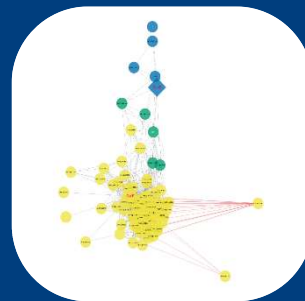
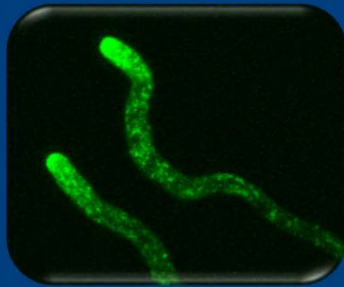


Harnessing transcriptomic data to predict the function of proteins in the microbial cell factory *Aspergillus niger*



Paul Schäpe, Timothy C Cairns, Vera Meyer

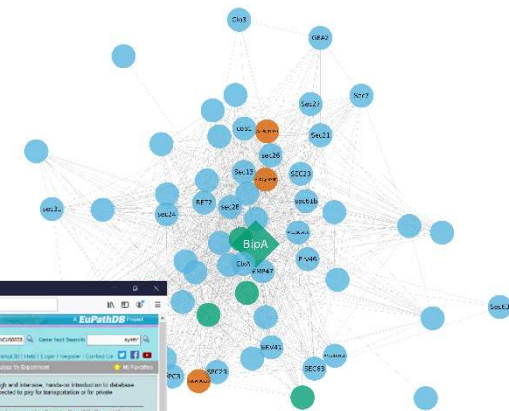
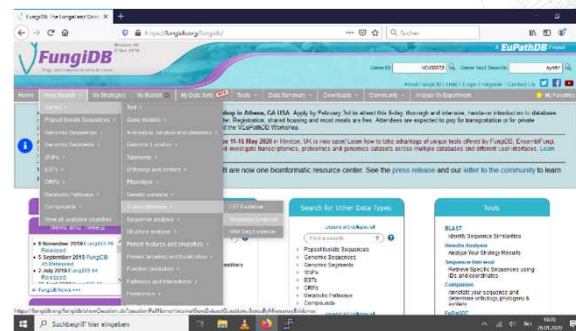
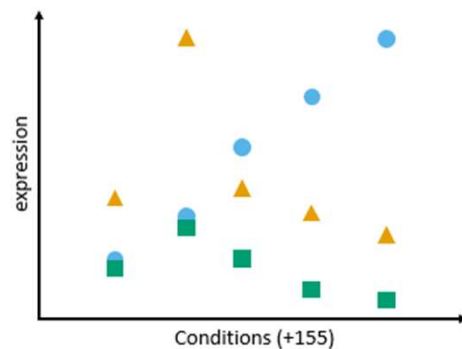
19.02.2020

Technische Universität Berlin
Chair of Applied and Molecular Microbiology

Harnessing transcriptomic data to predict the function of proteins in the microbial cell factory *Aspergillus niger*

Introduction

- How to correlate genes based on their expression
- What insights can be gained from these analysis
- How to successfully apply it
- Where to find the coexpression-network

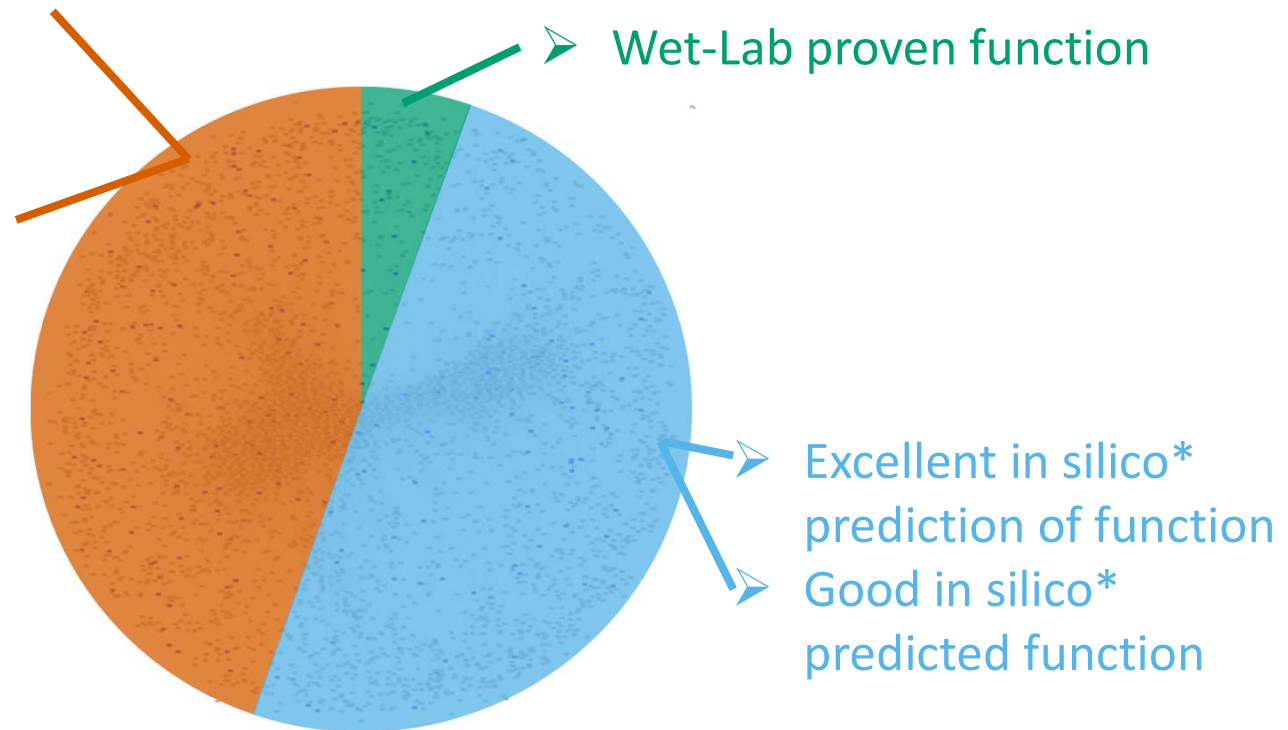


Why new annotation method?

