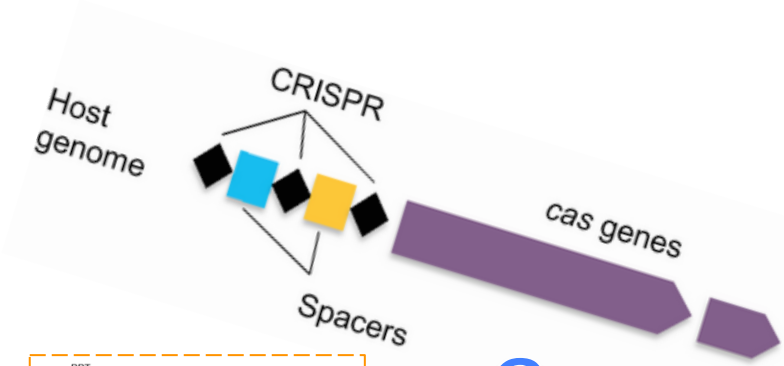
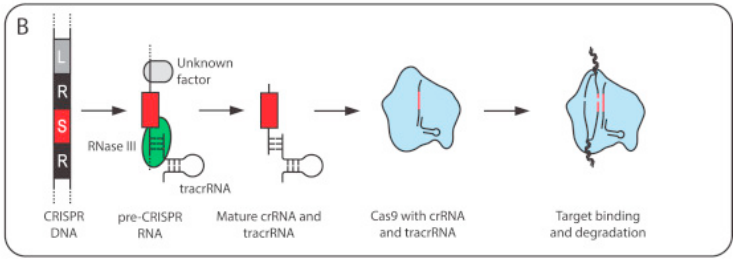
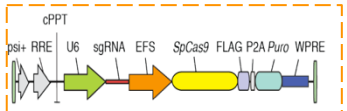


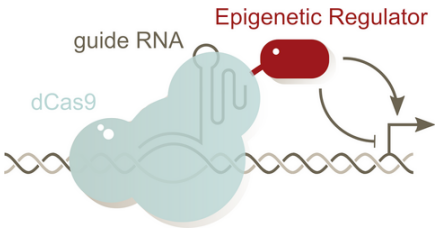




Genome wide CRISPR applications in Toxicology

Chris Vulpe

Amin Sobh, Graduate Student
Mani Tagmount, Scientist
Alex Loguinov, Bioinformatics
Quan Lu, Collaborator, Harvard University
Luoping Zhang, Collaborator, UC Berkeley
Martyn Smith, Collaborator, UC Berkeley



UF | Center *for* Environmental & Human Toxicology
College *of* Veterinary Medicine

UC BERKELEY SUPERFUND RESEARCH PROGRAM
SCIENCE FOR A SAFER WORLD

Srp Superfund Research Program

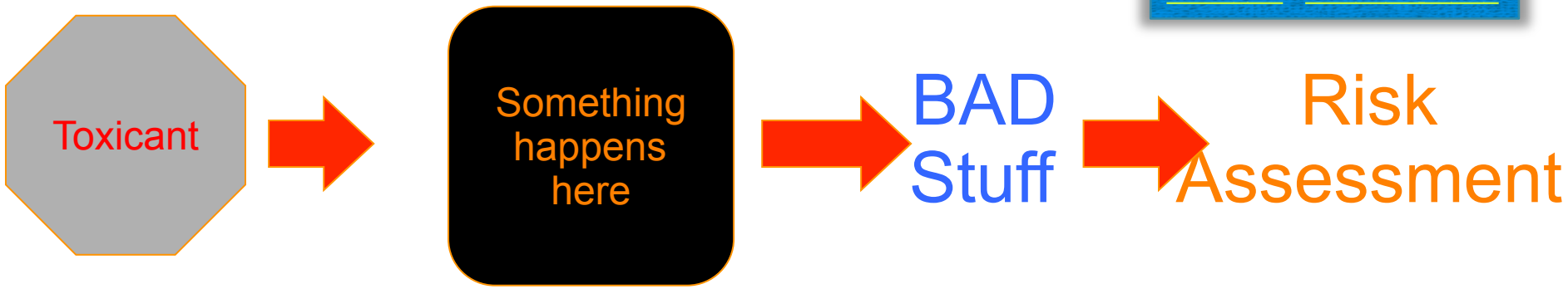
UF | UNIVERSITY of FLORIDA

Outline

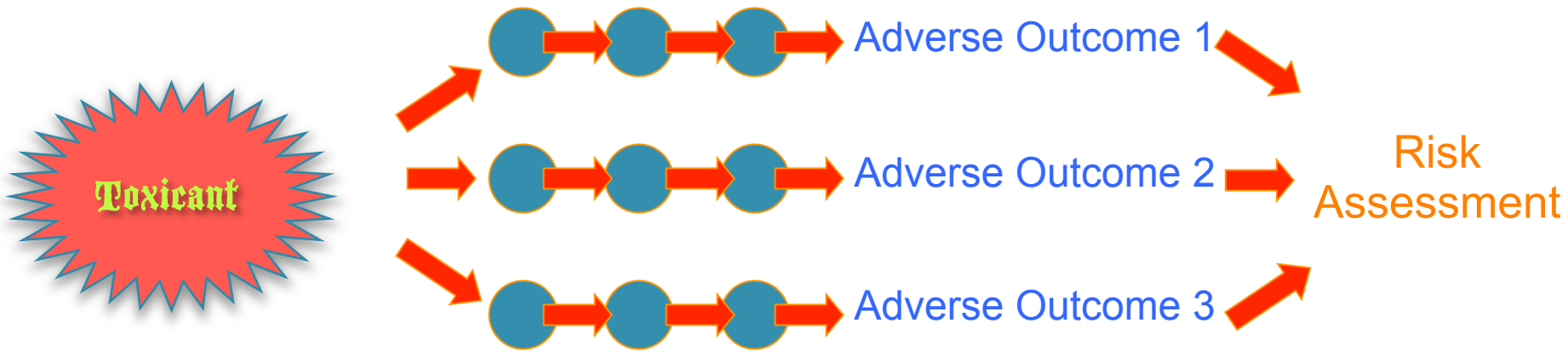
1. The world's shortest introduction to Adverse outcome pathways
2. Functional Toxicology – how to mess things up to learn something useful – you hope
3. Genome wide screens – how to mess up a whole lot of things, okay genes, all at the same time and of course, learn a whole lot more, really
4. Some examples from the lab -
   Acetaldehyde – Ethanol's nasty metabolite
   Arsenic – Toxic metalloid and cancer drug
5. Where we are headed – this is going to be an exciting and bumpy ride

