Engineering Programmable Nucleases: Applications in the Study of Gene Function and eventually Gene Therapy

Scot Wolfe
Molecular, Cell and Cancer Biology
UMass Medical School
Frontiers in Science Presentation
March 18th, 2015
One of our fundamental areas of exploration is the development of *tools* for the targeted engineering of genomes.
One of our fundamental areas of exploration is the development of tools for the targeted engineering of genomes.

In analogy with surgical tools - we endeavor to make genomic scalpels that permit the precise modification of the genome.
One of our fundamental areas of exploration is the development of tools for the targeted engineering of genomes.

In analogy with surgical tools - we endeavor to make genomic scalpels that permit the precise modification of the genome.

These tools have utility for:
1) The study of gene function during vertebrate development
2) The correction of disease-causing genetic lesions
How do we locate a specific “address” within the genome?

Gene X
How do we locate a specific “address” within the genome?

Gene X
How do we locate a specific “address” within the genome?
How do we locate a specific “address” within the genome?

Gene X

Sequence-specific DNA recognition module
How do we locate a specific “address” within the genome?

Sequence-specific DNA recognition module

Where do we begin? - By understanding how naturally-occurring DNA-binding proteins locate their sequences.
Cys$_2$His$_2$ Zinc Finger proteins (ZFPs) are the most diverse family of DNA-binding domains in vertebrates.
Cys$_2$His$_2$ Zinc Finger proteins (ZFPs) are the most diverse family of DNA-binding domains in vertebrates.

How can we understand DNA recognition by these proteins?

Pavletich & Pabo Science 1991
Elrod-Erickson, Rouid, Nekludova & Pabo Structure 1996
A bacterial one-hybrid system can be used to determine DNA-binding specificity.

Bacterial Selection Strain

\[ \Delta rpoZ \]
\[ \Delta hisB \]
\[ \Delta pyrF \]
A bacterial one-hybrid system can be used to determine DNA-binding specificity.
A bacterial one-hybrid system can be used to determine DNA-binding specificity

Zinc Finger Protein

ZFP

Match

RNA polymerase

HIS3

URA3

10 bp

-35

-10

Match

ZFP+target

ZFP+ctrl

ZFP+target

ZFP+ctrl

Bacterial Selection Strain

ΔrpoZ

ΔhisB

ΔpyrF

DNA binding site selections are performed in a single round to yield interacting partners.

Overview of a B1H binding site selection

- Cell
- ΔrpoZ ΔhisB ΔpyrF
- Co-transform prey library and Bait plasmid into the selection strain (one prey/cell)
- Randomized binding site library
- pol-ZFP
DNA binding site selections are performed in a single round to yield interacting partners.

Overview of a B1H binding site selection.
DNA binding site selections are performed in a single round to yield interacting partners

Overview of a B1H binding site selection
DNA binding site selections are performed in a single round to yield interacting partners

Overview of a B1H binding site selection

[Diagram showing the process of DNA binding site selection with labels for cell, randomized binding site library, co-transform prey library, select active bait-prey pairs, isolate prey and sequence identify motif w/MEME]
We have determined the DNA-binding specificity of more than 100 naturally-occurring zinc finger proteins (ZFPs) and many more synthetic ZFPs.
We have determined the DNA-binding specificity of more than 100 naturally-occurring zinc finger proteins (ZFPs) and many more synthetic ZFPs.
We have constructed an archive of characterized zinc finger modules can be assembled into ZFPs to recognize a specific “address” within the genome.
We have constructed an archive of characterized zinc finger modules can be assembled into ZFPs to recognize a specific “address” within the genome.
We have constructed an archive of characterized zinc finger modules can be assembled into ZFPs to recognize a specific “address” within the genome.

