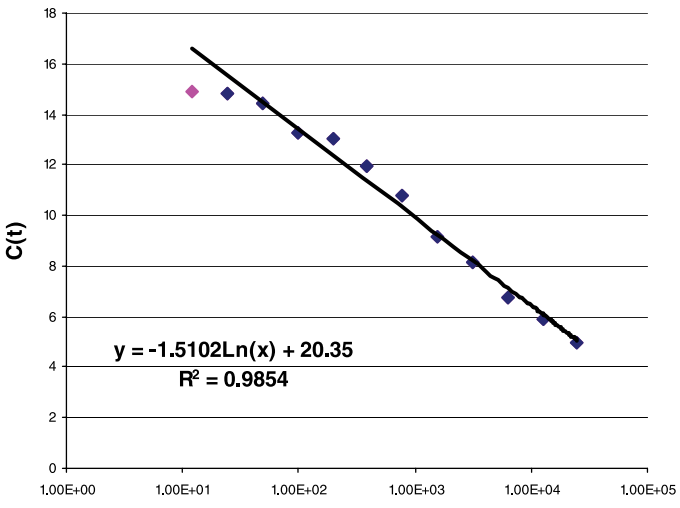


Figure 6.

The effect of library synthesis primer concentration on amplification of small quantities of input RNA was examined. Two-fold serial dilutions of human prostate total RNA were added to reactions, with input quantities ranging from 25 ng to 12.2 pg. In addition, decreasing concentrations of library synthesis primer were evaluated: the 1X Transplex WTA2 Kit concentration and three additional 2-fold serial dilutions. C(t) values measured at early exponential growth (threshold, 0.025) for each amplification reaction are plotted against the input quantity of RNA for each library synthesis primer concentration tested. The results for the kit concentration primer (1X) and a 2-, 4-, and 8-fold dilution of primer are shown. These results suggest that RNA input quantities less than ~800 pg are not efficiently amplified in the presence of 1X library synthesis primer concentration. Intermediate primer concentrations (0.5X and 0.25X) show increasing improvement of amplification efficiency. Overall efficiency appears to be maintained when the 8-fold dilution of library synthesis primer is used.

Figure 7A.

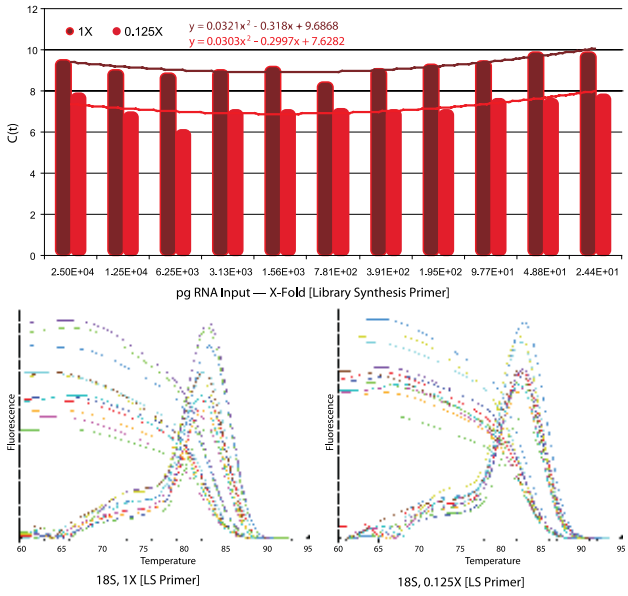


Figure 7B.

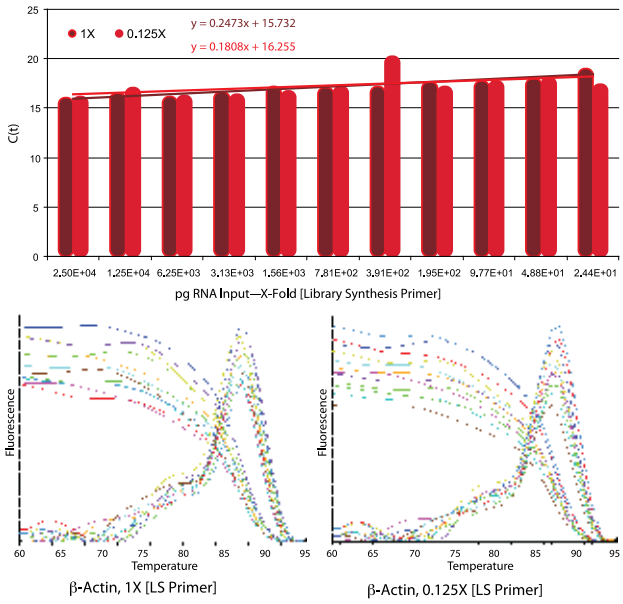


Figure 7C.

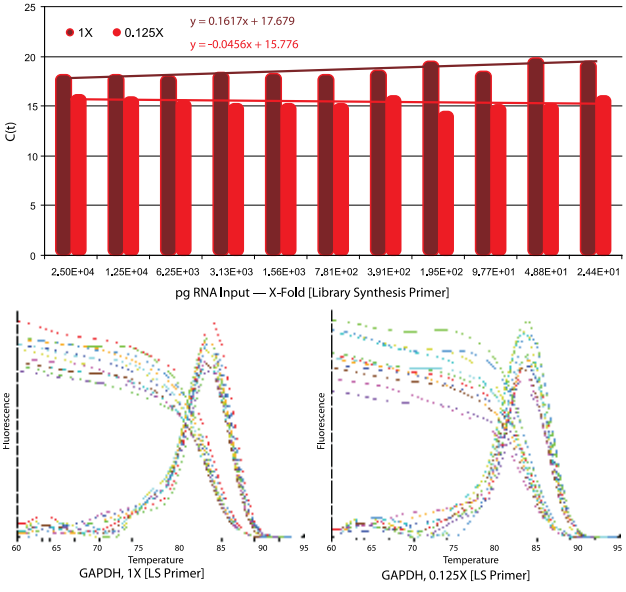


Figure 7D.