Toxicity

Old School



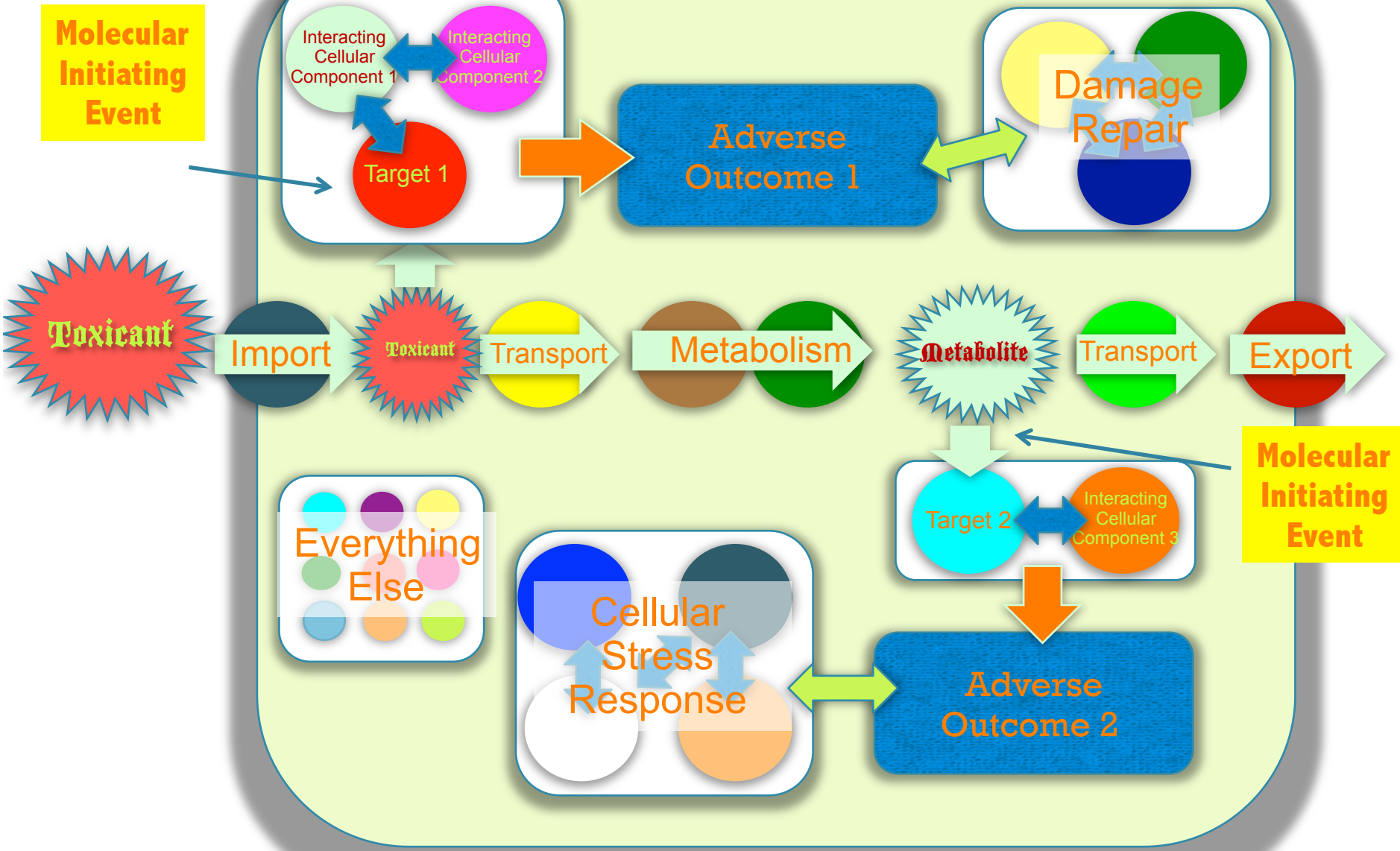
Cell, organism, population…

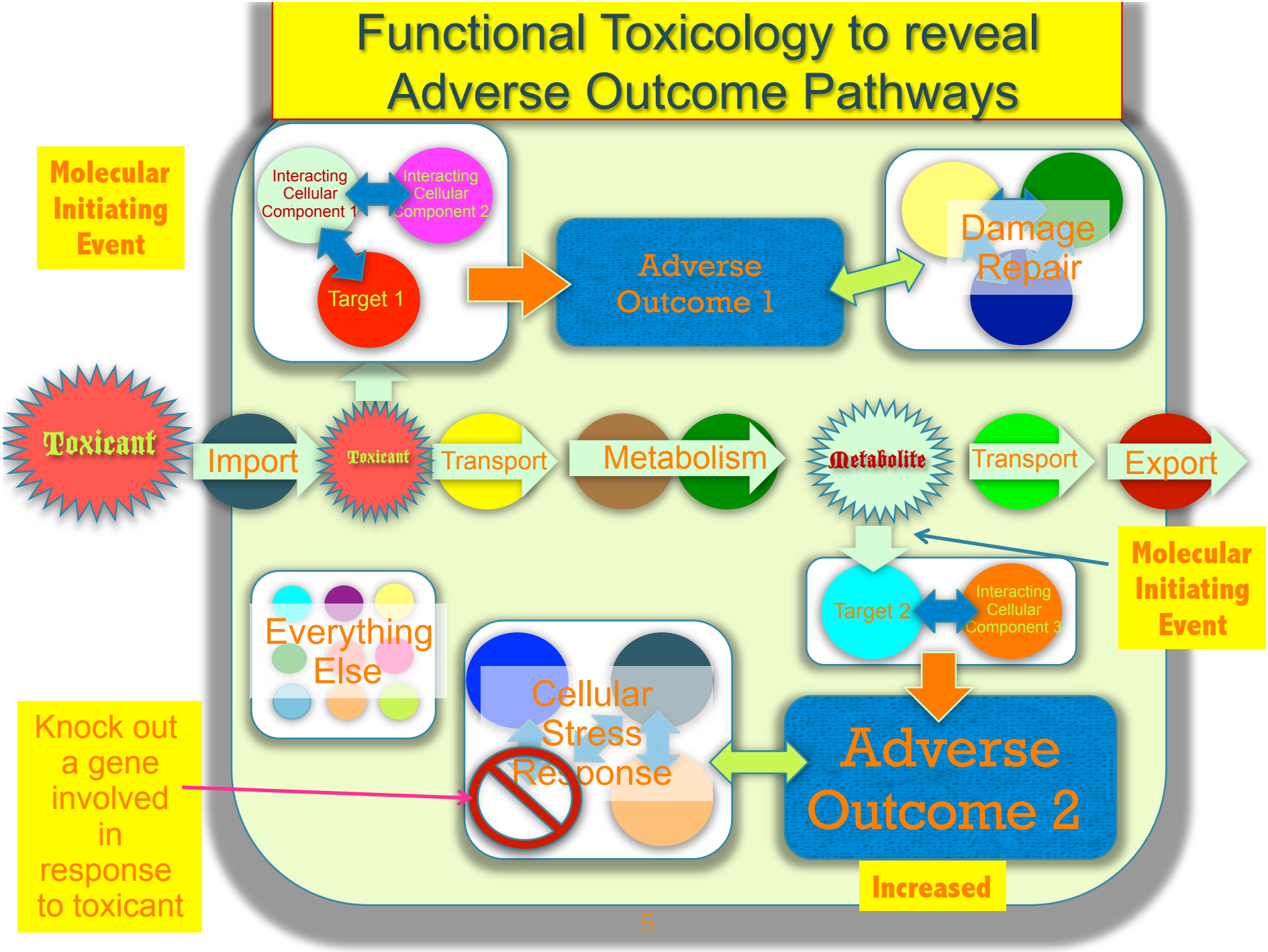


Adverse outcome pathways

Reformed School

Functional Toxicology to reveal Adverse Outcome Pathways



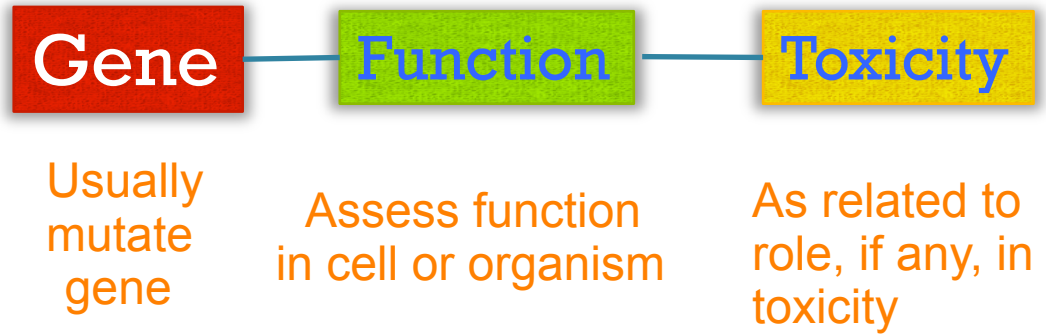

Functional Toxicology to reveal Adverse Outcome Pathways
Molecular Initiating Event
Interacting Cellular Component 1
Interacting Cellular Component 2
Target 1
Adverse Outcome 1
Damage Repair
Toxicant
Import
Toxicant
Transport
Metabolism
Metabolite
Transport
Export
Molecular Initiating Event
Target 2
Interacting Cellular Component 3
Everything Else
Cellular Stress Response
Adverse Outcome 2
Increased
Knock out a gene involved in response to toxicant

Functional Profiling in Toxicology

the study of the requirement for the biological activities of genes and corresponding proteins in the response to, and effect on, an organism by a toxicant

OR if you muck it up (the gene) & bad (or good) things happen, then it's probably important

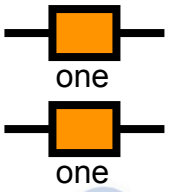
Functional Profiling – systematically testing multiple genes for their functional role, if any, by perturbing their function



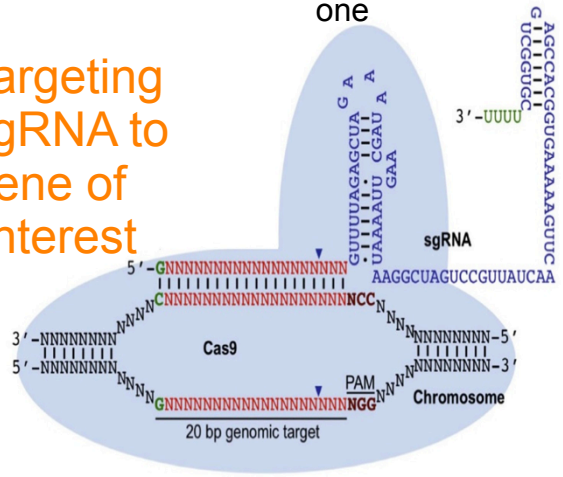
Goal is to link gene through its function to observed toxicity
e.g. molecular target, transport, metabolism, cellular response, etc

Targeted CRISPR vs Genome Wide CRISPR

Gene of Interest

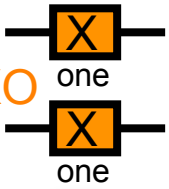


Targeting sgRNA to gene of interest



CAS9

Introduce into cell/organism

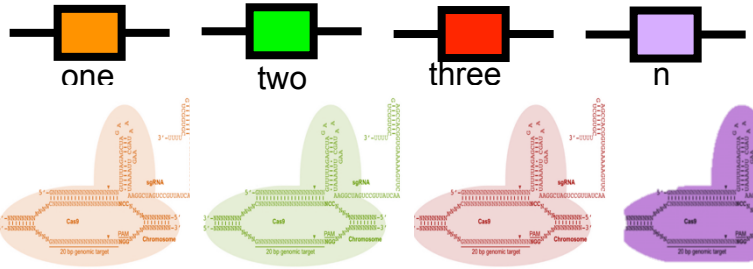


Screen for KO

Isolate KO cell/organism

Assess Function in cell/organism