➤ All predicted ORFs of *A. niger*

➤ Weak in silico*
predicted function

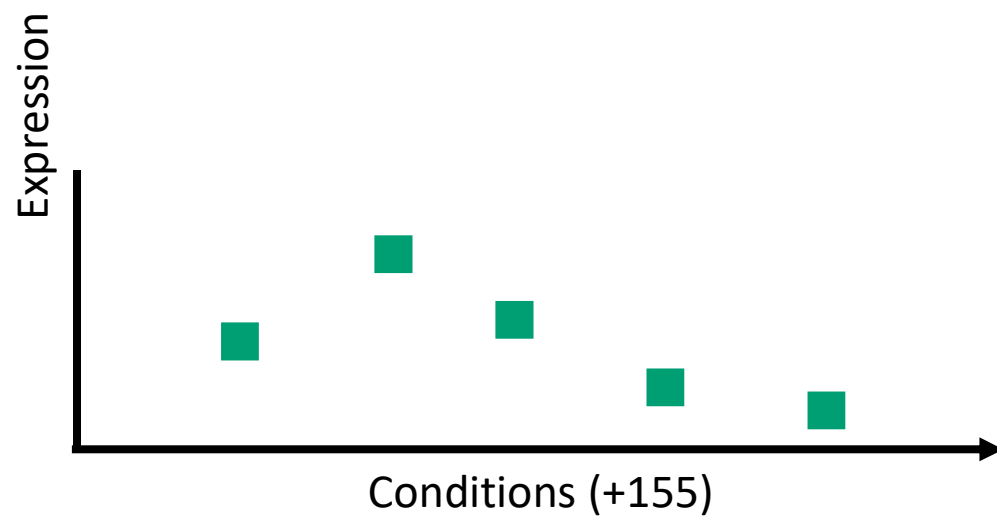
➤ Hypothetical function



*based on sequence homology

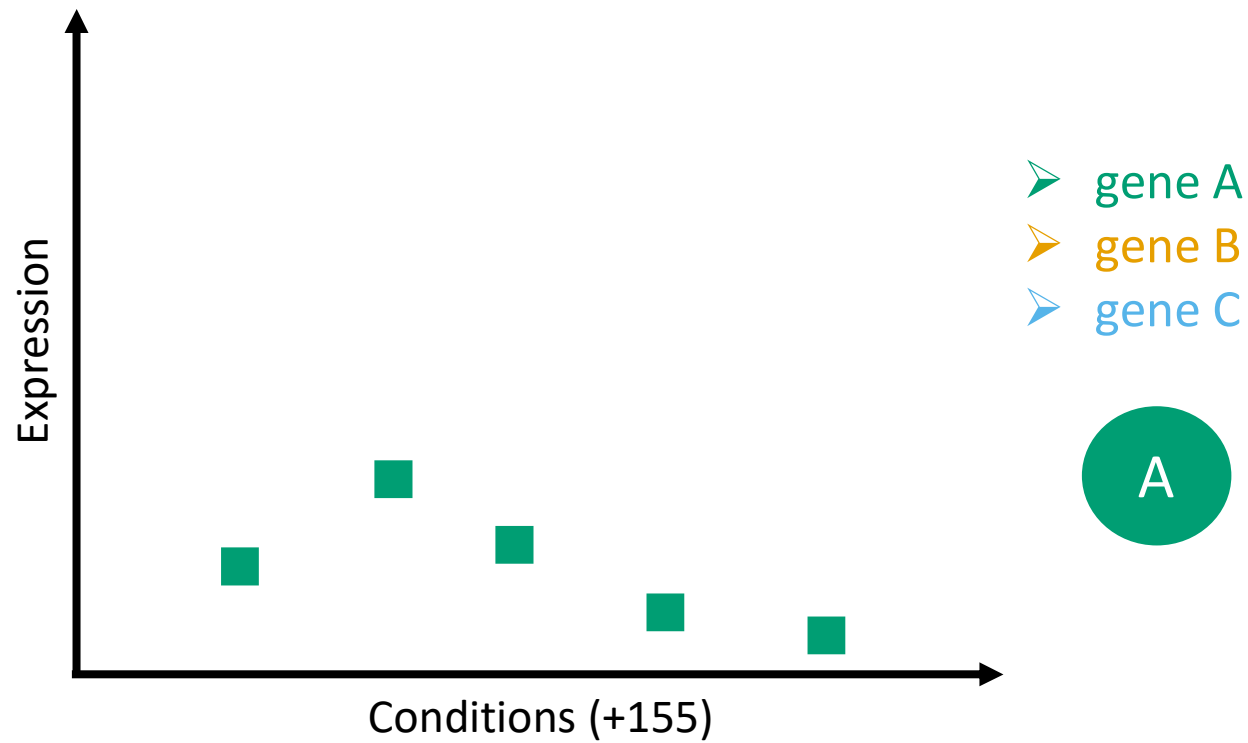
Expression Correlation

- Genes with comparable expression pattern are considered to have similar biological function
- Over 155 publicly available expression-profiles (microarray)
- Expression profile for every gene combination was investigated



