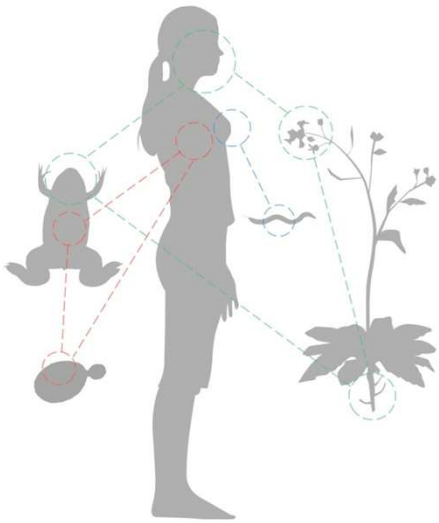
Phenologs

An example of using bioinformatics to find new genes for genetic traits

BIO337 Systems Biology / Bioinformatics – Spring 2014

Edward Marcotte, Univ of Texas at Austin

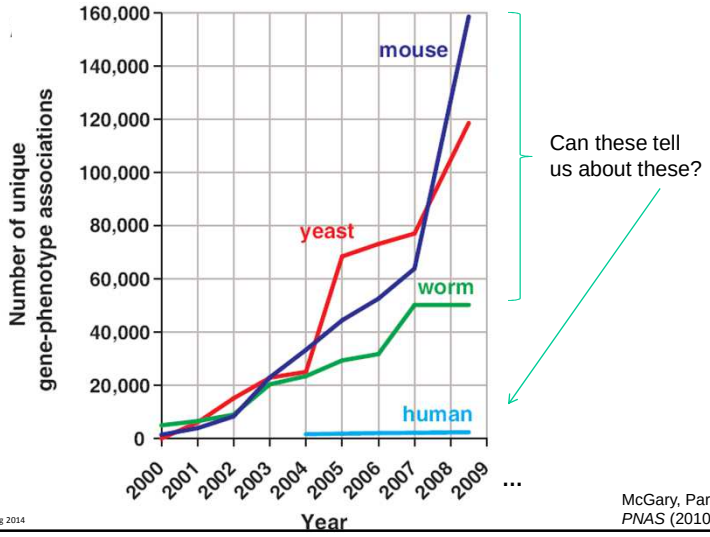


Virtually all genetic traits and diseases affect molecular structures that are evolutionarily conserved.

Consequently, human traits and diseases often have equivalents in other species, even distant ones.

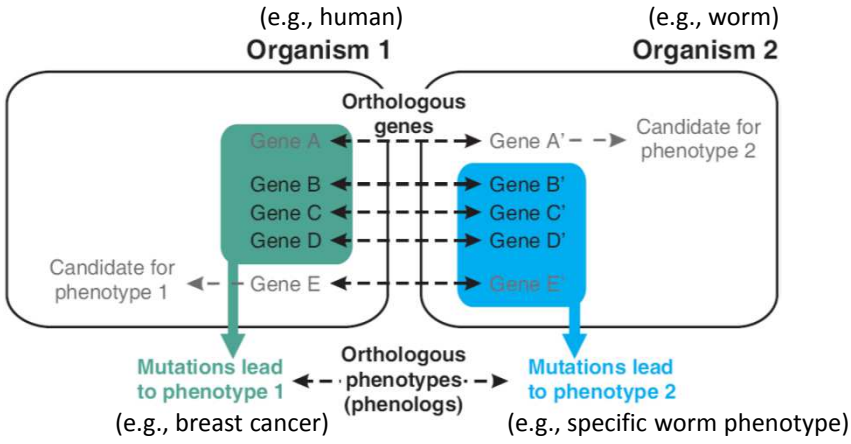
We know far more about genes & traits in lower organisms than in us.

How do deeply conserved gene networks relate to traits, structures, and diseases in different organisms?