Multiple genes of interest

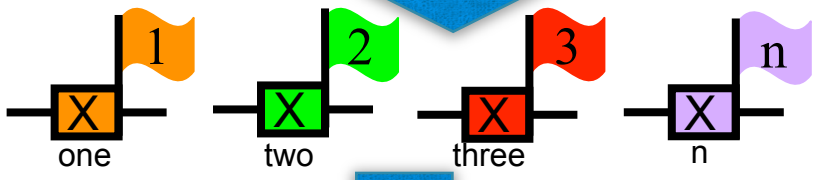


Targeting sgRNA to genes of interest

(CRISPR KO library)

Introduce into cell

One per cell

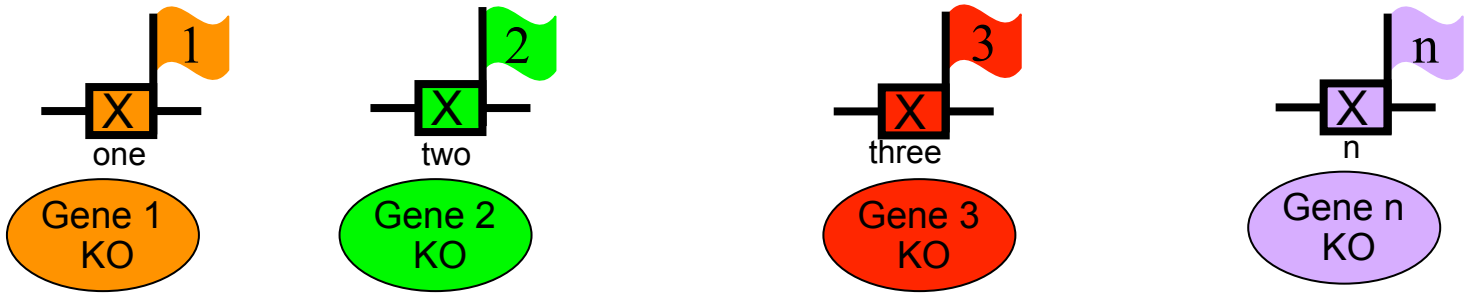


Generate pool of mutants

Screen for sensitivity to toxicant to identify and RANK the important genes

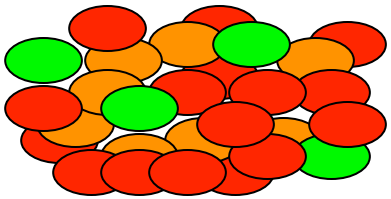
Genome Wide CRISPR in Toxicology

Each KO is individually flagged with a unique molecular barcode so they can be tracked



Compare growth of each KO in toxicant to growth without toxicant

Grow pool with



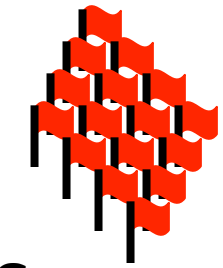
Collect and count flags

Grows okay

Not involved

Grows poorly

Sensitive



Grows well

Resistant

Genome-wide Loss-of-function Screening (the weeds)

Okay – you are convinced – GW CRISPR is way cool

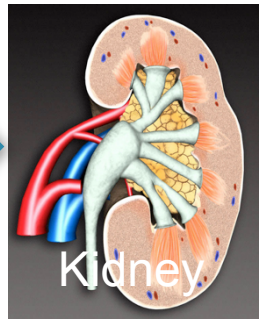
Decisions, Decisions, and Issues, even more issues

- What organism?
  - Pre-made Tools for Humans, Mice and Rats
  - But can generate CUSTOM tools for any organism with a sequenced genome
- What genes do you want to target?
  - All of the genes? Genome wide – every gene in the genome
  - Only a subset of genes – e.g the genes which encode proteins which carry out a specific set of functions – KREBS cycle
- How much do you trust a COMPUTER?
  - Selecting the targeting guides is done by computer algorithms
  - JUST TOO MANY to test them all with experiments
  - Some guides may not work – NO TARGET
  - Some guide may misdirect the CAS9 – OFF TARGET
  - Need to use MULTIPLE guides (3-10)
  - Redundancy – HOPE that some work and increases CONFIDENCE if you see same thing with different guides to the SAME gene.