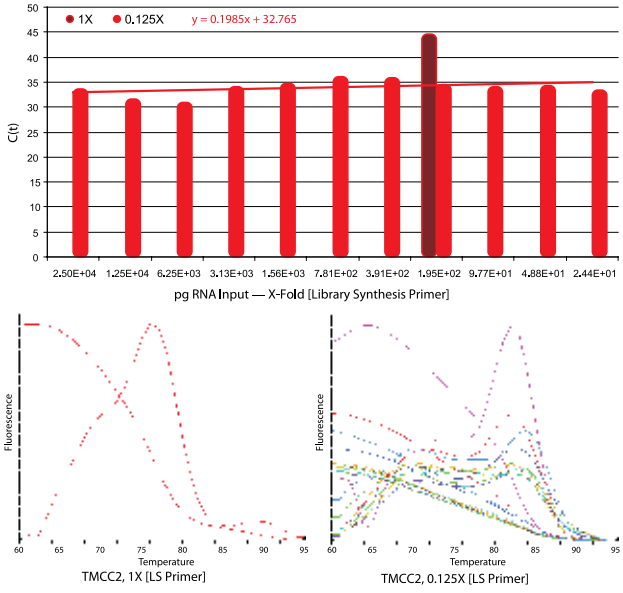


Figure 7.

Amplification products were subjected to qPCR evaluation using primer sets for 18S rRNA **(A)**, β-actin **(B)**, GAPDH **(C)**, and TMCC2 **(D)**. Inserts show the melt curves for amplification product at the decreasing RNA reaction input quantities, and for the 1X and 0.125X primer concentrations. **A–C** represent high-copy rRNA and medium-copy and house-keeping transcripts; TMCC2, human transmembrane and coiled-coil domain family 2 transcript, was noted to be inconsistently detectable in previous studies, with C(t)s in the mid-30s. Interestingly, a pattern of biphasic inefficient amplification is observed for 18S rRNA, at highest and lowest input quantities for both library synthesis primer concentrations. Generally consistent C(t) values are observed for β-actin and GAPDH at the 0.125X primer concentration, while a gradual decrease in efficiency is observed as input quantities decrease at the 1X primer concentration. At the 1X primer concentration, no amplification product is detected for TMCC2. Only a spurious signal at 75 degrees (primer-dimer, gel data not shown) is observed. However, at the low library synthesis primer concentration, TMCC2 amplicon is detectable at all input quantities.

Figure 8A.

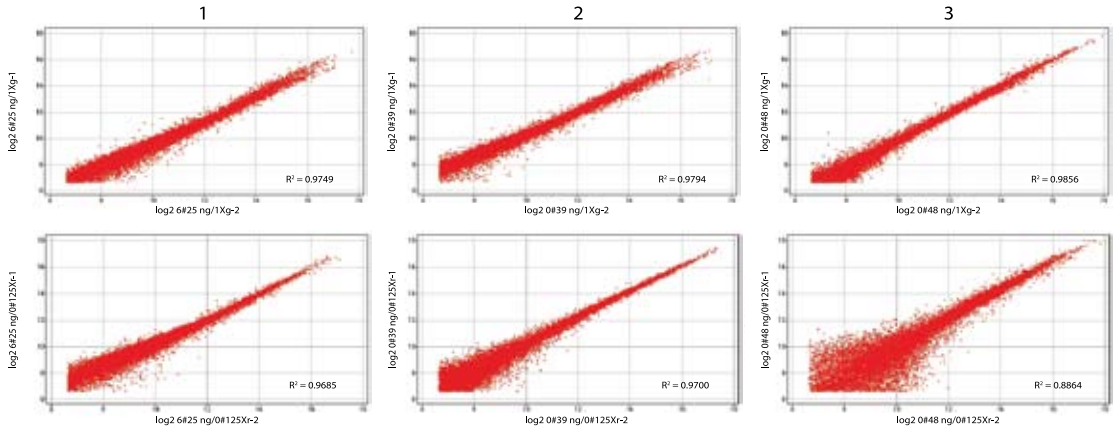


Figure 8B.

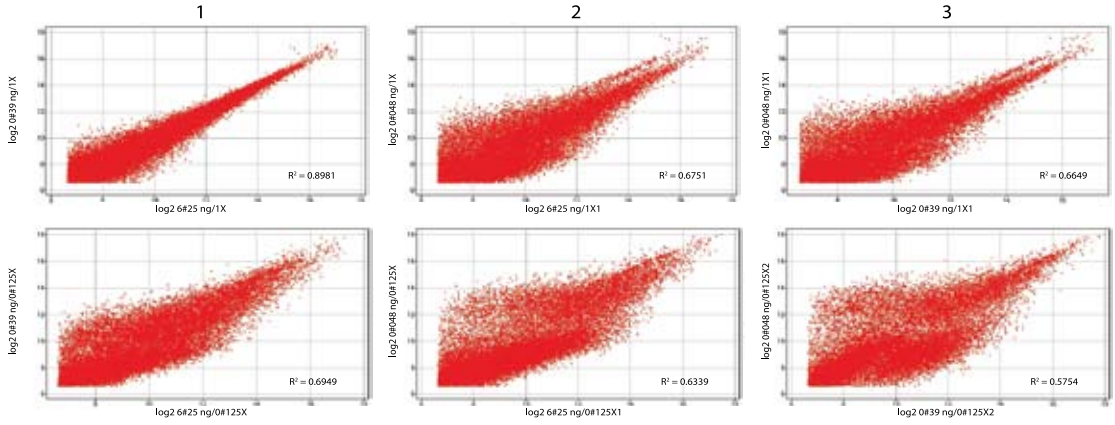


Figure 8C.

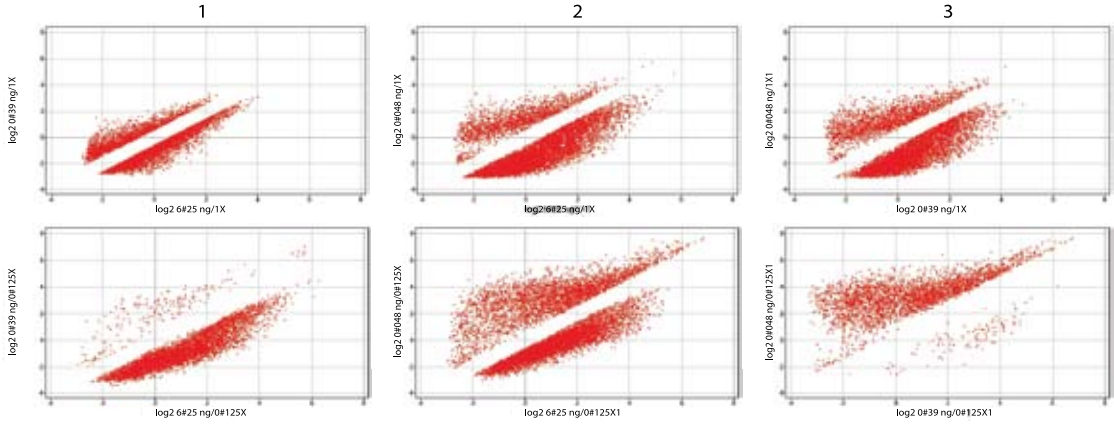


Figure 8.