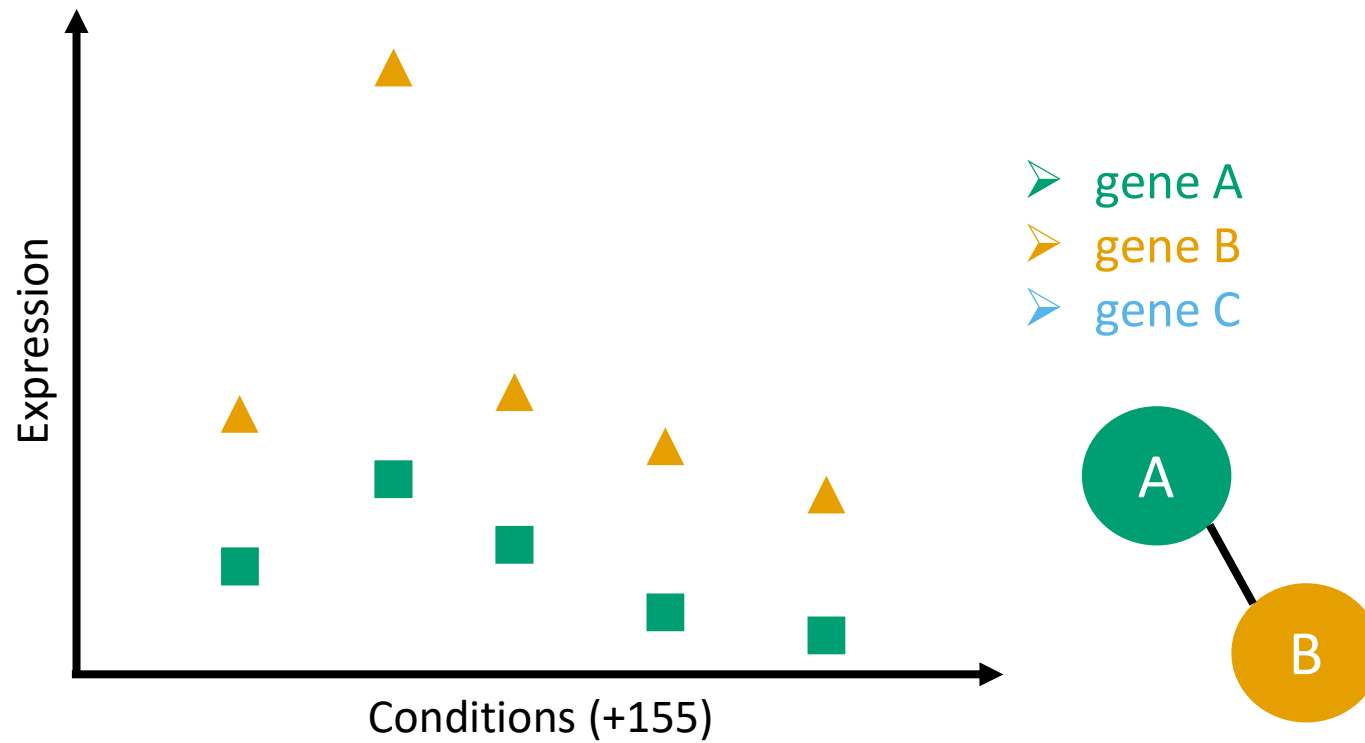
Expression Correlation

- Genes with comparable expression pattern are considered to have similar biological function



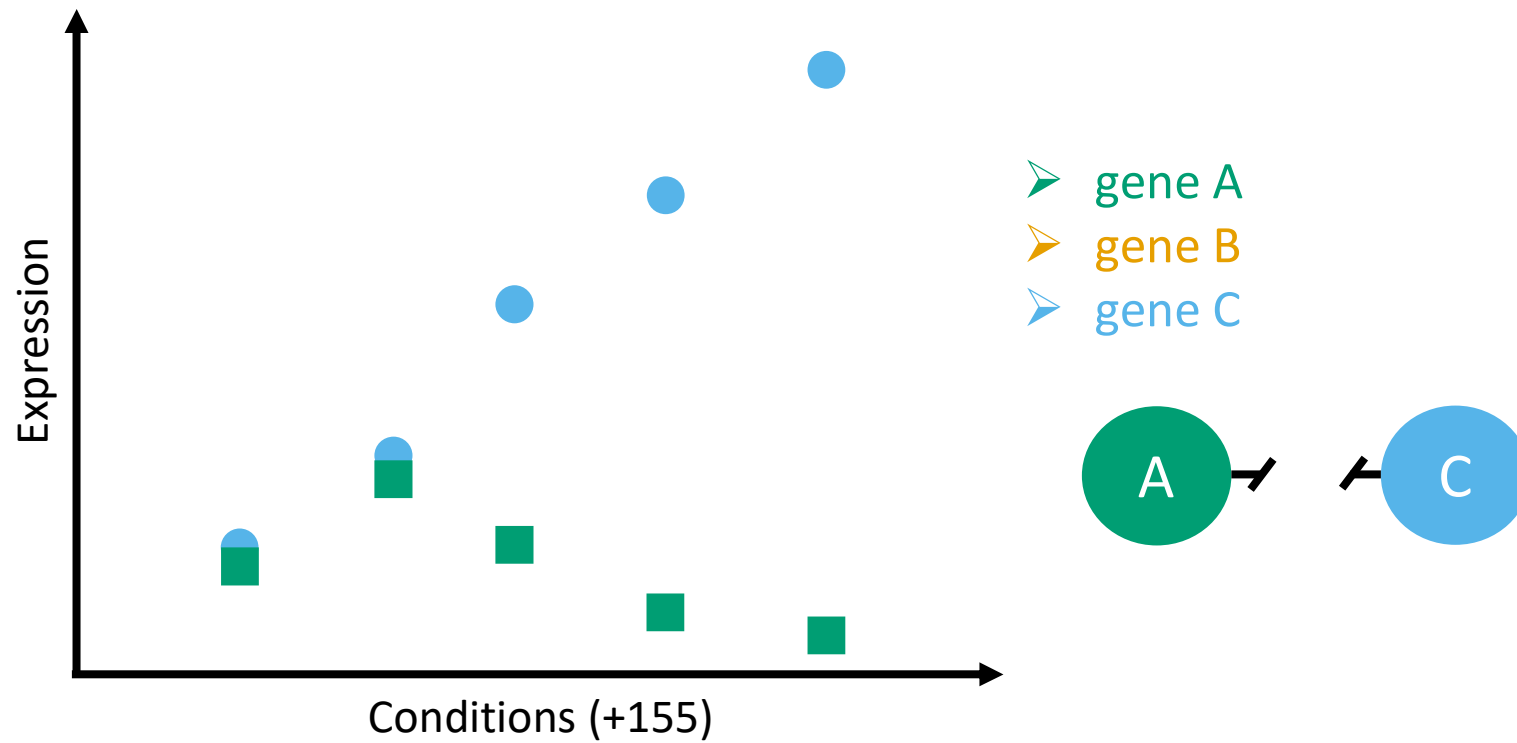
Expression Correlation

- Genes with comparable expression pattern are considered to have similar biological function