McGary, Park *et al*. *PNAS* (2010)

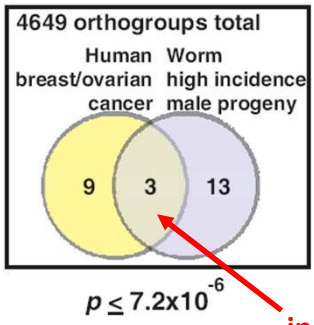
Phenologs = significantly overlapping sets of orthologous genes, such that each gene in a given set gives rise to the same phenotype in that organism



McGary, Park *et al*. *PNAS* 107:6544-9 (2010)

An example phenolog: a high incidence of male *C. elegans* maps to human breast/ovarian cancers

Human/Worm Ortholog	Linked to breast cancer in humans	Linked to more males in worms
ATM/atm-1	X	
BRIP1/dog-1	X	
KRAS/let-60	X	
PHB/phb-1	X	
PIK3CA/age-1	X	
RAD51/rad-51	X	
RAD54L/rad-54	X	
SLC22A18/C53B4.3	X	
TSG101/tsg-1010	X	
BARD1/brd-1	X	X
BRCA1/brc-1	X	X
CHEK2/chk-2	X	X
FAM82B/F33H2.6		X
GCC2/hcp-1,hcp-2		X
HMG20A,B/W02D9.3		X
HORMAD2,1/him-3,htp-1,2		X
KIF15/klp-10,18		X
MRE11A/mre-11		X
PIGA/D2085.6		X
RAD1/mrt-2		X
RAD21/coh-1		X
SEH1L/npp-18		X
SVIL/viln-1		X
TSPO,BZRPL1/C41G7.3		X
WDHD1/F17C11.10		X



includes BRCA1

McGary, Park *et al*. *PNAS* 107:6544-9 (2010)

Building & searching a collection of phenotypes

Mining available databases +
manual collection from the primary literature

↓

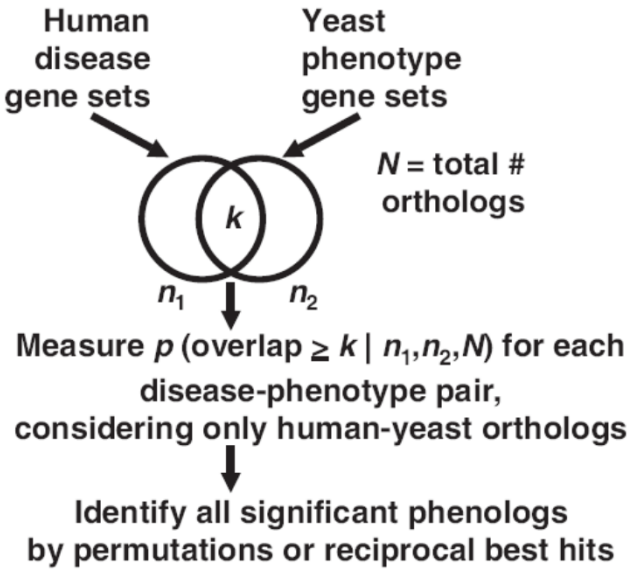
Organism	# gene-phenotype associations
human	1,923
mouse	74,250
worm	27,065
yeast	86,383
Arabidopsis	22,921

Spanning ~300 human diseases,
>7,000 model organism mutational phenotypes

↓

Computational scan phenotypes for novel models of a disease of interest, identify significant phenologs using permutation tests McGary, Park *et al*. *PNAS* (2010)

Discovering phenologs



McGary, Park *et al*. *PNAS* (2010)

There are 1,000's of phenologs between human diseases and mouse, yeast, worm, and even plant traits

Some cases we knew about already, serving as positive controls…

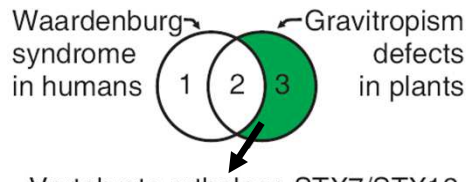
For example, genes for mouse cataracts suggest genes for human cataracts...

But many cases were surprising!