Amplification product generated from 6.25 ng, 390 pg, or 48 pg input quantities, prepared with 1X or 0.125X library synthesis primer, was labeled for microarray analysis as described in Materials and Methods. Results are shown in Panels A through C (log intensities vs. log2 intensities). **Panel A** shows reproducibility for duplicate 6.25 ng reactions (1), 390 pg reactions (2), and 50 pg reactions (3). In **Panel B**, 6.25 ng reaction intensities are plotted versus 390 pg reaction intensities (1), 6.25-ng reaction intensities versus 48-pg reaction intensities (2), and 390 pg reaction intensities are plotted versus 48 pg reaction intensities (3). Statistically analyzed and normalized intensities are shown in **Panel C**: 6.25 ng reaction intensities versus 390 pg reaction intensities (1), 6.25 ng reaction intensities versus 48-pg reaction intensities (2), and 390 pg reaction intensities versus 48 pg reaction intensities (3). The positive effect of lowered library synthesis primer concentration observed for quantitative PCR is not evident in microarray results.

is problematic at the 1X concentration, with only a background signal observed for an intermediated RNA input quantity. TRCC2 amplicon is detected for all RNA input quantities at an 8-fold LS primer dilution, indicating an 8-fold dilution of the WTA2 LS primer optimal for TRCC2 for downstream qPCR.

Amplification of Low RNA Input: Effect of Banding Pattern on Microarray Analyses

To determine whether the use of a reduced LS primer concentration is appropriate for microarray anaysis of low RNA input quantities, amplification product from three input quantities was tested (**Figure 8**): 6.25 ng (similar to the kit-equivalent of 5 ng per 75 ul reaction) (1), 391 pg (just less than the 800 pg demarcation observed for reduced amplification efficiency for 1X LS primer concentration in **Figure 6A**) (2), and 48.8 pg (approximately the RNA mass equivalent of a 'single' cell) (3). Amplification products were labeled and subjected to microarray analyses as described in Materials and Methods. In **Figure 8A** (non-normalized intensities) reproducibility between same-dye duplicate data for each original RNA input quantity is shown. **Figure 8B** (also non-normalized intensities) shows a similar comparison of duplicates for each input quantity (n = 4). Each comparison shows an increase in variance between the two sets of data (see Results and Discussion section on pg. 32). However, one would postulate that there is a considerably less loss of quantitative and qualitative correspondence between the duplicates generated in 1X LS primer than those prepared with the 8-fold dilution. This observation is extended to **Figure 8C**, where the data sets have been statistically tested (t-test, false discovery rate correction, and +1.5 threshold). Unquestionably, the 1X library synthesis primer concentration performs better for microarray analysis.

Conclusions

These observations indicate that reduced library synthesis primer concentration enhances detection of low-copy transcripts when performing downsteam qPCR application. At 1X LS primer concentration, medium-copy transcripts can be reliably detected. However, at this concentration, low-copy transcripts may be overprimed, resulting in template too short for both qPCR primers of a set to find their respective annealing sites. This would not appear be a problem for microarray amplifications, where such size reduction of the amplification product would be essentially equivalent to target fragmentation. However, reduced primer concentration decreases the probability of the LS primer to annealing to all potential templates during library synthesis, thus preventing uniform amplification of the transcriptome. This is not to suggest that the observed effect of higher LS primer concentration on amplification efficiency (**Figure 6**) is not affecting microarray results. This effect manifests itself in predictable and manageable small increases in expression differentials for low-copy transcripts, and a similar minor compression of those of high-copy transcripts. Therefore, the low library synthesis primer concentration is optimal for detection of low-copy transcripts by qPCR, and the standard kit concentration for representative amplification for all RNA input quantities when following with microarray analysis.

For the investigator utilizing TransPlex WTA2 amplification product for qPCR, it should be noted that the 8-fold dilution of the LS primer was shown here to be optimal for TRCC2. This could vary with individual amplicons. It is suggested that an investigator test serial dilutions of the LS primer to find the concentration optimal for a given amplicon, as described in this presentation. (Directions for LS primer dilution are provided in Materials and Methods.)

In conclusion, the TransPlex Complete WTA2 kit provides a robust, sensitive method for amplifying RNA quantities approaching the single cell level, while maintaining the representative pattern of expression of higher quantities of RNA input.

Acknowledgements:
Our appreciation to Albert Donnenberg, University of Pittsburgh for providing bone marrow stem cell for this study. We would like to acknowledge the members of Sigma-Aldrich Biotechnology R&D, in particular Patrick Sullivan, Ernie Mueller and current supervisor, Carol Kreader, for their assistance in the development of this product; and Sigma Marketing Communication Group for their assistance in the preparation of this poster.

References:

(1) Amplification and analysis of whole genome and whole transcriptome libraries generated by a DNA polymerization process. Kamberov, E., et al., US Patent Application 2007054311, 3/8/2007. (Previously, US Patent Application 20040209298, 12/14/2005.) EPO Application EP1604040, 12/14/2005. WIPO PCT WO2004081225, 9/23/2004).

(2) Modeling subgenomic hepatitis c virus RNA kinetics during treatment with alpha interferon. Uprichard, S.L., et al. *Journal of Virology Methods*, **83**, 6383-6390 (2009).

(3) Outbreak of hepatitis E virus infection in Darfur, Sudan: effectiveness of real-time reverse transcription-PCR analysis of dried blood spots. Mérens, A., et al., *Journal of Clinical Microbiology*, **47**, 1931-33 (2009).

(4) Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA. Iscove, N.N., et al., *Nature Biotechnology*, **20**, 940-43 (2002).