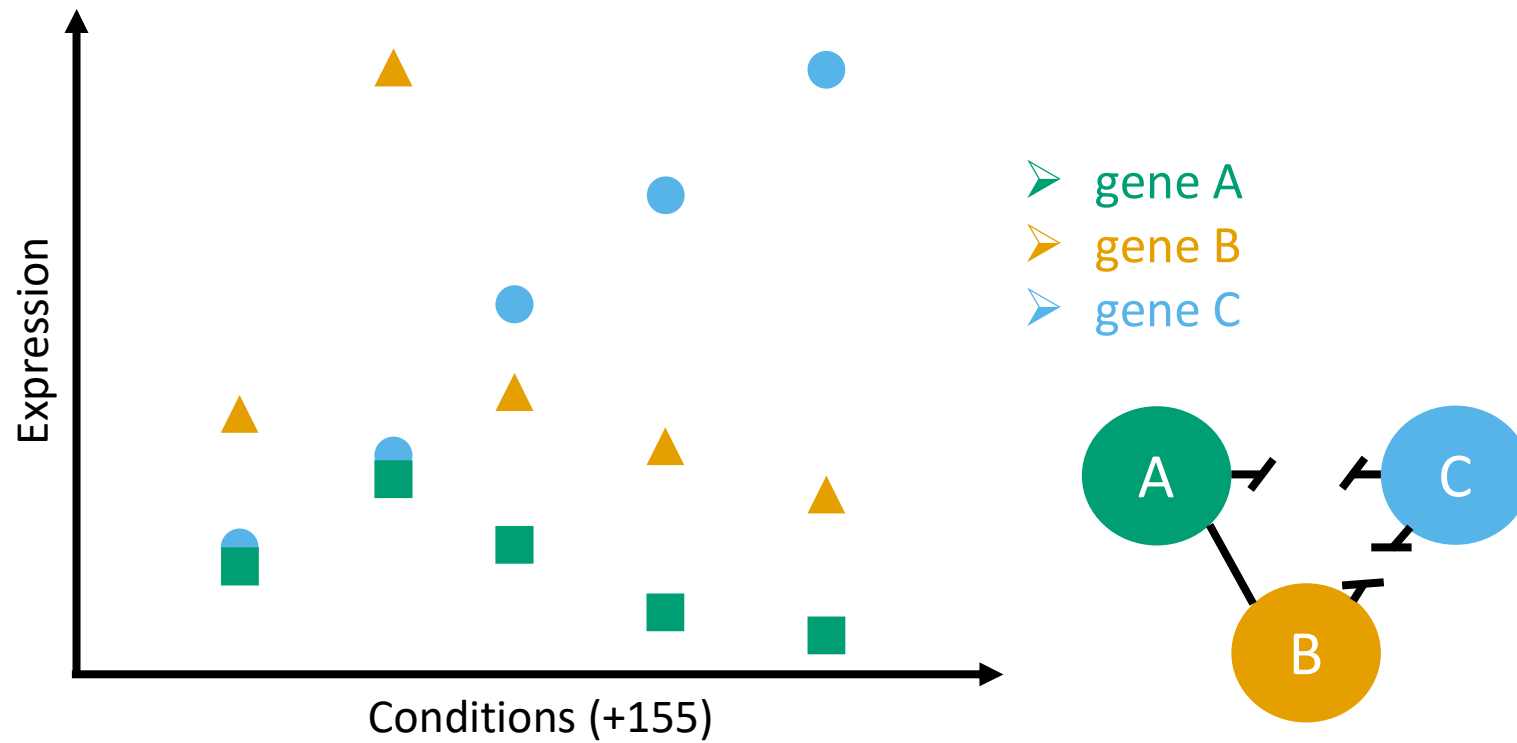
Expression Correlation

- Genes with comparable expression pattern are considered to have similar biological function



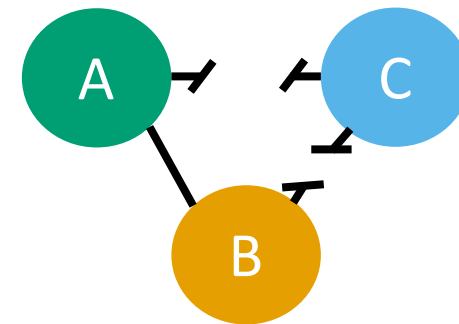
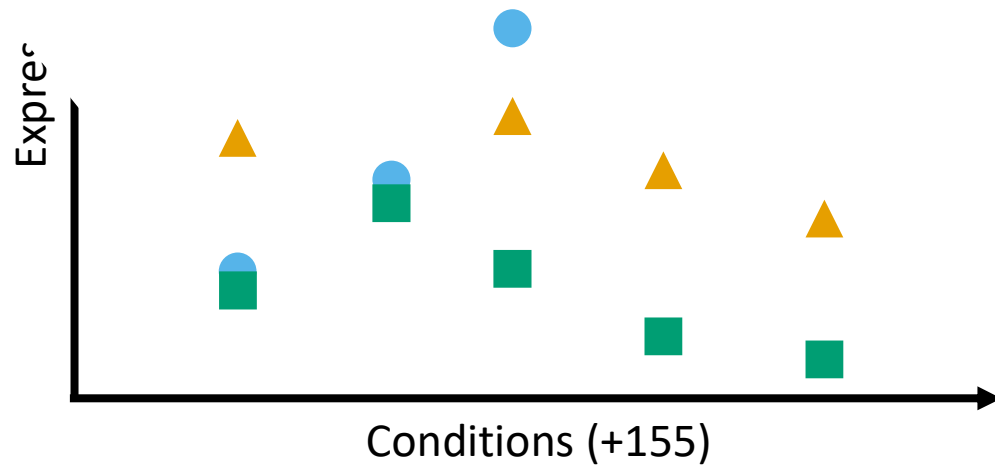
Expression Correlation

- Genes with comparable expression pattern are considered to have similar biological function



Expression Correlation

- Genes with comparable expression pattern are considered to have similar biological function
- Over 155 publicly available expression profiles (microarray)
- Expression profile for every gene combination was investigated
- Spearman-Coefficient with cutoff of > 0.5 was used