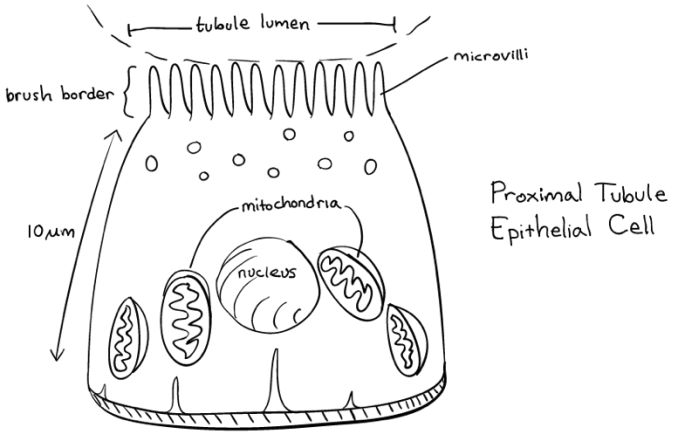
Science. 2014;343(6166):84-7.

Genome-wide Loss-of-function Screening (more weeds)

- You picked the organism, thought you were clever, well now what tissue, what cell?



Tissue/Organ



Cell type

In fact, most whole genome CRISPR KO screens use CANCER Cell Lines

Why? They grow *in vitro,* they grow *FAST*, and *its CHEAPER*
But lots of issues with USING CELL LINES
FOR TOXICOLOGY – a big one is POOR or NO Metabolism

PRIMARY CELLS ARE POSSIBLE BUT HARD TO GROW, SLOW, EXPENSIVE

- AND finally what DOSE to USE? And for HOW LONG?

NOT TOO HIGH – KILL EVERYTHING / NOT TOO LOW – EVERYBODY HAPPY
NOT TOO LONG – Yup, same reason / NOT TOO SHORT – ahh, cleansing bath

CONTROL – Always comparing to growth of each KO in absence of toxicant

Only a few published Genome wide CRISPR screens related to Toxicology

Cell line	# Genes	Toxicant	Reference
HL-60 (Human AML Leukemia)	7114	6-TG, etoposide	Science 2014, 343, 80–84
A375(melanoma) HUES62 (ES cell) Human	18000	BRAF inhibitor	Nature 2015, 517, 583–588
Mouse ES	18000	6-TG	Nat. Biotechnol. 2014, 32, 267–273.
K562 Human red blood cell leukemia	16000	DPT	Cell 2014, 159, 647–661
HepG2 Human liver cancer	18080	Triclosan	EST,2016;50(19): 10682-92

Generally CANCER cells and CANCER drugs

Acetaldehyde and Arsenic Trioxide Toxicity

Acetaldehyde

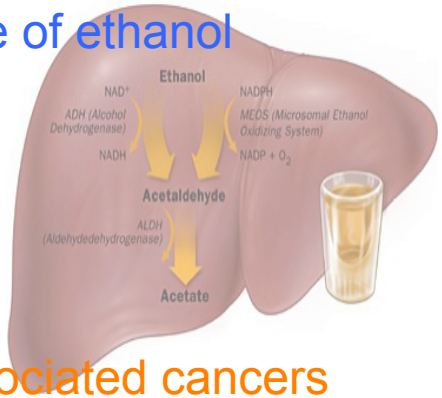
- Primary oxidative metabolite of ethanol

  Genotoxic

- Group 1 carcinogen (IARC)

- Likely underlies alcohol-associated cancers

- Mechanisms of toxicity are poorly understood



Arsenic Trioxide

-Arsenic Trioxide used in blood cance treatment – metabolised to usual As metabolites

- Arsenics - IARC type I carcinogen

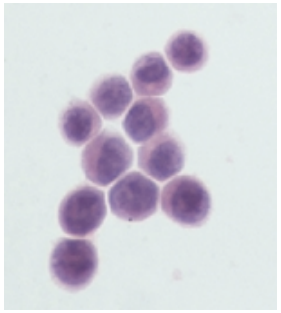
- Drinking water exposure

- Mechanisms still controversial (depending on who you ask)

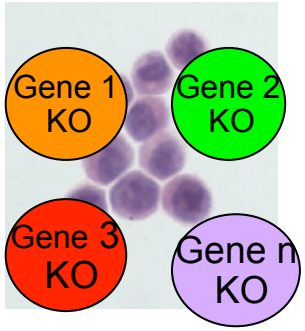
An abundance of mechanisms – there are too many mechanisms and it is unclear which are the most important

Can whole genome CRISPR give us some insights into the cellular mechanisms and maybe their relative importance?

What did we do?



K562 (human PRE-RED BLOOD CANCER cell line)



Knock out Library

~18,000 genes

Arsenic Trioxide Control

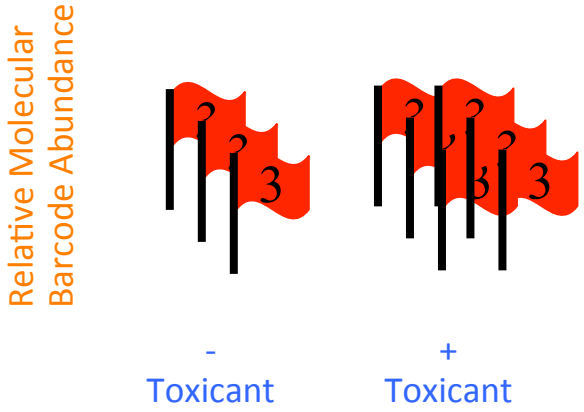
Acetaldehyde Control

Use Dose that decreases growth by ~50% Of NORMAL cell

Collect the cells that survive

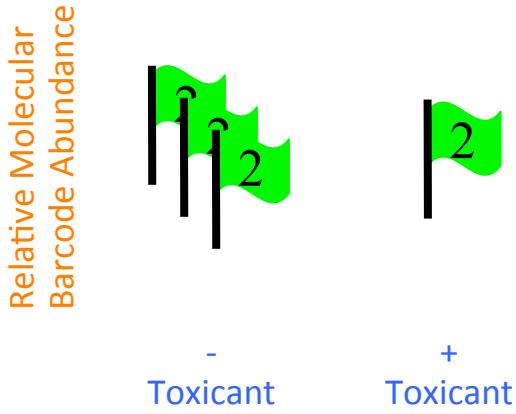
Count molecular BARCODES from Toxicant/Control (SEQUENCING)