(5) Direct comparison of microarray gene expression profiles between non-amplification and modified cDNA amplification procedure applicable for needle biopsy tissues. Li, Y., et al., *Cancer Detection and Prevention Journal*, **27**, 405-11 (2003).

(6) Functional expression of the extracellular-Ca2+-sensing receptor in mouse taste cells. Bystrova, M.F., et al., *Journal of Cell Science*, **123**, 972-82 (2010).

(7) A comparison of RNA amplification techniques at subnanogram input concentration. Lang, J.E., et al., *BMC Genomics*, **10**, e326 (2009).

(8) An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. Kurimoto, K., et al., *Nucleic Acids Research*, **34** (5) e42 (2006).

(9) Options available for profiling small samples: a review of sample amplification technology when combined with microarray profiling. Nygaard, V., et al., *Nucleic Acids Research*, **34**, 996-1014 (2006).

(10) *A Biologist's Guide to Analysis of Microarray Data*. Steen, K., Wiley-Interscience: New York, 18-20 (2002).

ZFN DONOR DESIGN

Codon and Single Base Genome Editing Using Zinc Finger Nucleases

A guide for the design of targeting donors for creating point mutations, codon changes, SNP corrections, and other small site-directed genomic modifications

Gregory D. Davis , Ph.D., and Shondra M. Pruett-Miller, Ph.D.
Sigma Life Science

ZFN Design and Positioning Relative to the Desired Mutation Site

Homologous recombination (HR) is a process by which double-stranded breaks (DSBs) are repaired in cells. In the context of normal cellular function, HR is a "copy and paste" mechanism by which a DSB is repaired using the information contained in a homologous sister chromatid. However, a user-defined mutation can be introduced by transfecting a donor plasmid that contains both the user-specified mutation and sequence homologous to the regions flanking the desired site of modification. In mammalian cells, the spontaneous rate of HR between a user-defined donor plasmid and a homologous locus is very low (10^{-6} to 10^{-5}) and is often outside the range for practical experimentation, even when using antibiotic selection. However, several groups have shown that the creation of a DSB at or near the site of desired mutation can increase the rate of HR by several orders of magnitude (Rouet, et al., 1994). Until recently, the creation of a DSB at a desired locus was not feasible. Zinc finger nucleases (ZFNs) are engineered endonucleases that can now be rapidly designed to create DSBs in the vicinity of a desired mutation site with routine success (Urnov, et al., 2005). Following successful design and manufacturing of a site-specific ZFN, the next key step to create targeted mutations via HR is to design and construct a targeting donor plasmid.

An overall workflow for donor plasmid design is shown in **Figure 1**. If the goal is to create a point mutation, the ZFN should be designed to cut within 200 bp of the desired mutation site. Existing literature (and unpublished data) shows that the penalty for moving the point mutation 100 bp away from the cut site is an approximate 4X drop in mutation frequency (Elliot, et al., 1998). These authors documented a point mutation 511 bp from a nuclease cut site, however, it occurred at a frequency that would likely require screening of >1,000 clones to isolate a mutant in the absence of selection. So, whenever possible, it is advisable to use a ZFN that is located <200 bp away from the cut site to maintain high efficiency and thus minimize clone screening. Using selection methods, it is certainly reasonable to expect

Figure 1.

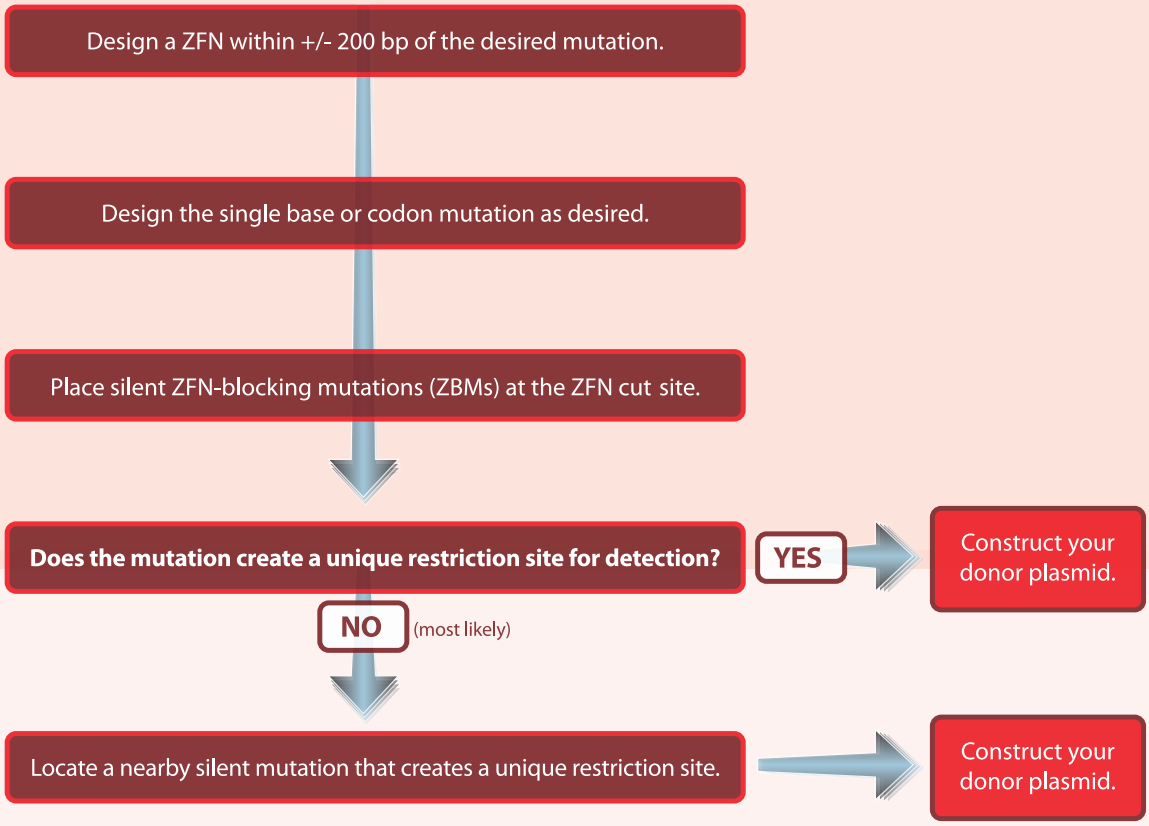


Figure 1. Workflow for designing plasmids to create small mutations near ZFN cut sites.

that mutations can be incorporated >500 bp from the cut site, but these applications require position-restricted integration of large selectable gene cassettes that complicate ZFN design and risk affecting gene expression and regulation.