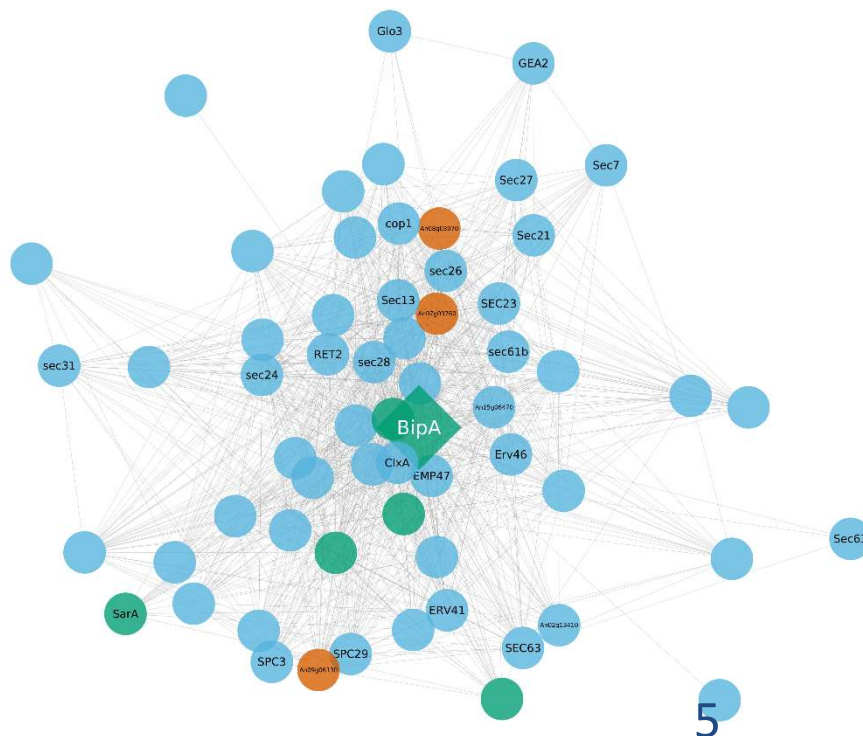
Functionally annotate a hypothetical gene

- 9,597 of the 14,003 predicted ORFs are part of the network
- Roughly 4,000,000 correlations calculated
- Too much information to consider everything for every research question
- Subnetworks were created



Functionally annotate a hypothetical gene

- Creating Subnetworks for further analysis
- Gene Ontology for fast interpretation
- The hypothetical gene likely plays a similar role as the known ones
- **Guilt By Association**

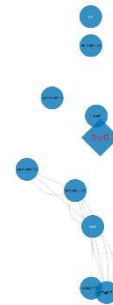


- Wet-Lab proven function
- Excellent in silico* prediction of function
- Good in silico* predicted function
- Weak in silico* predicted function
- Hypothetical function

*based on sequence homology

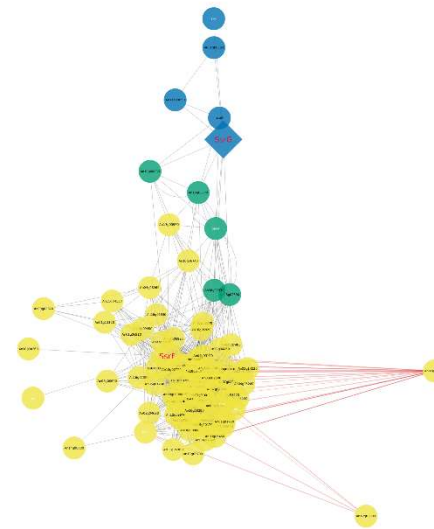
Understanding Biology

- Combining Subnetworks to investigate interactions



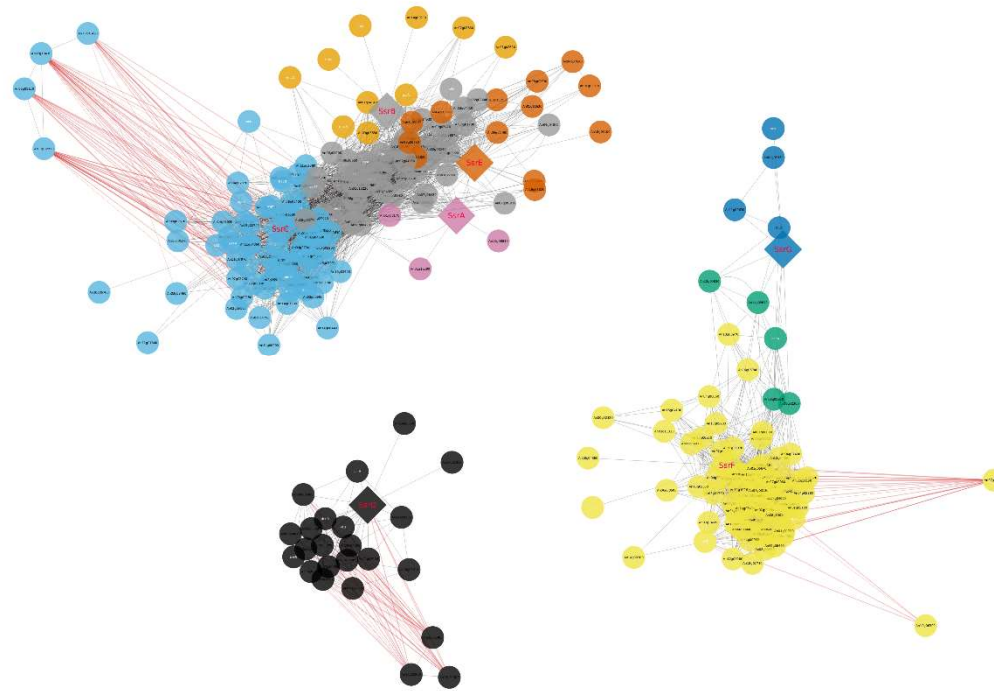
Understanding Biology

- Combining Subnetworks to investigate interactions
- Find genes that have a more central role



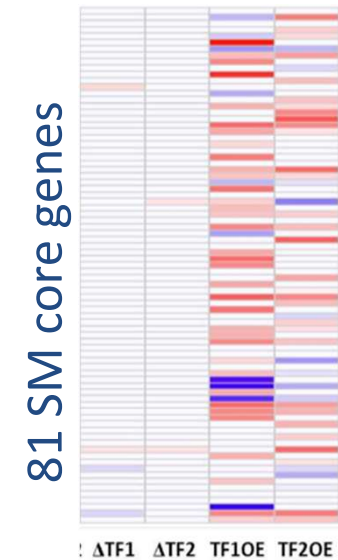
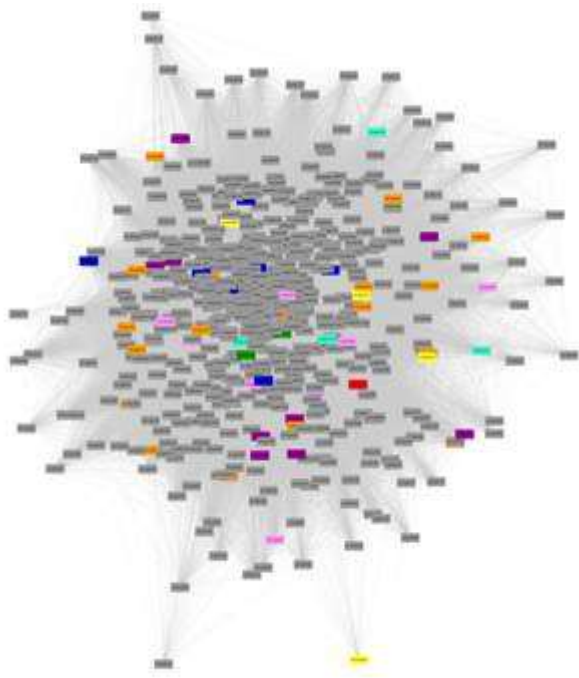
Understanding Biology