A defect in...	suggests genes for ...
yeast lovastatin sensitivity	**angiogenesis defects**
worm abnormal body wall muscle cell polarization	**gastrointestinal hemorrhage**
yeast hydroxyurea sensitivity	**hemolytic anemia**
plant cotyledon development defects	**mental retardation**
***E. coli* chemical sensitivies**	**chemically-induced seizures**

McGary, Park *et al*. *PNAS* 107:6544-9 (2010)
Woods, Blom *et al. BMC Bioinformatics*, 14:203 (2013)

Example #2: plant *negative gravitropism* defects predict *Waardenburg syndrome*, a congenital disease with characteristic craniofacial, hearing, and pigmentation alterations



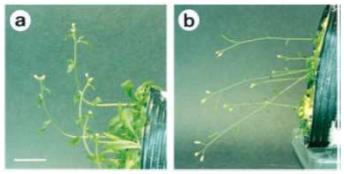
Vertebrate orthologs STX7/STX12, DDHD2/SEC23IP, and DNAJC13 are candidate Waardenburg genes

Michael Murphy, M.D.

Assorted websites

Waardenburg syndrome
(accounts for ~2-5% of cases of deafness)

≈

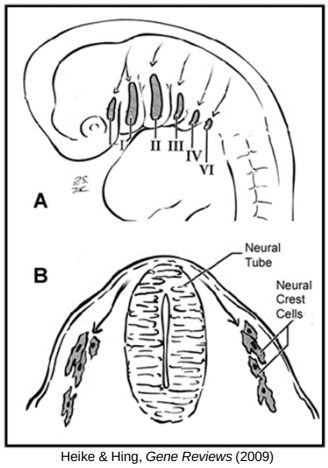


Fukaki *et al., The Plant Journal* 14, 425–430 (1998)

Plants failing to grow upwards

Waardenburg syndrome is a defect of neural crest cells

Neural crest cells migrate during embryonic development



Heike & Hing, *Gene Reviews* (2009)

Some WS correlates in other animals:

Deafness in Dalmatian dogs (22% unilaterally deaf)

www.petplanet.co.uk

Variations in the Blenheim spot of Cavalier King Charles Spaniels

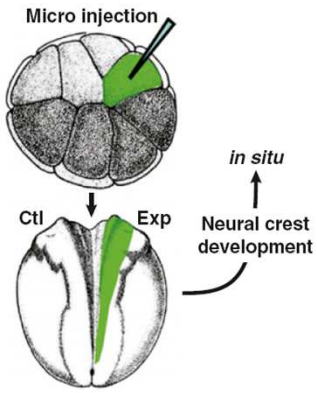
www.silvarcea.co.uk

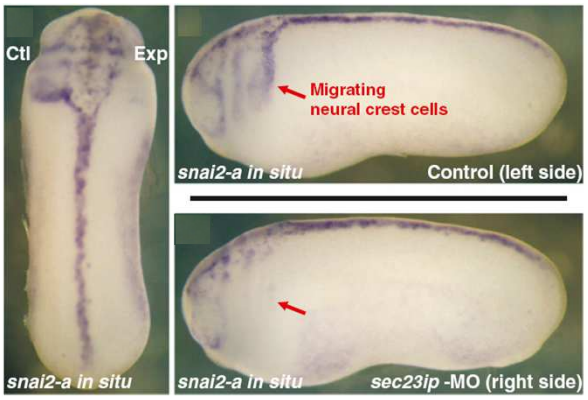
Association between white blue-eyed cats and deafness (noted by Darwin in 1859)

White forelock and deafness/bowel bowel blockage in foals

& many more...

Inactivating SEC23IP—predicted from *Arabidopsis*—in a tadpole disrupts neural crest cells, consistent with Waardenburg syndrome

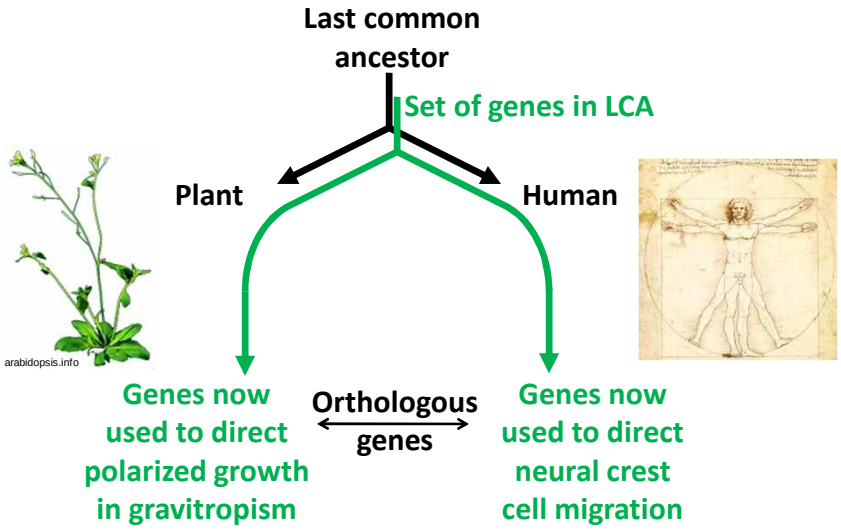




SEC23IP localizes to the neural crest cells & induces neural crest defects upon knockdown