Two Finger Archive (194 modules)

One Finger Archive (27 modules)

-TCGTTTCGGGCTCGCGCGTTAAAGCAAGTGCAGAA-
-AGCAAGGCCGGAGCGGCAATTTTCGTTTCACGTCTT-
ZFPs or other platforms can be used to create artificial nucleases that allow targeted gene editing.
ZFPs or other platforms can be used to create artificial nucleases that allow targeted gene editing
ZFPs or other platforms can be used to create artificial nucleases that allow targeted gene editing.
ZFPs or other platforms can be used to create artificial nucleases that allow targeted gene editing.
Zinc Finger Nuclease construction of disease models in zebrafish

**Diabetes:** leptin receptor *(db)*

miR-375 (1 & 2)

**Obesity:** melanocortin receptor 4 *(mc4r)*

*melanocortin receptor 3 *(mc3r)*

*leptin (A & B) *(ob)*
Zinc Finger Nuclease construction of disease models in zebrafish

Diabetes: \textit{leptin receptor (db)}

\textbf{miR-375 (1 & 2)}

Obesity: \textit{melanocortin receptor 4 (mc4r)}

\textit{melanocortin receptor 3 (mc3r)}

\textit{leptin (A & B) (ob)}
Pancreatic islet development is largely conserved between humans and zebrafish

miRNAs can inhibit translation and destabilize their target mRNAs.

Two Copies of *miR-375* are present in the zebrafish genome

zebrafish *miR-375-2* pri-miRNA hairpin

<table>
<thead>
<tr>
<th>ZFN target site</th>
<th>5’ ZFP site</th>
<th>3’ ZFP site</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>miR375-1</strong></td>
<td>TCGTTCCGGCTCGcgttaaGCAAGTGCAGAA</td>
<td>CGAGCCGAACGA</td>
</tr>
<tr>
<td><strong>miR375-2</strong></td>
<td>TCGTTCCGGCTCGcgttaGCAGATGCAGAC</td>
<td>CGAGCCGAACGA</td>
</tr>
</tbody>
</table>
Injecting nucleases into zebrafish embryos

Inject both ZFN mRNAs into 1-cell stage embryos
ObLiGaRe: precise NHEJ-mediated ligation of exogenous DNA sequences within the genome

ObLiGaRe: precise NHEJ-mediated ligation of exogenous DNA sequences within the genome

Maresca, et. al. (2013). Genome Research, 23(3), 539.
ObLiGaRe: precise NHEJ-mediated ligation of exogenous DNA sequences within the genome

Maresca, et. al. (2013). Genome Research, 23(3), 539.
ObLiGaRe: precise NHEJ-mediate ligation of exogenous DNA sequences within the genome

Maresca, et. al. (2013). Genome Research, 23(3), 539.
ObLiGaRe insertion into mir-375-2 locus

embryos (N/D)  N  N  N  N  D  N  D
ZFNs (pg)  0  63  0  63  63  125  125
Donor (pg)  0  0  25  25  25  25  25

1.3%
miR-375-2 knockin fish crossed to insulin:mcherry
miR-375-2 knockin fish allows pancreatic progenitor cells to be followed during development (23 to 30 hpf)
miR-375-2 knockin fish allows pancreatic progenitor cells to be followed during development (23 to 30 hpf)
A subset of miR-375 positive cells express insulin at 2 dpf

miR-375-2 Gal4FF; UAS: EGFP

ins: mcherry

overlay
miR-375-1^-/- miR-375-2^-/- animals display defects in pancreatic islet formation.

3 dpf - insulin:mCherry

WT double mutant
miR-375 null animals have reduced β-cell numbers at 3dpf
miR-375 nulls show elevated glucose levels at 6 dpf

WT

Nulls

miR-375 Genotype

whole embryo glucose concentration

p<0.001

18 μM

36 μM

Ankit Gupta
miR-375 positive cells can be sorted from wild-type and mutant animals to identify genes that are potentially responsible for the observed phenotypes.
Upregulated genes that are potential miR-375 targets