Length of Homology Arms and Plasmid Backbone

When designing donor plasmids for making small mutations, it is recommended to use ~400-800 bp of homology in each arm centered close to the ZFN cut site (**Figure 2**). Donor plasmids with 800 bp of homology in each arm have been shown to work robustly in different applications requiring point mutations. Donor plasmids with shorter arm lengths (50-100 bp) have been shown to work (Orlando, et al., 2010), but are generally less effective than 400-800 bp across a broad range of cell types. It is reasonable to assume that selection methods might enable use of 50 bp homology arms in a broader range of cell types, but this has yet to be explored.

For plasmid backbones, it is recommended to use small pUC-based plasmids (<2,600 bp) to limit the mass of DNA that is transfected to cells. To date, multiple variations of pUC based backbones have been used to successfully execute ZFN-based gene targeting experiments. Total gene synthesis companies (that can be used for donor synthesis) generally prefer plasmid backbones which minimize the cost of their cloning and sequencing operations. It is perfectly fine to use whatever standard cloning vector vendors offer as long as the plasmid is small, lacks common restriction sites, and lacks significant sequence homology to the target region of interest.

Detecting desired mutations in pooled and clonal cell populations

Prior to investing the effort to derive and screen a clonal cell population, it is useful to estimate the mutation frequency at the pooled cell level immediately following

Figure 2.

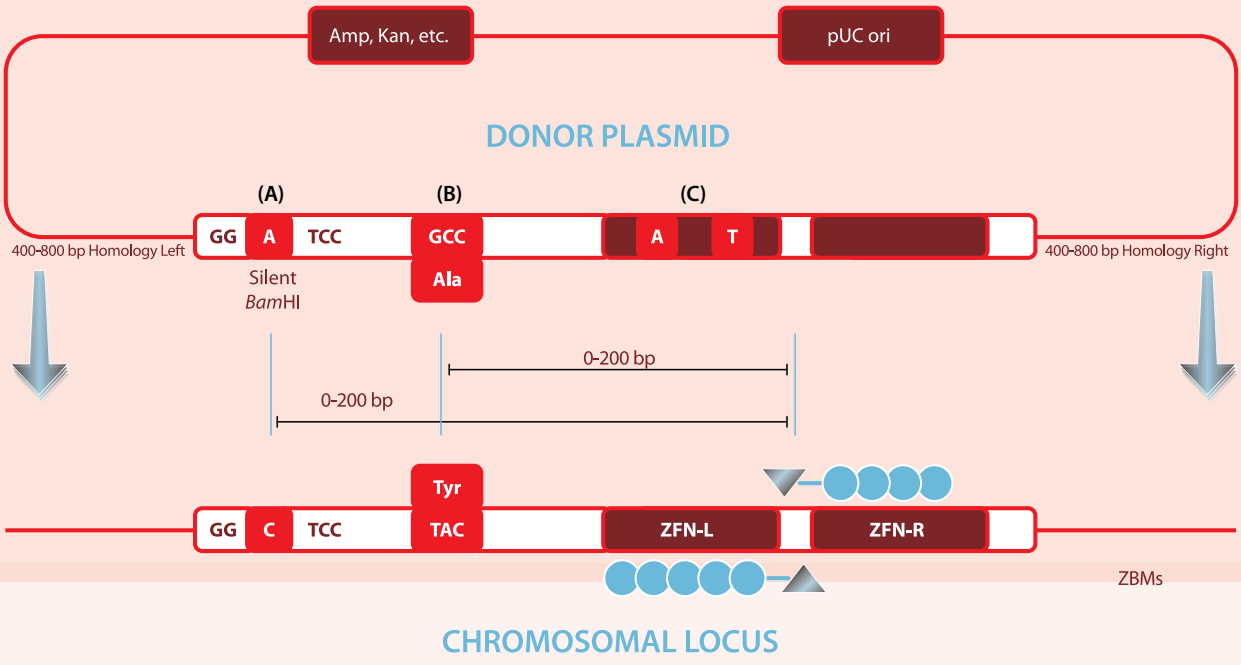


Figure 2. Example donor plasmid for creating a codon change near a ZFN cut site.

Three key mutations need to be considered: (**A**) a silent mutation that allows detection and quantification via PCR and restriction enzyme cleavage (RFLP), (B) the desired mutation for the biological experiment, and (**C**) mutations (ZBMs) within the ZFN binding site that disrupt intracellular ZFN cleavage of the plasmid or cleavage of the chromosome post-integration of the donor plasmid. Note: the desired mutation (**B**) should be flanked by the detection mutation (A) and the ZFN cut site. In rare cases, the desired mutation will create a new unique restriction site.

transfection of the ZFN and donor plasmid. Addition of a silent restriction site somewhere in the donor will facilitate easy detection of mutant clones (e.g., an RFLP site). The RFLP site should be as close as possible to the desired mutation and always positioned so that the desired mutation is flanked by the RFLP site and the ZFN cut site (see **Figure 2A**). Homologous recombination inserts new sequence directionally in relation to the ZFN cut site, and this method of positioning the RFLP site will help ensure that when analyzing clones the desired mutation is present in all RFLP positive clones. In rare cases, the desired mutation may actually create a unique restriction site. An excellent resource for finding silent mutations quickly is WatCut: (http://watcut.uwaterloo.ca/watcut/watcut/template.php).

When creating point mutations to create RFLP sites and amino acid changes, it is recommended to check the codon usage in the gene of interest and to use the most frequently used codons in that gene where feasible. In rare cases, even silent mutations can have drastic effects on functional gene expression (Kimchi-Sarfaty, et al., 2007). A good site for determining codon usage frequency is: http://www.ebioinfogen.com/biotools/codon-usage.htm. After finding a good RFLP site, check that it is not located at another site within the donor, since this will complicate downstream analysis.