McGary, Park *et al*. *PNAS* 107:6544-9 (2010)

Phenologs identify evolutionarily conserved systems of proteins relevant to particular traits/diseases.



McGary, Park *et al*. *PNAS* (2010)

Summary of a plant model of Waardenburg syndrome

Waardenburg syndrome genes…

…suggest a relevant plant system. Plant genes…

…suggest new WS genes, confirmed in frogs, validating the plant model.

?

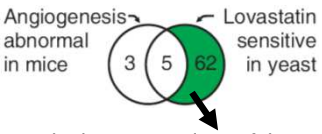
?

PREDICT AND VALIDATE
new gravitropism genes

SEARCH FOR MUTATIONS
in humans

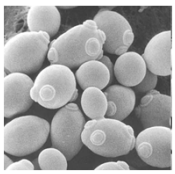
McGary, Park *et al*. *PNAS* (2010)

Example #3: Yeast genes linked to statin drug sensitivity predict mammalian blood vessel defects

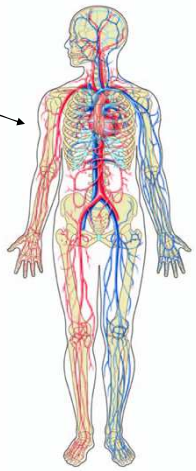


The human versions of these yeast genes are <u>candidate angiogenesis genes</u>

Can these really tell us about these?