<table>
<thead>
<tr>
<th>Gene</th>
<th>Fold Change relative to WT</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELAVL4</td>
<td>4.3 3.5</td>
</tr>
<tr>
<td>ELAVL3</td>
<td>104.0 147.0</td>
</tr>
<tr>
<td>Ebf3</td>
<td>5.3 6.1</td>
</tr>
</tbody>
</table>
The CRISPR/Cas9 system provides a simple system for genome editing.
CRISPR/Cas9 system provides RNA guided nuclease activity

The primary constraint on target recognition is the PAM specificity of the Cas9 protein

S. pyogenes Cas9 (SpCas9): nGG >> nAG; nGA >> others

Development of chimeric Cas9 - DNA-binding domain (DBD) fusion

**Hypothesis:** Fusion of Cas9 to a programmable DNA-binding domain will provide improved genome editing precision and broader targeting range.
Development of chimeric Cas9 - DNA-binding domain (DBD) fusion

Hypothesis: Fusion of Cas9 to a programmable DNA-binding domain will provide improved genome editing precision and broader targeting range.

Goals:
1) Identify parameters to create functional Cas9-DBD framework
2) Attenuate Cas9 to make nuclease activity DBD dependent
GFP reporter assay to monitor nuclease activity in cell culture

Wilson et al. MTNA (2013) 2, e87
GFP reporter assay to monitor nuclease activity in cell culture

Wilson et al. MTNA (2013) 2, e87
GFP reporter assay to monitor nuclease activity in cell culture

mammalian cell

with functional nuclease

Wilson et al. MTNA (2013) 2, e87
Identifying parameters to create a functional Cas9-DBD framework
Identifying parameters to create a functional Cas9-DBD framework
Identifying parameters to create a functional Cas9-DBD framework
A DBD fusion expands the targeting repertoire of spCas9
Cas9-ZFP is functional on genomic targets with suboptimal PAMs

\[ \text{Cas9-ZFP} = \text{SpCas9-Zif268} \]
Cas9-ZFP is functional on genomic targets with suboptimal PAMs

Cas9-ZFP = SpCas9-Zif268

T7 Endonuclease I (T7EI) assay
To attenuate independent Cas9 activity mutagenesis was performed on the PAM contact residues.
Arg1333 and Arg1335 mutants cause loss of function which can be restored by a ZFP fusion
Arg1333 and Arg1335 mutants cause loss of function which can be restored by a ZFP fusion.
Arg1333 and Arg1335 mutants cause loss of function which can be restored by a ZFP fusion.
Can Cas9-ZFPs improve precision?

---

**Target ID** | **Gene** | **Full-length target (20 nt)** | **Indel mutation frequency (%)**
--- | --- | --- | ---
Target site 3 | VEGFA | GGTGAGTGAGTGTTGTGCGTGtGG | 54.08 (55.10, 53.06)
OT3-1 (abParts) | MAX | GGTGAGTGAGTGTTGTGaGG | 6.16 (6.71, 5.60)
OT3-2 | MAX | AGTGAGTGAGTGTTGTGgGG | 19.64 (18.58, 20.70)
OT3-4 | TPCN2 | GCTGAGTGAGTGATGCCTGtGG | 7.95 (7.84, 8.06)
OT3-9 | TPCN2 | GGTGAGTGAGTGCGTGgGG | Not Detected
OT3-17 | SLIT1 | GTTGAGTGAGTGTTGTGgGG | 1.85 (1.77, 1.92)
OT3-18 | COMDA | TGTGGGTGTGAGTGGTGcGaGG | 6.16 (6.72, 5.60)
OT3-20 | | AGAGATGAGTGAGTGACATgGaGG | 10.47 (9.39, 11.55)

---

**VEGF-A Cas9 TS3**

---

**ZFP target site**

---

Can Cas9-ZFPs improve precision?