Resistant Mutants



Some mutant CRISPR KO mutants will enriched in presence of toxicant as compared to its frequency in control

Sensitive Mutants



Some mutant CRISPR KO mutants will depleted in presence of toxicant as compared to its relative abundance in control

Arsenic Trioxide whole genome CRISPR screen

(DATA SLIDE – PROVES WE ACTUALLY DID SOMETHING)

Top 10 - Whole Genome Screen Candidates

Log FC – relative abundance in treated vs control

Gene ID	Gene name	logFC	P Value	FDR
KEAP1	kelch-like ECH-associated protein 1	2.05	3.13E-59	6.87E-55
SEPHS2	selenophosphate synthetase 2	1.77	1.88E-23	2.06E-19
EEFSEC	eukaryotic elongation factor, selenocysteine-tRNA-specific	1.25	1.09E-17	7.97E-14
PSTK	phosphoseryl-tRNA kinase	1.49	3.23E-17	1.77E-13
KRT73	keratin 73	-2.5	2.88E-15	1.26E-11
ARID1B	AT rich interactive domain 1B (SWI1-like)	1.42	5.44E-13	1.99E-09
TXNDC17	thioredoxin domain containing 17	0.9	3.20E-10	1.00E-06
SLC6A12	solute carrier family 6 (neurotransmitter transporter), member 12	0.92	8.66E-10	2.37E-06
DCLRE1A	DNA cross-link repair 1A	-1.1	5.52E-09	1.34E-05
DLGAP5	discs, large (Drosophila) homolog-associated protein 5	-1.1	2.91E-08	6.38E-05

KEAP1 – NRF2 Partner – involved in oxidative stress

Selenocysteine metabolism

DNA Repair

FDR – False Discovery Rate

Confirmatory Screen

Resistant

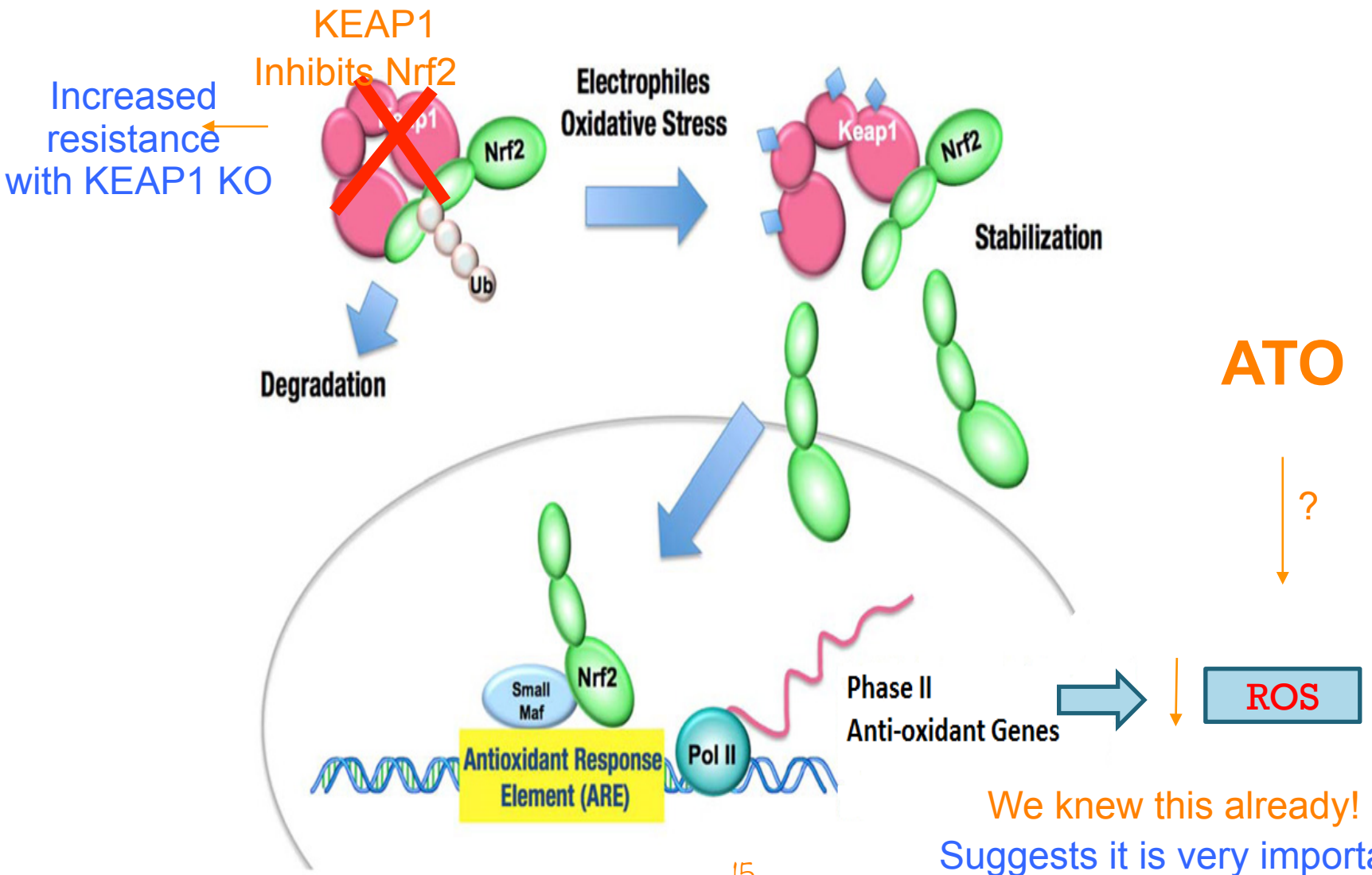
Gene	sgRNA	FDR	Log FC
KEAP1	8/8	0.000354	3.6
TXNDC17	8/8	0.000354	1.4
PSTK	7/7	0.000354	1.6
GFI1B	7/7	0.000354	1.1
SLC30A1	7/7	0.000354	1
FLCN	7/7	0.000354	1.3
EED	7/7	0.000354	0.7
RRAGC	8/8	0.000354	1
EEFSEC	6/7	0.000354	1.6
C15orf41	7/7	0.000354	0.6
SET	7/8	0.000354	0.8
SEPHS2	6/7	0.000354	1.4
SEPSECS	7/8	0.000354	0.7
DPH6	6/7	0.000354	0.8
NAA38	8/8	0.000928	0.7

Sensitive

Gene	sgRNA	FDR	Log FC
ABCC1	8/8	0.000619	-2.1
MTPN	7/7	0.000619	-0.7
NCAPD3	6/7	0.000619	-0.7
DEPDC5	7/7	0.000619	-0.4
UBE2H	7/8	0.000619	-0.6
NPRL2	6/6	0.000619	-0.3
CNOT2	7/7	0.000619	-0.6
NDE1	7/8	0.000619	-0.7

ATO Toxicity: Reactive Oxygen Species

Nrf2 primary anti-oxidant transcription factor – KEAP1 is REPRESSOR of NRF1

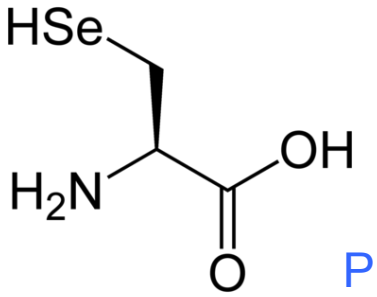


We knew this already!