- Combining Subnetworks to investigate interactions
- Find genes that have a more central role
- Find independent parts of a biological process
- Find highly linked parts of a biological process



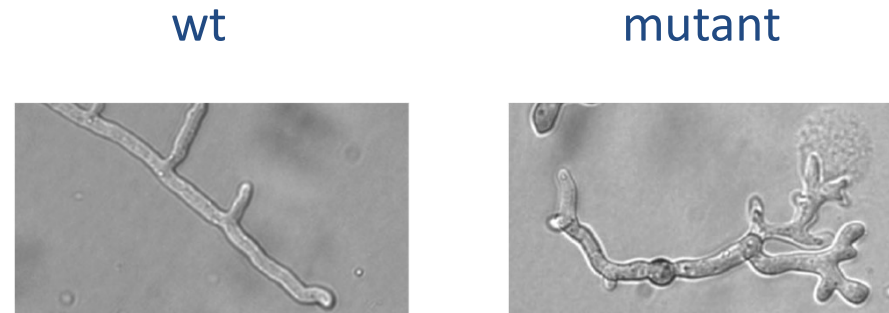
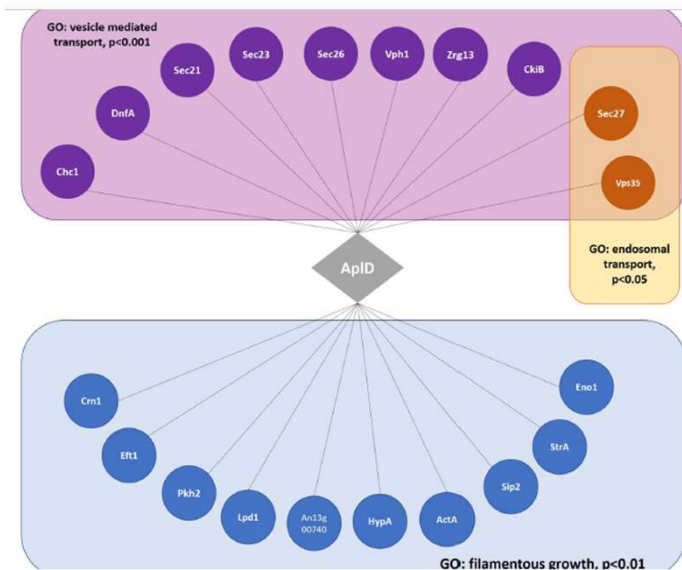
Understanding Biology: SM

- Schäpe, Kwon *et al.* 2019
- Investigate regulation of secondary metabolites
- Apply coexpression-network on predicted transcription factors
- Two novel transcription factors regulating 43/45 SM core genes were found



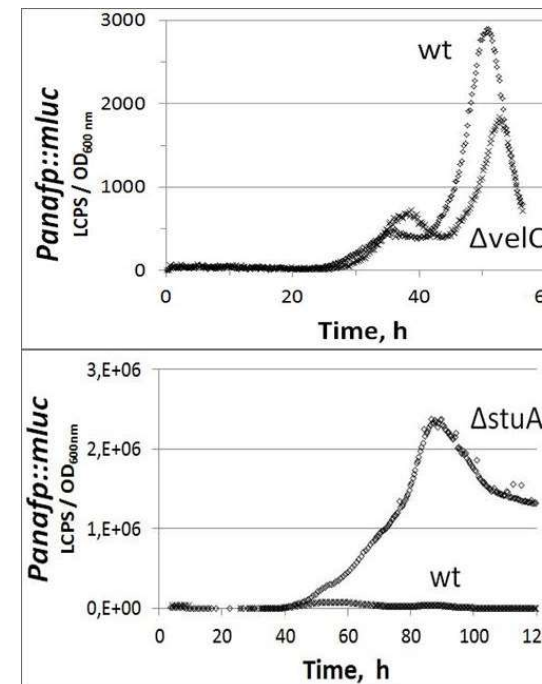
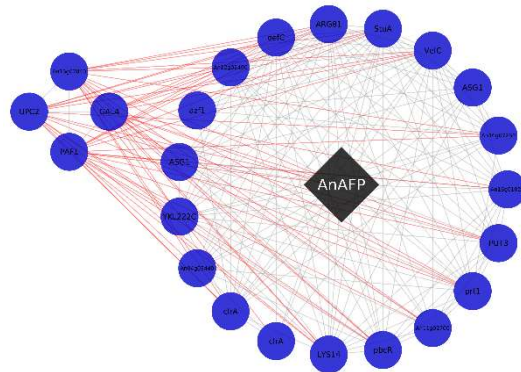
Understanding Biology: Morphology

- Cairns, Feurstein *et al.* 2019
- Find a target to gene playing a role in morphology (*aplD* An01g02600)
- Combining coexpression data with gene ontology analysis
- Mutants showed: swollen tips, higher branching rate, longer hyphae, more