www.chemistryland.com

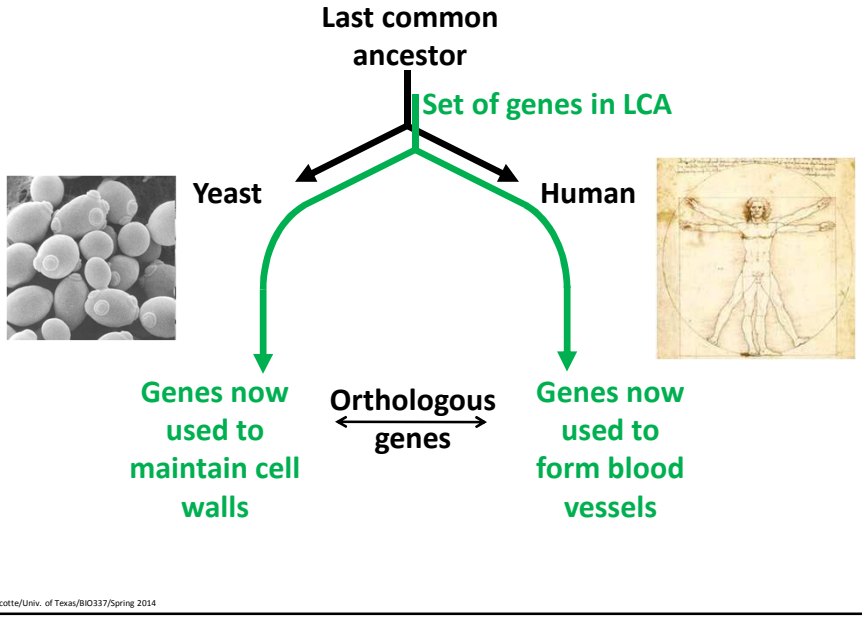


Dorling Kindersley

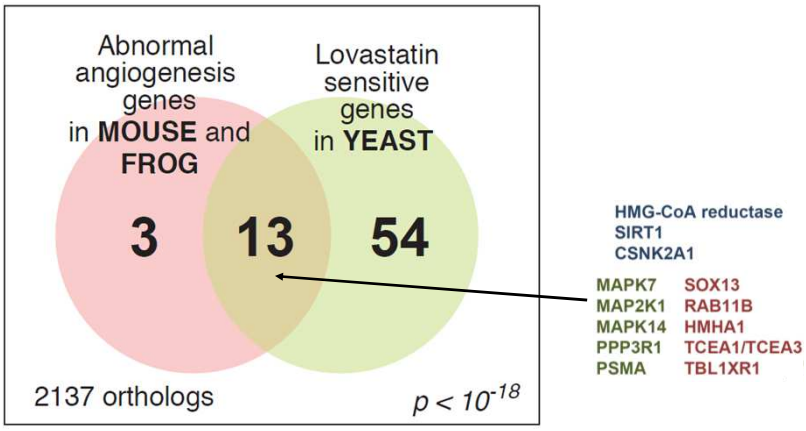
McGary, Park *et al*. *PNAS* (2010)

Disrupting the SOX13 gene causes strong blood vessel defects

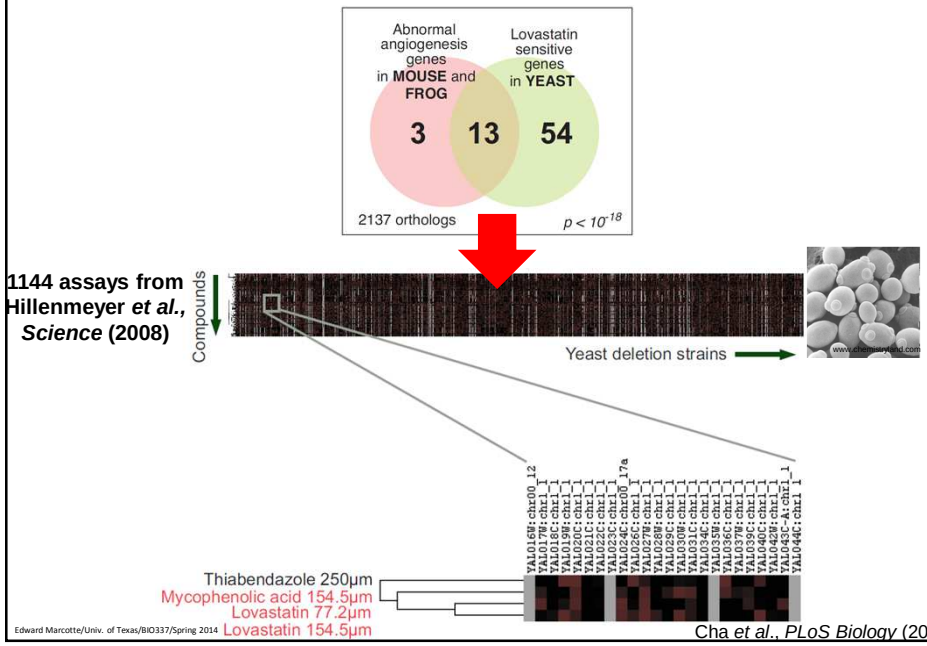
McGary, Park *et al*. *PNAS* (2010)



The yeast/angiogenesis gene module



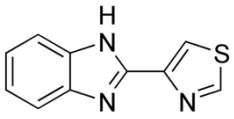
Chemicals that interact genetically with this module are candidate angiogenesis inhibitors



Cha *et al.*, *PLoS Biology* (2012)

Screening for drugs that interact genetically with this yeast module led us to identify a new angiogenesis inhibitor

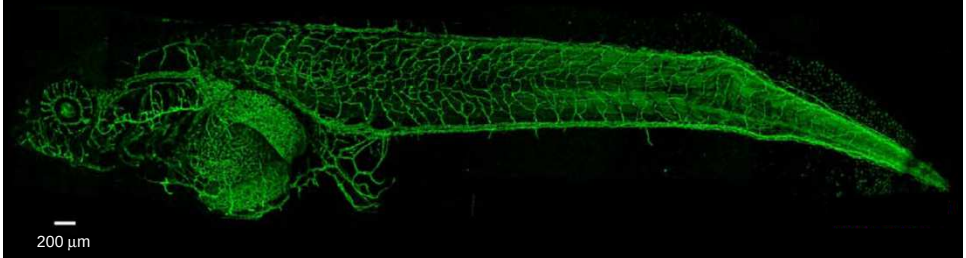
TBZ = thiabendazole
An FDA-approved antifungal drug with 40 years of safety data