Preventing unwanted NHEJ mutations when creating precise point mutations

If the ZFN cut site is present within an open reading frame (ORF) and the goal is to make a nearby codon change or SNP correction, it is possible that the ZFN will also cause undesirable secondary mutations at the cut site via misrepair arising from non-homologous end joining (NHEJ). In our work with donor plasmids, some configurations resulted in 30-40% of the correctly modified clones also having secondary mutations at the ZFN cut site, while other applications had >90% secondary mutations. The level of secondary mutations appears to be higher when the desired mutation is located further away from the

Figure 3.

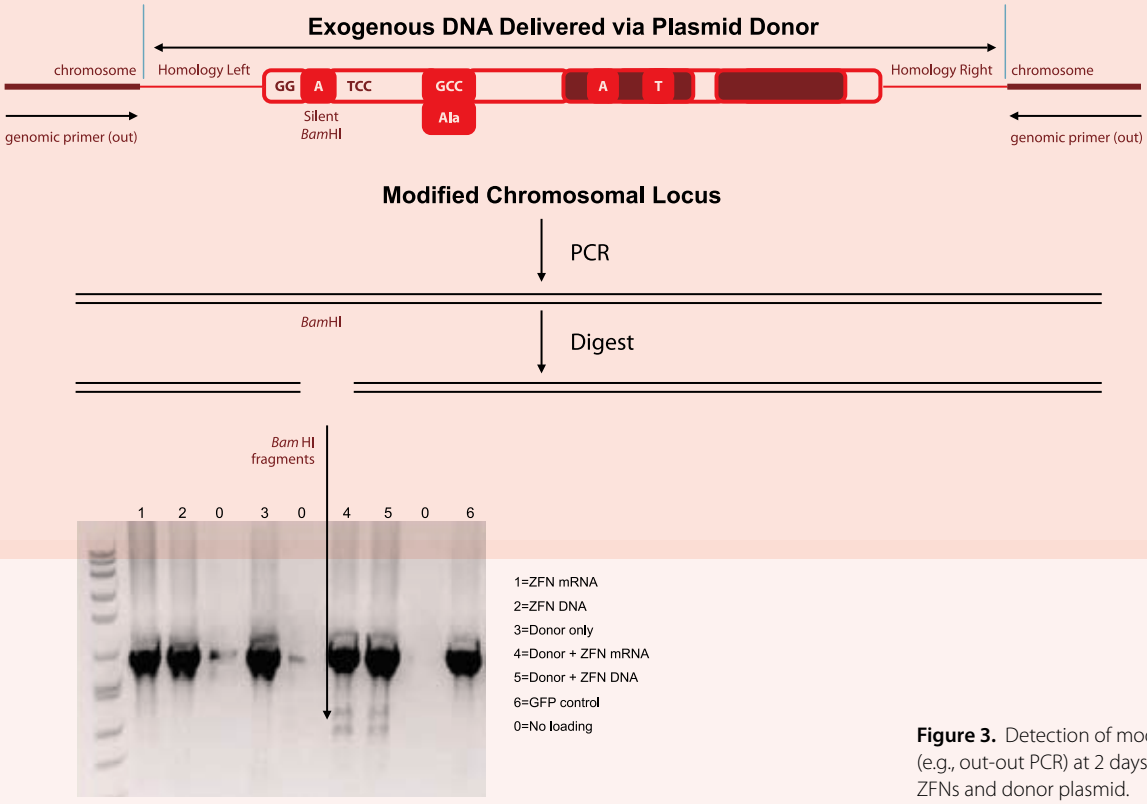


Figure 3. Detection of modified alleles by RFLP (e.g., out-out PCR) at 2 days post-transfection of ZFNs and donor plasmid.

cut site. As a general practice, we recommend incorporating silent mutations in the ZFN binding site to prevent ZFN cleavage of the donor plasmid and/or recleavage of the target site post-integration of the plasmid. At least two ZFN-blocking mutations (ZBMs) should be incorporated into the ZFN binding site - separated by at least 3 bp if they are within the same ZFN arm, or one placed in each arm (**Figure 2C**). The 3 bp spacing ensures each ZBM lies in the binding sequence of a different zinc finger, thereby maximizing the chances for disruption of binding. These mutations should be selected in a way that minimizes changes in codon usage frequency. This can be done by placing the ZBMs at amino acids with highly degenerate codons, such as Ala, Val, Thr, Pro, Ser, Leu, Gly, and Arg. This will give more options for preserving codon usage frequency since they all have approximately four different codons to select from.

Detecting mutations at the pooled and clonal cell level

Following transfection of the ZFN and donor construct, the next step is to assess the rate of mutation in the pooled cell population prior to diluting (or FACS sorting) the cells for single cell cloning. Two different methods maybe used to detect mutations: (1) RFLP, using a silent restriction site (**Figure 3**), and/or (**2**) mutation-specific junction PCR (**Figures 4A and 4B**). Because both assays are based on PCR, it is essential to always include a "donor only" transfected sample to control for PCR artifacts, especially at the pooled level when copy number of the donor plasmid will be high. For analyzing a pool of cells, the RFLP approach is the most quantitative method and best predicts success in follow-up single cell cloning efforts. For RFLP detection, two PCR primers should be designed that prime outside the region of homology present on the donor plasmid (out-out PCR; **Figure 3**). Genomic DNA is isolated 2-3 days post-transfection and used as a PCR template. The PCR product is then digested with the RFLP restriction enzyme (in this case *Bam*HI) and resolved by electrophoresis on an agarose or polyacrylamide gel. The rate of targeted integration can be estimated by performing densitometry on the gel and comparing the intensity of the fragments released by restriction digestion against the intensity of the parental band. If the mass of the parental PCR band is in excess of 700 ng, then the signal from the band is likely saturated and out of the linear range required for accurate quantitation. In our experience, RFLP signals that are visually observable by EtBr staining, regardless of parental band intensity, have high chances of success. **Figure 3** shows a good example of overloading of the parental PCR product,

Figure 4A.