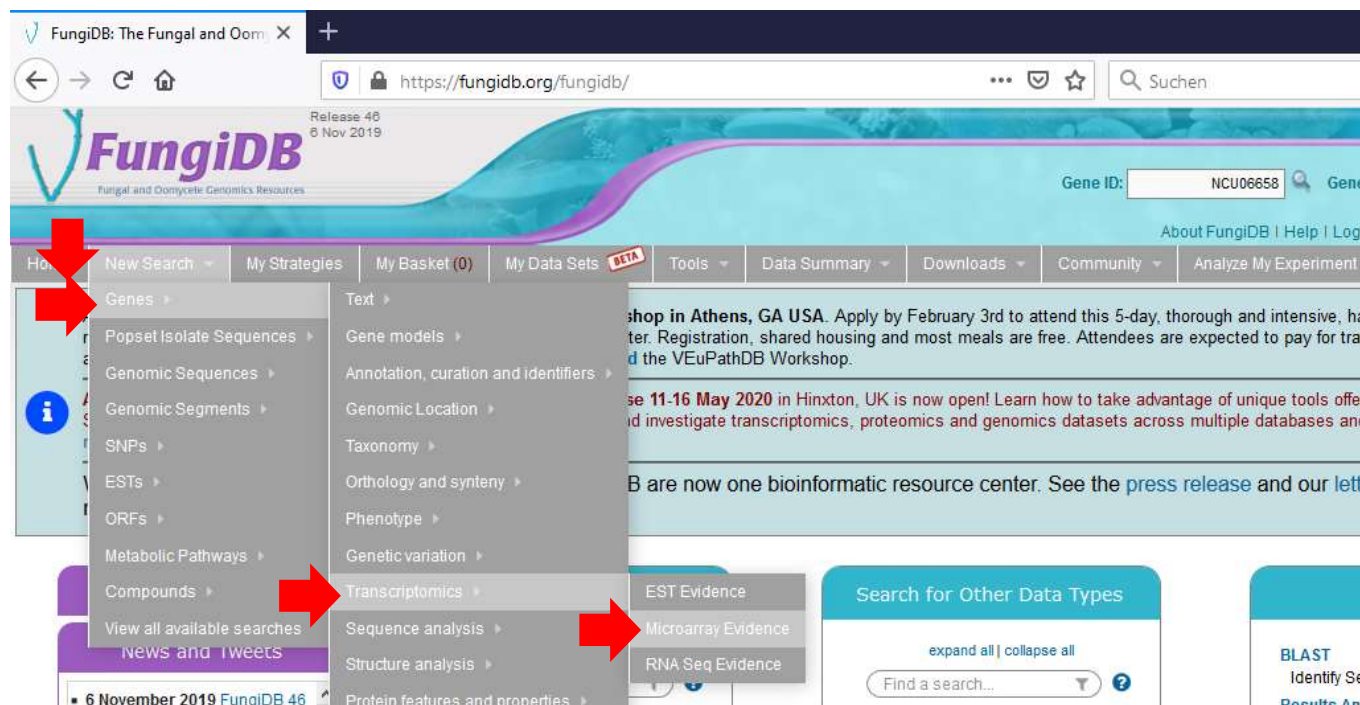
Understanding Biology: Antifungals

- Paeye *et al.* 2016
- Investigating the mode of action of an antifungal peptide (AnAFP)
- While analyzing expression correlation, links to secondary metabolism regulation and starvation appeared
- Further analysis of an internal role of AnAFP is ongoing



FungiDB

- Resource was made to be used by anyone interested
- Publicly available at FungiDB
- “New Search -> Genes -> Transcriptomics -> Microarray Evidence”



FungiDB

- “New Search -> Genes -> Transcriptomics -> Microarray Evidence”
- “*Aspergillus Niger* co-expression network (Vera Meyer)”

Identify Genes based on Microarray Evidence

Filter Data Sets:

Legend: **C** Coexpression **DC** Direct Comparison **FC** Fold Change **P** Percentile

Organism	Data Set	Choose a search	
<i>A. fumigatus</i> Af293	Response to hypoxia (Barker et al.)	FC	P
<i>A. niger</i> CBS 513.88	<i>Aspergillus niger</i> gene co-expression network (Vera Meyer)	C	
<i>C. albicans</i> SC5314	Antifungal Benzimidazole Derivative Response (Steffen Rupp)	DC	P
<i>C. cinerea</i> okayama7#130	Microarray analysis of fruiting body development in <i>Coprinopsis cinerea</i> (Anderson et al.)	FC	P
<i>C. gattii</i> WM276	Comparison of gene expression in <i>Cryptococcus gattii</i> clinical isolates (Ma et al.)	FC	
<i>C. immitis</i> RS	Gene expression changes during mycelia differentiation in <i>Coccidioides immitis</i> (Viriyakosol et al.)	FC	P
<i>C. neoformans</i> var. <i>grubii</i> H99	Flucytosine Susceptibility (Yong-Sun Bahn)	FC	P
<i>C. tropicalis</i> MYA-3404	Transcriptome of <i>C. tropicalis</i> cells under in vitro biofilm-forming conditions (Jones et al.)	DC	P
<i>F. graminearum</i> PH-1	Transcriptome of <i>Fusarium graminearum</i> during the infection of wheat (Lys?e et al.)	FC	P
<i>F. graminearum</i> PH-1	<i>Fusarium graminearum</i> transcriptome during symptomless and symptomatic wheat infection (Brown et al.)	FC	P