Suggests it is very important in ACUTE short term toxicity

Mitsuishi *et al., Front. Oncol.*, 2012

Selenocysteine Incorporation into Proteins Increases Susceptibility to Arsenic Trioxide



Selenocysteine
The 21st amino acid

Selenium
Previously known As binds Se

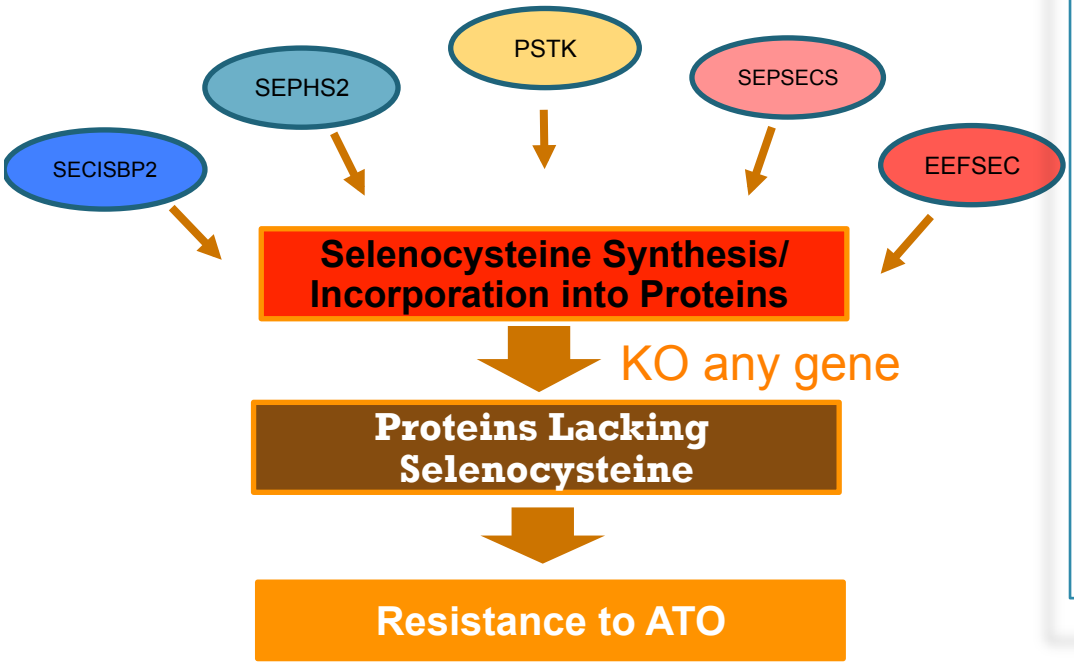
Specialized, totally cool, system for putting them in proteins

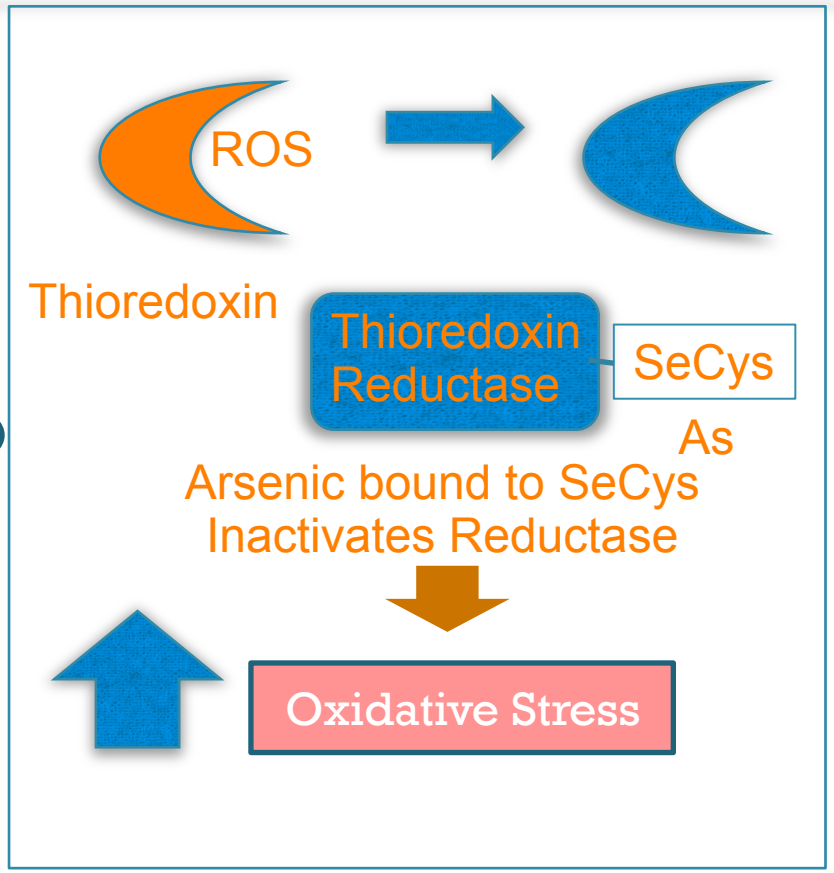
Many of these proteins involved in response to oxidative stress

What did we find?