<table>
<thead>
<tr>
<th>Target ID</th>
<th>Gene</th>
<th>Full-length target (20 nt)</th>
<th>Indel mutation frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target site 3</td>
<td>VEGFA</td>
<td>GGTGAGTGAGTGAGTGTCGCTGtGG</td>
<td>54.08 (55.10, 53.06)</td>
</tr>
<tr>
<td>OT3-1</td>
<td>(abParts)</td>
<td>GGTGAGTGAGTGAGTGTGTCGCTGtGG</td>
<td>6.16 (6.71, 5.60)</td>
</tr>
<tr>
<td>OT3-2</td>
<td>MAX</td>
<td>AGTGAGTGAGTGAGTGTCGCTGtGG</td>
<td>19.64 (18.58, 20.70)</td>
</tr>
<tr>
<td>OT3-4</td>
<td>TPCN2</td>
<td>GCTGAGTGAGTGAGTGTGTCGCTGtGG</td>
<td>7.95 (7.84, 8.06)</td>
</tr>
<tr>
<td>OT3-9</td>
<td>SLIT1</td>
<td>GGTGAGTGAGTGAGTGTCGCTGtGG</td>
<td>Not Detected</td>
</tr>
<tr>
<td>OT3-17</td>
<td>COMDA</td>
<td>TGTGAGTGAGTGAGTGTCGCTGtGG</td>
<td>1.85 (1.77, 1.92)</td>
</tr>
<tr>
<td>OT3-18</td>
<td></td>
<td>AGAGAGTGAGTGAGTGTCGCTGtGG</td>
<td>6.16 (6.72, 5.60)</td>
</tr>
<tr>
<td>OT3-20</td>
<td></td>
<td>AGAGAGTGAGTGAGTGTCGCTGtGG</td>
<td>10.47 (9.39, 11.55)</td>
</tr>
</tbody>
</table>

GTG GGTGAGTGAGTGAGTGTCGCTGtGG

VEGF-A Cas9 TS3

ZFP target site

Cas9-ZFPs increase the precision of genome editing
Cas9-ZFPs increase the precision of genome editing
Cas9-ZFPs increase the precision of genome editing
Artificial nucleases are powerful tools for manipulating vertebrate genomes in a directed manner.
Artificial nucleases are powerful tools for manipulating vertebrate genomes in a directed manner

1) Artificial nuclease platforms have been developed that can target the majority of sequences with a vertebrate genome.
Artificial nucleases are powerful tools for manipulating vertebrate genomes in a directed manner

1) Artificial nuclease platforms have been developed that can target the majority of sequences with a vertebrate genome.

2) These tools allow the study of gene function in model organisms and the creation of disease models to understand dysfunction at the systemic and molecular level.
Artificial nucleases are powerful tools for manipulating vertebrate genomes in a directed manner

1) Artificial nuclease platforms have been developed that can target the majority of sequences with a vertebrate genome.

2) These tools allow the study of gene function in model organisms and the creation of disease models to understand dysfunction at the systemic and molecular level.

3) More precise nucleases are being developed that will permit the realization of genetic correction of aberrant loci for the treatment of disease.
ACKNOWLEDGEMENTS

UMass Medical School (UMMS)
Molecular, Cell and Cancer Biology (MCCB)

Wolfe Lab
Fatih Bolukbasi
Ankit Gupta
Dana DeSantis
Victoria Hall
Selase Enuameh
Cong Zhu
Yuna Asriyan
Amy Rayla
Stephanie Chu
Heather Bell
Marcus Noyes

Brodsky Lab
• Michael Brodsky
Sarah Oiekemus

Nathan Lawson (MCCB)
Bioinformatics MCCB
Julie Zhu
Jianhong Ou

Bioinformatics Core (UMMS)
Manuel Garber
Alper Kucukural

Nick Rhind (BMP)

Maehr Lab (DCOE)
René Maehr
Nicola Kearns
Ryan Genga
Hannah Pham

Pederson Lab (BMP)
Thoru Pederson
Hanhu Ma

Jack Leonard
Karl Simin
Erik Sontheimer
Laura Alonso
Jeremy Luban

Funding:
NIGMS, NHLBI, NIAID & NHGRI