- **Approved by the U.S. Food and Drug Administration in 1967**

- fungicide and parasiticide
- No mutagenic or carcinogenic effects
  2 year safety trials in animals
- Off-patent, now marketed as a generic drug

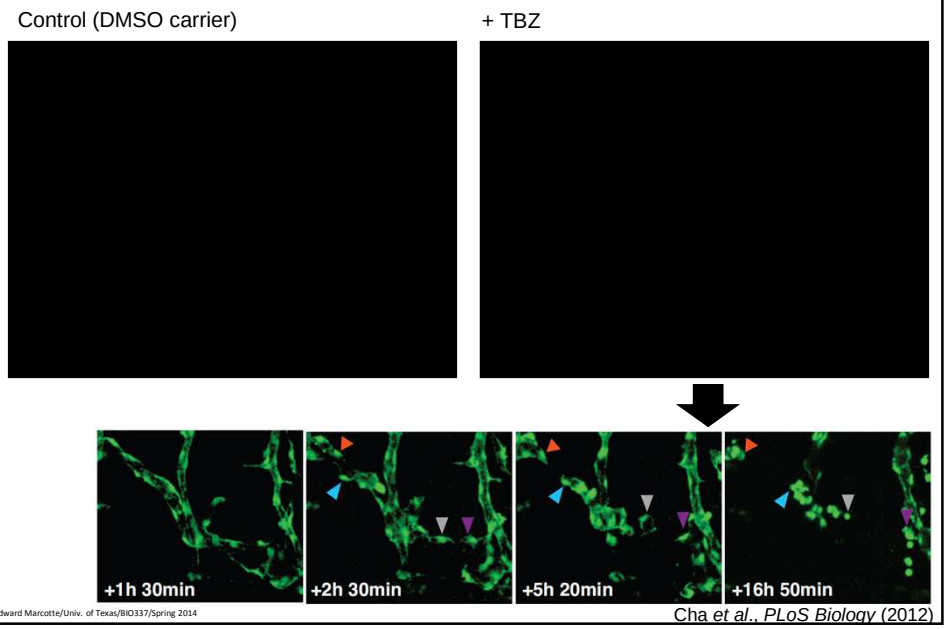
Imaging the blood vessels of a living, transgenic tadpole in a dish of water