Loss of any gene needed for SeCys Incorporation in proteins leads to RESISTANCE

THIS IS NEW AND KINDA UNEXPECTED



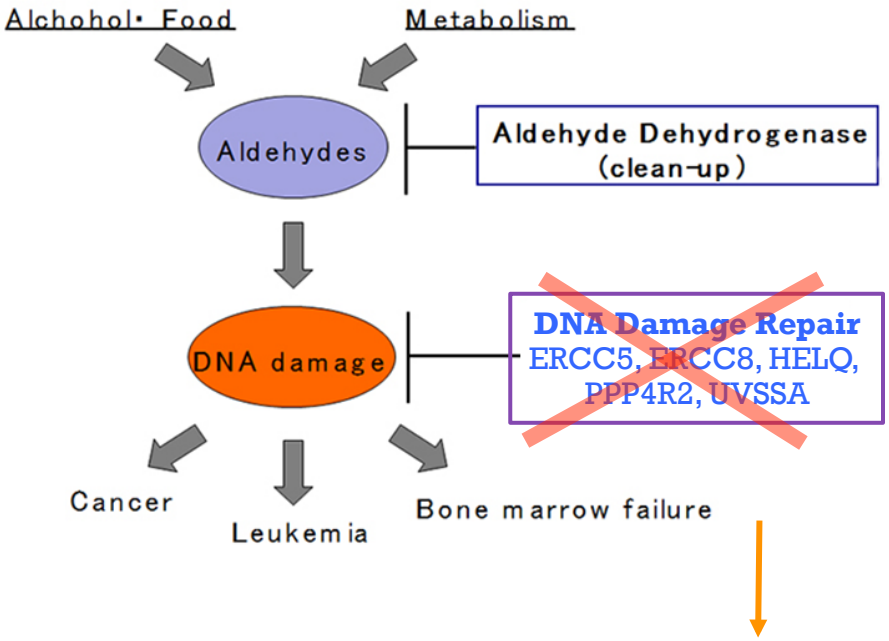


Blocking DNA Repair Increases Susceptibility to Acetaldehyde

Validated Acetaldehyde Susceptibility Candidates

Gene	sgRNA sequence	FDR
OVCA2	GCCGAGCTCGTGTGCCTCAG	4.78E-12
OVCA2	GACACCAAGAGGATAAACCG	8.91E-10
OVCA2	TGTCTCACCGAAGTCTGATC	8.91E-10
HELQ	TGCTGGAATAGATACTATTG	1.39E-08
HELQ	GGAGTTGCCTATCACCACAG	1.20E-07
OVCA2	CGGGGCTTCCGTGAGAAGAC	9.18E-07
HELQ	GTTGACAGCAAAGCTGAGAA	9.22E-07
HELQ	TCCTGATCACTTGGTAGCAT	9.73E-07
OVCA2	TTCCAATGCGGAGAAAACGT	9.73E-07
OVCA2	GGGCTTCCGTGAGAAGACCG	1.66E-06
HELQ	TGAAGTATATCATCCAATCA	4.16E-06
ERCC8	GCCAAGATATAGTCATAACG	0.000198
ERCC5	TTAATGGCTGAAAGAGTCCG	0.0003243
ERCC8	CAGTGGTATCCTCATGACAC	0.0003705
PPP4R2	CATGACAAAGAAACTGATCC	0.0003731
PPP4R2	TCACATTGTTTCTCCAGTCT	0.0005561
OVCA2	GAGGGCGCCAGATCAGACTT	0.0008455
HELQ	ACCAATGCTACCAAGTGATC	0.0022578
OVCA2	CAACTGGCCAGCCAATTTCC	0.0025663
ERCC8	TGTAAAGCAGTGTGTTCCAT	0.0028369
NANS	GAGATCGGCCAGAACCACCA	0.020249
NANS	TATGTGACGTTCCAACACCT	0.0524015
FBXO40	AACCTCCGGCTTAATGGCAA	0.0648084
UVSSA	AATTGAATCCTGCTTGACGG	0.0769292
HELQ	CTTATCTCTTACCTTCGAGC	0.0776389

Acetaldehyde-induced DNA Damage in Blood Precursors



Increased Susceptibility to Acetaldehyde Toxicity

Something OLD

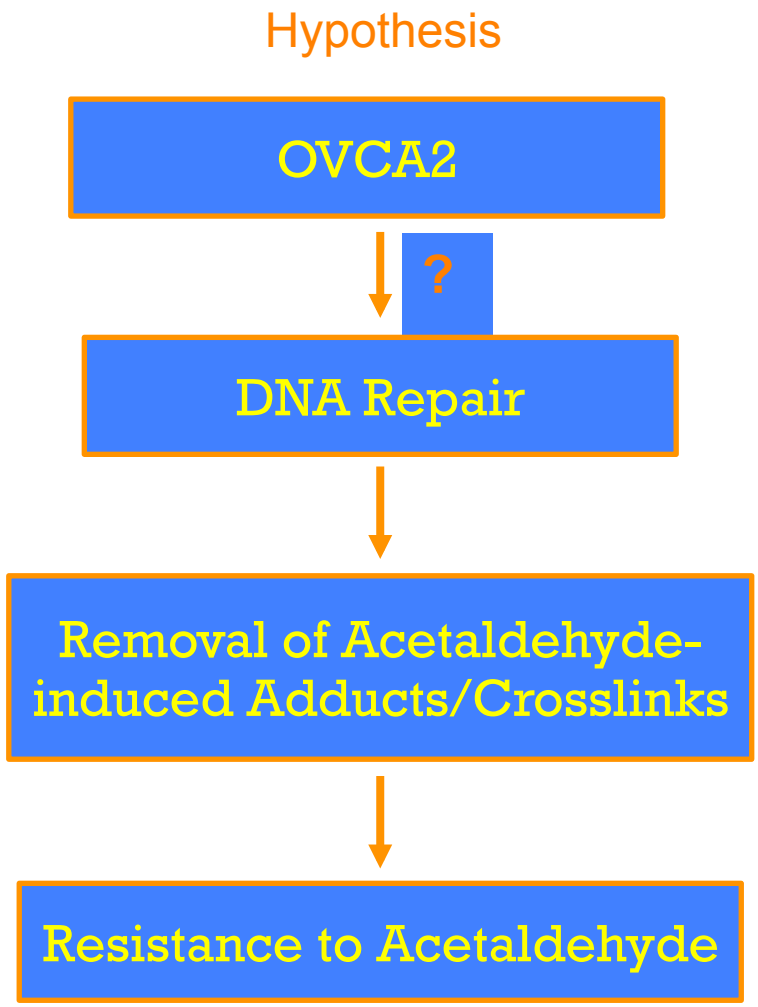
Modified from: Hira *et al.*, *Blood*, 2016

Potential Role of *OVCA2* in DNA Repair

OVCA2

- Ovarian tumor suppressor candidate 2
- Strongest candidate in our screen
- Validated with 8 sgRNAs out of 8 in a secondary screen
- Loss-of-function increases sensitivity to Acetaldehyde
- Predicted hydrolase (esterase) activity
- Downregulated in multiple cancer types
- Yeast homolog (*FSH1*) is essential for growth in ethanol media
- What the heck is it?

Something NEW, and blue of course



What did we learn?

- Oxidative stress is (the) major player in acute arsenic toxicity
- Selenium metabolism is important in acute arsenic toxicity
- DNA damage is important in acute acetaldehyde toxicity
- Unexpected insight into OVCA2 – a new DNA repair gene?

Implications?

- Help fill in adverse outcome pathway for Arsenic

As → ROS → BAD THINGS

Antioxidants could help?

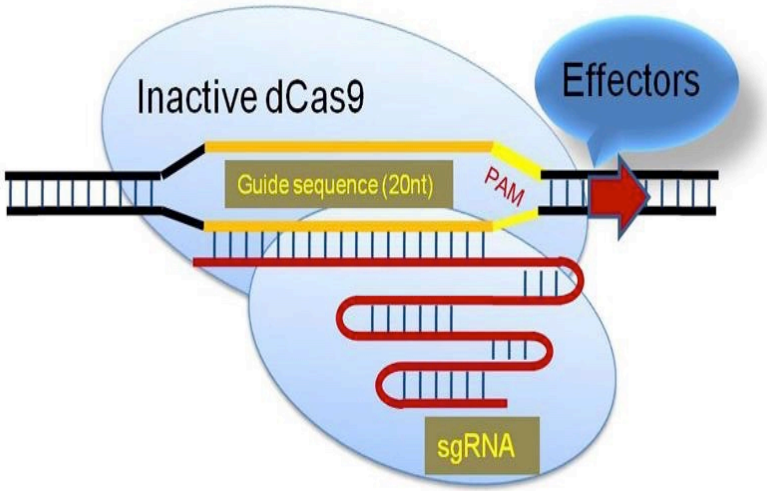
- Suggest selenium deficiency could decrease acute effects and selenium sufficiency could increase acute effects –Public Health Implications?
- Acetaldehyde – confirm genotoxic mechanism – suggest DNA damage also important for Acute toxicity
- OVCA2 is tumor suppressor gene– maybe role in DNA repair explains why