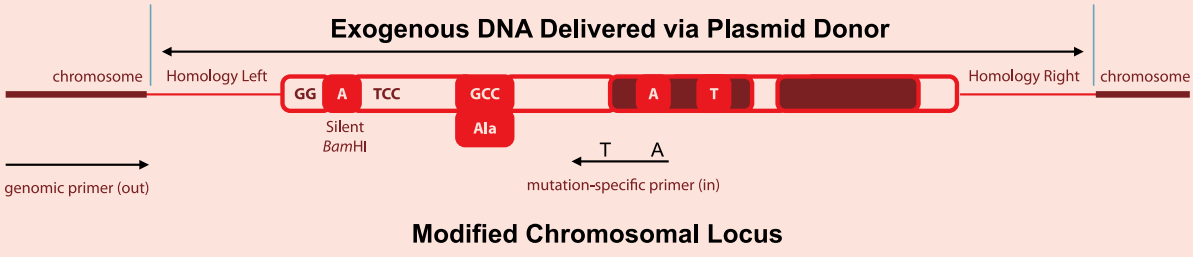


Figure 4B.

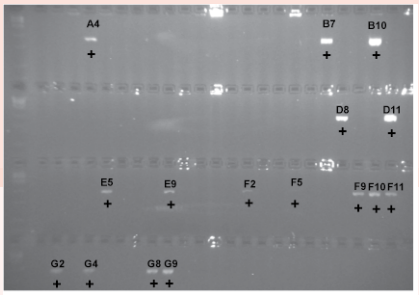


Figure 4A-B. Detection of modified alleles by junction PCR using a mutation specific primer (e.g., out-in PCR).

Figure 5.

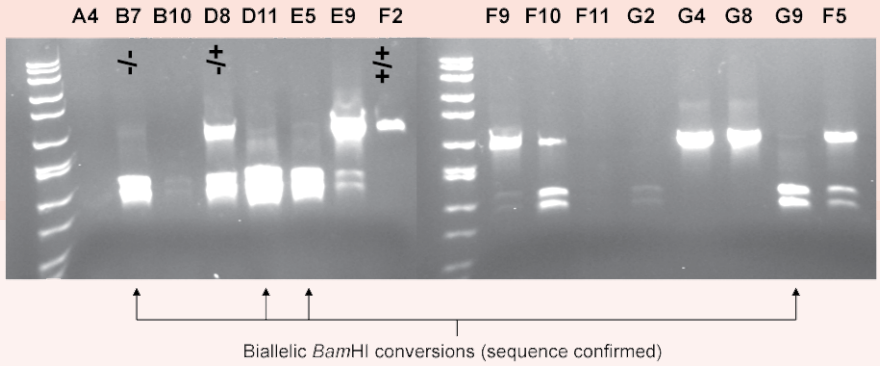


Figure 5. Detection of modified alleles by RFLP (e.g., out-out PCR) of individual clones at 3-weeks post-transfection of ZFNs and donor plasmid.

with observable RFLP fragments. Despite overloading, this experiment yielded correctly targeted biallelic K562 clones at a rate of 6 of 112 clones screened.

When screening clones, either the RFLP or junction PCR approach may be used. **Figures 3-5** outline an experiment in which RFLP was performed on the transfected cells (**Figure 3**), then "out-in" junction PCR was used to screen clones to identify those with the ZBM mutation introduced at the cleavage site (**Figure 4A** and **4B**). A subset of these ZBM-positive clones will also contain the desired mutation. Finally, RFLP was used (**Figure 5**) to assess whether selected clones were monoallelic or biallelic for the *Bam*HI conversion. When screening clones, the RFLP assay becomes more stochastic in that a small defined number of modified alleles will contribute to the RFLP signal (2 to 6 alleles, depending on ploidy). The typical experimental goal is to find a clone that completely lacks wildtype sequence, resulting in complete digestion of the parental PCR product (**Figure 5**, e.g., clone B7), though heterozygous mutations maybe biologically relevant depending on circumstances. Following confirmation of biallelic (or complete) conversion by RFLP, the fidelity of the mutation should be confirmed via DNA sequencing of all alleles. DNA sequencing of clones showing biallelic *Bam*HI conversions revealed that all alleles containing a *Bam*HI conversion also contained the desired codon mutation. This result provides further justification for always flanking the desired mutation by a silent RFLP and the ZFN cut site as illustrated in **Figure 2**.

References

Gene conversion tracts from double-strand break repair in mammalian cells. Elliott, B., et al., *Molecular & Cellular Biology*, **18** (1), 93-101 (1998).

A "silent" polymorphism in the MDR1 gene changes substrate specificity. Kimchi-Sarfaty, C., et al., *Science*, **318** (5811), 525-8 (2007).

Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors within limited chromosomal homology. Orlando, S.J., et al., *Nucleic Acids Research* (2010).

Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Rouet, P., et al., *Molecular & Cellular Biology*, **14** (12), 8096-106 (1994).

Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Urnov, F.D., et al., *Nature*, **435** (7042), 646-51 (2005).

Introducing a GLOBAL EDUCATION PROGRAM

from Sigma® Life Science

Designed to Help you Grow

At Sigma's Biouniversity we are passionate about learning. Our goal is to equip researchers with the essential tools for unraveling complex biomolecular interactions. This is accomplished using a variety of educational resources, including technical courses, symposia, seminars, webinars, and e-learning materials.

Technical Courses

Our team of skilled facilitators and scientists make it their mission to deliver technical courses that provide both "hands-on" experiences with state-of-the-art technologies and a forum for scientific discussion. Our courses have trained scientists from around the world on the latest technologies, and the global reach of our workshops continues to expand.

The EMBL Advanced Training Centre, Heidelberg, Germany

Sigma Life Science recently entered into a corporate partnership with the European Molecular Biology Laboratory (EMBL) Advanced Training Centre, based in Heidelberg, Germany. This partnership provides scientists with the unique opportunity of training in key Sigma Life Science technology areas in EMBL's state-of-the-art teaching laboratories. The collaboration features custom-designed technical workshops and training courses focused on pertinent technologies. The inaugural workshop, Targeted Genome Editing using Zinc Finger Nucleases (ZFNs), was attended by 16 scientists from around Europe and, by all accounts, was a complete success. Through participation in this course, attendees became proficient in ZFN fundamentals, various techniques associated with the workflow, and applications for the technology.

For more information or to register for future workshops, go to sigma.com/workshops

Seminars and Symposia

Sigma offers a wide range of seminars and webinars to enhance your understanding of the latest technologies. We provide focused, technical presentations on the latest applications including targeted genome editing, transgenics, RNAi, regenerative medicine, and biomarker discovery.