kdr:GFP transgenic *Xenopus laevis*

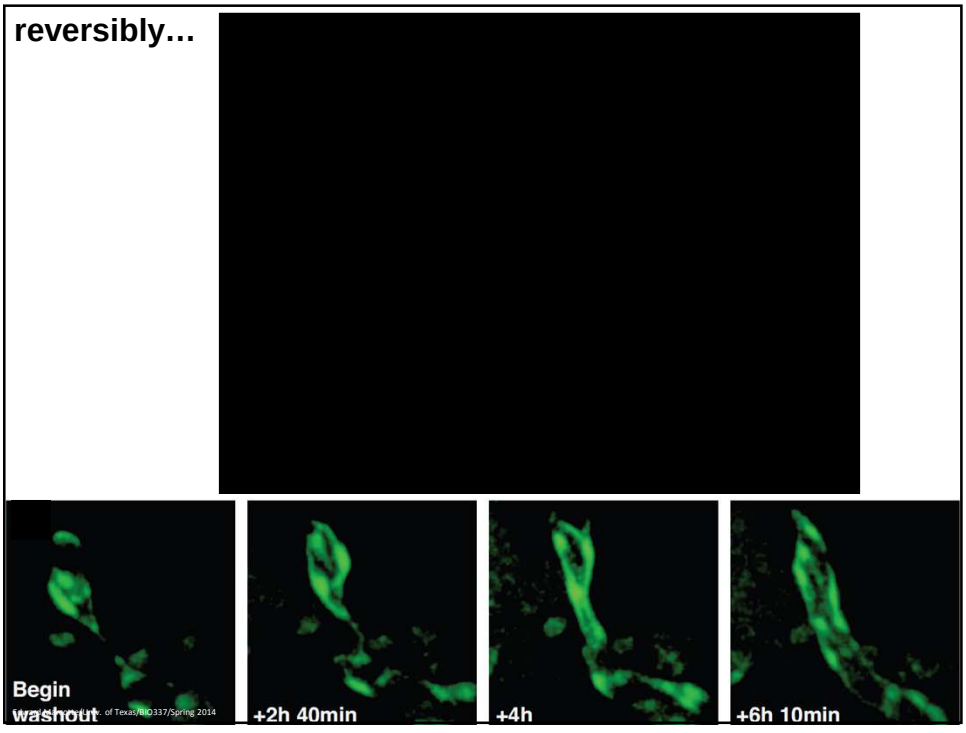
Image: Hye Ji Cha

Thiabendazole disrupts vascular integrity, causing retraction and rounding of vascular endothelial cells

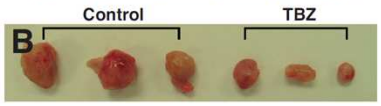


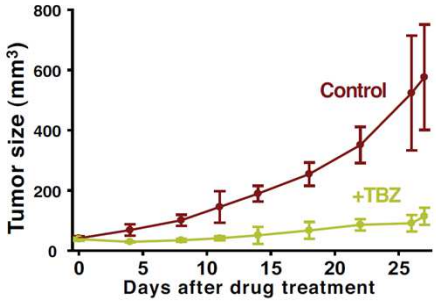
Cha *et al*., *PLoS Biology* (2012)

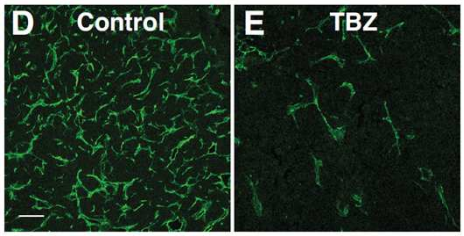
reversibly…



TBZ slows the growth of human fibrosarcoma tumors transplanted into immune-compromised mice





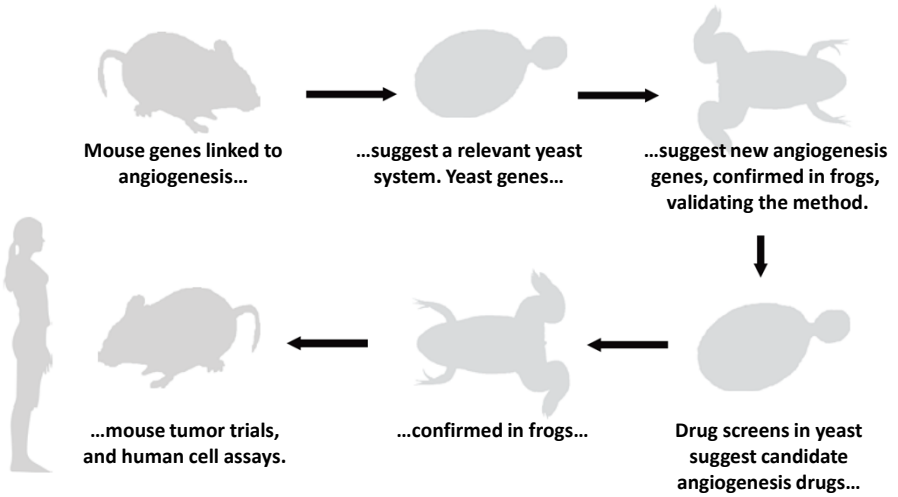


Vasculature in tumor sections

Edward Marcotte/Univ. of Texas/BIO337/Spring 2014

Cha *et al*., *PLoS Biology* (2012)

Summarizing the "road map" to a new vascular disrupting agent



Edward Marcotte/Univ. of Texas/BIO337/Spring 2014

Cha *et al*., *PLoS Biology* (2012)

Summary of the major themes

- Genetic traits and diseases often arise from perturbing any one (or more) of a set or module of genes, e.g. components of the same pathway or protein complex
- Pathways and complexes can be deeply evolutionarily conserved, often more deeply than the diseases or traits they are linked to
- Knowing the underlying module of genes thus predicts new candidate genes for any of the linked traits across organisms, e.g. as for yeast lovastatin sensitivity predicting vertebrate angiogenesis genes