The FUTURE

ToxCRISPR

- Quan Lu – Harvard
- 3675 Toxicology-related genes
- Subset CRISPR library for probing toxicology mechanisms
- Can use less cells – need only 7.5 million vs 30 million for whole genome
- Enable more rapid screening of chemicals with more doses

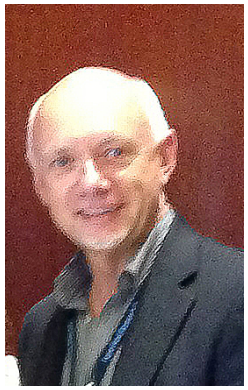
Alternative CRISPR-Cas9 Platforms



- Activate transcription (VP64)
- Repress transcription (SRDX)
- Cargo delivery
- DNA labeling (GFP)
- Epigenetic modification (DNA demethylase)

Acknowledgments

Luoping Zhang



Martyn

UF | Center *for* Environmental & Human Toxicology
College *of* Veterinary Medicine

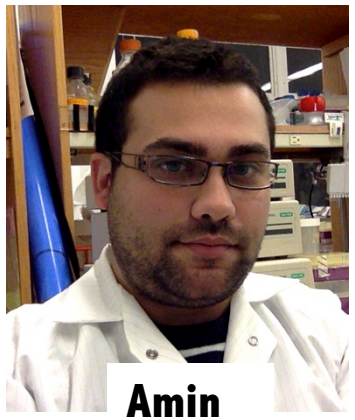
NIEHS

UC BERKELEY SUPERFUND RESEARCH PROGRAM
SCIENCE FOR A SAFER WORLD

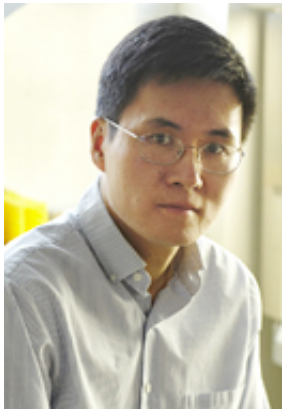
Srp Superfund Research Program

CRISPR

Scientist



Graduate Student



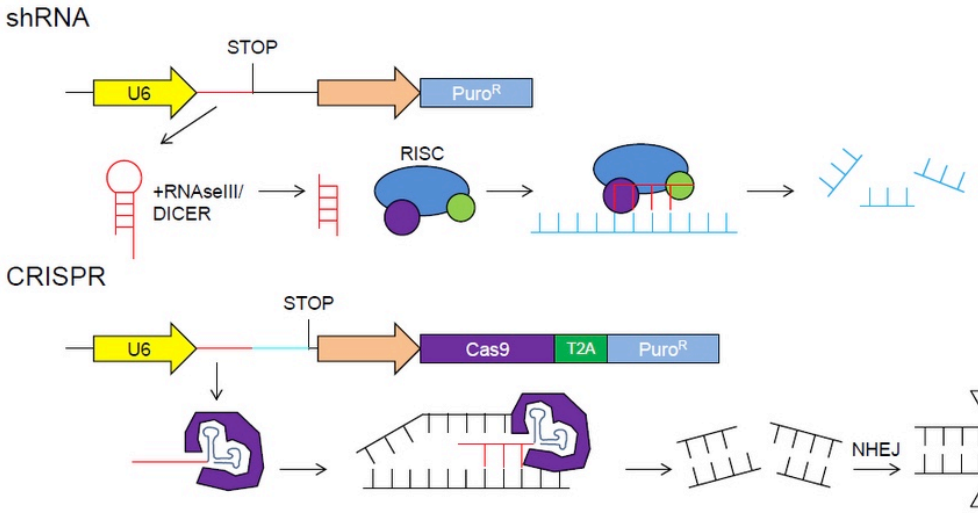
Quan Lu
Collaborator
Harvard

Bioinformatics

UF | UNIVERSITY of FLORIDA

How does whole genome CRISPR KO compare to other genome wide functional approaches?

Two recent papers compared RNA silencing with CRISPR KO



RNA silencing

Knock down of RNA level

CRISPR KO

Deletion/Mutation of gene

Opposite conclusions

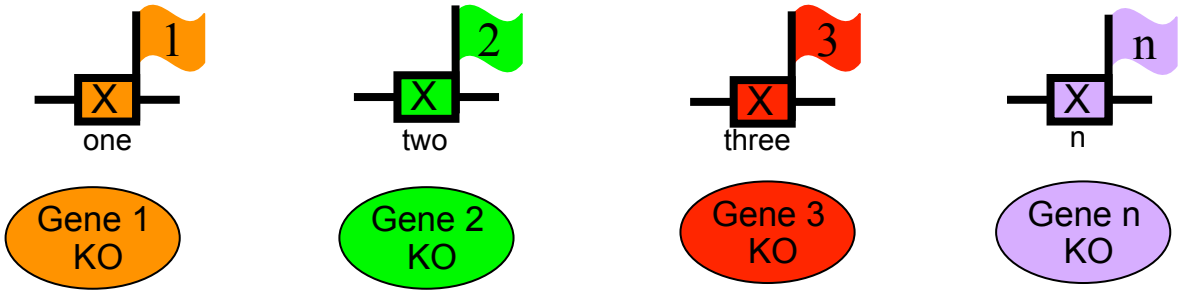
One paper concluded CRISPR is superior
Other paper found shRNA more reliable

1. Evers B, Jastrzebski K, Heijmans JPM, Grernrum W, Beijersbergen RL, Bernards R. CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. Nat Biotech. 2016;34(6):631-3. doi: 10.1038/nbt.3536 http://www.nature.com/nbt/journal/v34/n6/abs/nbt.3536.html - supplementary-information.
2. Housden BE, Perrimon N. Comparing CRISPR and RNAi-based screening technologies. Nat Biotech. 2016;34(6):621-3. doi: 10.1038/nbt.3599.
3. Morgens DW, Deans RM, Li A, Bassik MC. Systematic comparison of CRISPR/Cas9 and RNAi screens for essential genes. Nat Biotech. 2016;34(6):634-6. doi: 10.1038/nbt.3567 http://www.nature.com/nbt/journal/v34/n6/abs/nbt.3567.html - supplementary-information.

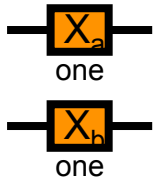
Key concepts/confusions in genome wide CRISPR screening

- *" In vitro "* - Using cell lines with all the accompanying issues and caveats –
  - e.g. metabolism, immortalized cells, toxicokinetics

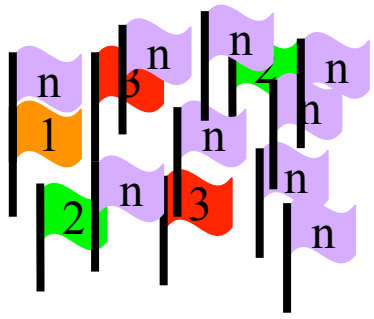
- Any or every gene can be targeted in your library BUT
- Only a single gene is inactivated (KO) in each cell
- A pool (library) of individual mutant cells each containing a KO of single gene represents all genes



- The gene on each chromosome are KOd but the mutations are different on each chromosome



- Each cell with a KO is TAGGED/FLAGGED with unique DNA barcode (sgRNA) so you can see it in a crowd (pool)



- Generally measuring growth advantage or disadvantage of mutant cells in response to environmental exposure such as toxicant