The seminar program has been expanded to include the Understanding Biology Symposia Series. These events highlight a research area and feature a series of technical presentations delivered by thought-leaders and Sigma scientists. This symposia series covers topics such as reprogramming strategies, screening and target validation, and targeted genome editing, and are delivered globally in Asia, Europe, and North America.

Upcoming Symposia

Gene Regulation Symposium and Zinc Finger Nuclease Users Group Meeting

This event is being held on November 2, 2010 at the Boston Cambridge Marriott. The morning symposium will highlight technologies including gene silencing and epigenetics. The afternoon session will be the Zinc Finger Nuclease Users Group Meeting, and will feature presentations on the uses of zinc fingers in cell lines and animal models, and will also include roundtable discussions. For more information or to register go to **wherebiobegins.com/biosymposia**.

The goal of Sigma Life Science is to provide researchers with leading technologies and the educational resources requisite for enhancing the understanding of all biological disciplines. We invite you to come learn with us by taking part in one of our technical workshops or symposia. ■

Best of the WEB!

Social Media and its place in science communications is a hotly debated topic. People on the social media side tout its usage as a way to simplify complex collaborative communications. Detractors say that social media is a bit of a time waster, and the value isn't quite there for them.

Here at Sigma® Life Science, we love web-enabled communications. Tools like Blogs, Facebook, Twitter, LinkedIn, and YouTube have allowed us to talk with our customers and other biologists in a way we never have before.

Skeptical?
Please weigh in with our poll, "What's your opinion on scientists using social media?" at wherebiobegins.com/biounity.

Biowire extended online!

In this issue of *Biowire*, we highlighted Dana Carroll of the University of Utah, one of the originators of the Zinc Finger Nuclease technology. His lab works in C. Elegans and Drosophila, and is constantly refining their use of ZFN technology.

Find more information on Dr. Carroll's work on our blog at sigmabioblogs.com, and on wherebiobegins.com/biounity.

Watch a video as Dr. Carroll describes his views on where the ZFN technology is today, and where it is headed in the future. Also learn from members of his lab as they talk through the ZFN projects that make up their day.

Interesting Blogs!

We'd like to let you know who we are reading. Weblogs can be a great way to see a new point of view, or to argue your point of view!

BayBlab – "Interesting news in science from a bunch of degenerate grad students"
bayblab.blogspot.com

A Blog around the Clock – Rockin' Around the Circadian Clock
blog.coturnix.org

Bitesize Bio – Brainfood for Biologists
bitesizebio.com

"You know you've worked too long in a lab when…"

A Facebook group in which scientists have fun completing the sentence…You know you've worked too long in a lab when…

… when your scariest story deals with a walking ultracentrifuge that destroys walls and eats through thermocyclers.

… when you get more excited with a significant difference than with an invitation for dinner…

To find this and follow it, just go to Facebook.com and search "You know you've worked too long in a lab when…"

What is your opinion on scientists using social media?

- ☐ It helps me to collaborate
- ☐ Only in my free time
- ☐ I use it to keep up with scientific news and advancements
- ☐ I've tried it, but don't get it
- ☐ I don't get it, and don't plan to spend any time on it.
- ☐ Other

Participate
Please weigh in with our poll at wherebiobegins.com/biounity.

We bring our favorite social media tid bits to you

#scienceconfessions on Twitter. Scientists share their secret confessions (It's a great way to crack each other up!)

The smell of an overnight culture of E. coli makes me feel at home. #scienceconfessions
by JATetro (Jason Anthony Tetro)
The smell of an overnight culture of E. coli makes me feel at home. #scienceconfessions

RT @drisis: I secretly judge you when you misapply linear regression #scienceconfessions
by Canageek (Danforth Inferno)
RT @drisis: I secretly judge you when you misapply linear regression #scienceconfessions

To follow #scienceconfessions and more:
Go to search.twitter.com and search for #scienceconfessions

Bioabsolutely.

Create knockout rats and mice in as few as 5 months with Sigma's exclusive CompoZr® ZFN technology.

Utilize Sigma's SAGEspeed™ process to produce a custom knockout animal according to your specifications or make knockout animals in your own lab by using CompoZr ZFN reagents.

Create the knockout of your dreams with Sigma® Life Science.

sigma.com/transgenics

SAGE™ LABS
Sigma Advanced Genetic Engineering

Look for the next issue of *Biowire* featuring SAGE™ Labs

Enabling Science to Improve the Quality of Life

Order/Customer Service (800) 325-3010 • Fax (800) 325-5052
Technical Service (800) 325-5832 • sigma-aldrich.com/techservice
Development/Custom Manufacturing Inquiries SAFC® (800) 244-1173
Safety-related Information sigma-aldrich.com/safetycenter

World Headquarters
3050 Spruce St.
St. Louis, MO 63103
(314) 771-5765
sigma-aldrich.com

©2010 Sigma-Aldrich Co. All rights reserved. SIGMA, SAFC, SIGMA-ALDRICH, ALDRICH, FLUKA, and SUPELCO are trademarks belonging to Sigma-Aldrich Co. and its affiliate Sigma-Aldrich Biotechnology, L.P. Sigma brand products are sold through Sigma-Aldrich, Inc. Sigma-Aldrich, Inc. warrants that its products conform to the information contained in this and other Sigma-Aldrich publications. Purchaser must determine the suitability of the product(s) for their particular use. Additional terms and conditions may apply. Please see reverse side of the invoice or packing slip. CompoZr, Ingenuity, MISSION, Prestige Antibodies, and Seppro are registered trademarks of Sigma-Aldrich Biotechnology LP and Sigma-Aldrich Co. GenElute, Hybri-Max, ReadyMix, JumpStart, and SAGE are trademarks of Sigma-Aldrich Biotechnology LP. and Sigma-Aldrich Co. Genesifter is a trademark of Geospiza, Inc., Spotfire is a trademark of TIBCO Software Inc., Microsoft Excel is a trademark of Microsoft Corporation in the United States and other countries. TransPlex and OmniPlex are trademarks of Rubicon Genomics, Inc., SYBR is a registered trademark of Molecular Probes, Inc. Luminex is a registered trademark of Luminex Corp., STATLIA is a registered trademark of Brendan Technologies, Inc.

SIGMA-ALDRICH®

SIGMA® Life Science *Where bio begins*™

MXI
74457-510980
1090