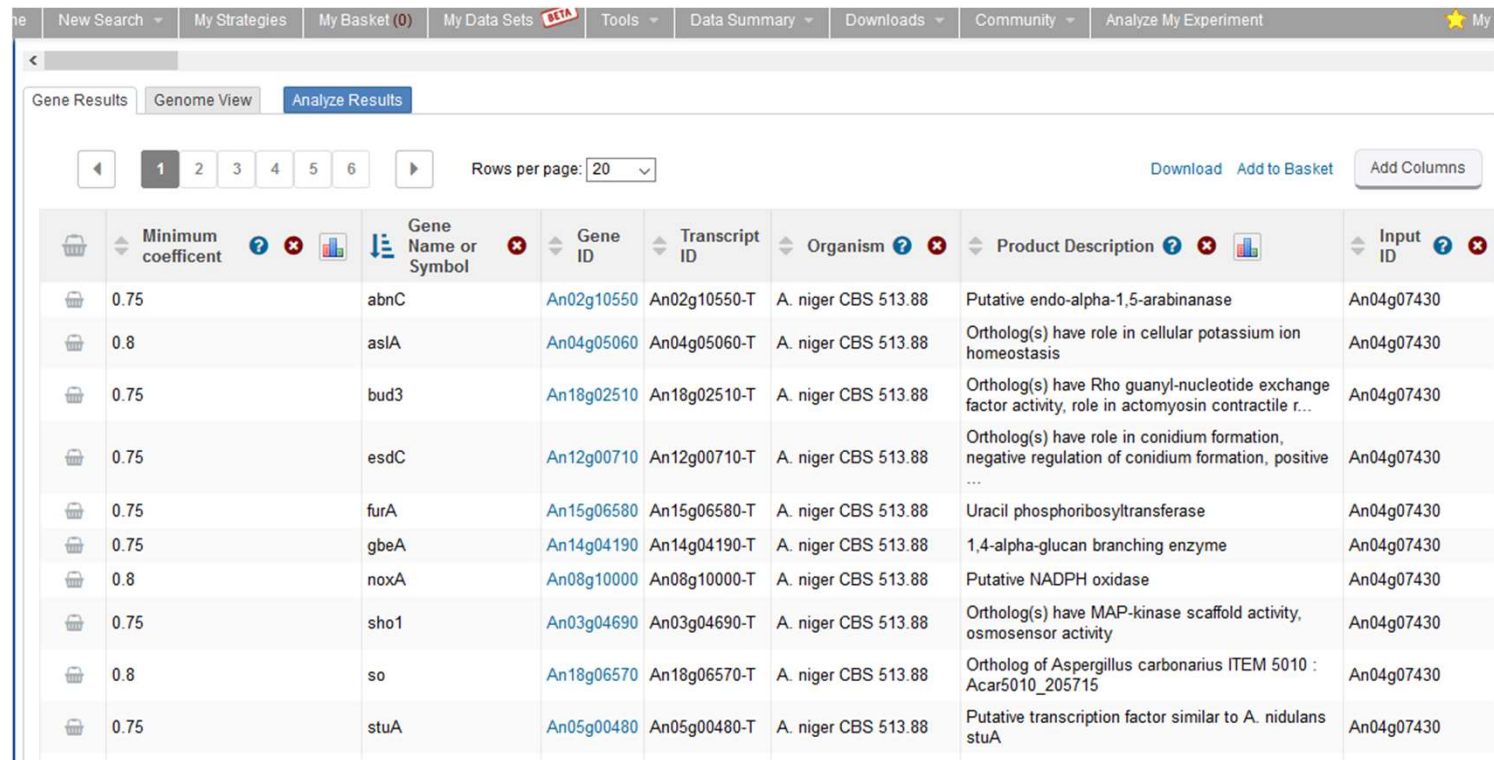
FungiDB

- “*Aspergillus Niger* co-expression network (Vera Meyer)”
- Enter your gene of interest
- Or the *Aspergillus Niger* orthologue of your gene of interest

The screenshot shows the FungiDB search interface. At the top, there is a navigation bar with links: New Search, My Strategies, My Basket (0), My Data Sets (with a 'BETA' badge), Tools, Data Summary, Downloads, Community, and Analyze My Experiment. Below the navigation bar is the 'Gene ID input set' section, which has three radio button options: 'Enter a list of IDs or text' (selected), 'Upload a text file', and 'Copy from My Basket'. The 'Enter a list of IDs or text' option has a text input field containing 'An04g07430'. A red arrow points to this input field. Below this are two more radio button options: 'Copy from My Strategy' and 'Copy from My Strategy'. The 'Correlation' section has a dropdown menu set to 'Positive Correlation'. The 'Spearman coefficient (greater or equal to)' section has a text input field containing '0.75'. At the bottom right, there is a 'Get Answer' button with a red arrow pointing to it. Below the 'Get Answer' button is a small text input field with a question mark icon and the text 'Give this search a name (optional)'.

FungiDB

- A list of all correlating genes will become available for download
- Or further analysis within FungiDB

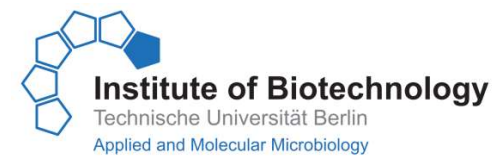


The screenshot shows the FungiDB Gene Results page. The table displays the following columns: Minimum coefficient, Gene Name or Symbol, Gene ID, Transcript ID, Organism, Product Description, and Input ID. The data is as follows:

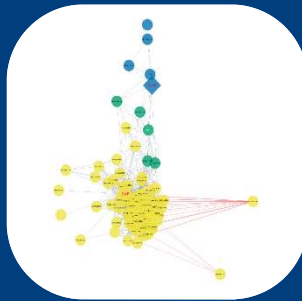
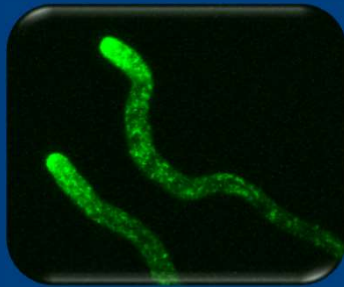
Minimum coefficient	Gene Name or Symbol	Gene ID	Transcript ID	Organism	Product Description	Input ID
0.75	abnC	An02g10550	An02g10550-T	A. niger CBS 513.88	Putative endo-alpha-1,5-arabinanase	An04g07430
0.8	aslA	An04g05060	An04g05060-T	A. niger CBS 513.88	Ortholog(s) have role in cellular potassium ion homeostasis	An04g07430
0.75	bud3	An18g02510	An18g02510-T	A. niger CBS 513.88	Ortholog(s) have Rho guanyl-nucleotide exchange factor activity, role in actomyosin contractile r...	An04g07430
0.75	esdC	An12g00710	An12g00710-T	A. niger CBS 513.88	Ortholog(s) have role in conidium formation, negative regulation of conidium formation, positive ...	An04g07430
0.75	furA	An15g06580	An15g06580-T	A. niger CBS 513.88	Uracil phosphoribosyltransferase	An04g07430
0.75	gbeA	An14g04190	An14g04190-T	A. niger CBS 513.88	1,4-alpha-glucan branching enzyme	An04g07430
0.8	noxA	An08g10000	An08g10000-T	A. niger CBS 513.88	Putative NADPH oxidase	An04g07430
0.75	sho1	An03g04690	An03g04690-T	A. niger CBS 513.88	Ortholog(s) have MAP-kinase scaffold activity, osmosensor activity	An04g07430
0.8	so	An18g06570	An18g06570-T	A. niger CBS 513.88	Ortholog of Aspergillus carbonarius ITEM 5010 : Acar5010_205715	An04g07430
0.75	stuA	An05g00480	An05g00480-T	A. niger CBS 513.88	Putative transcription factor similar to A. nidulans stuA	An04g07430

The Meyer Lab

- Technische Universität Berlin – Chair of Applied and Molecular Microbiology
- Lab of Vera Meyer



Harnessing transcriptomic data to predict the function of proteins in the microbial cell factory *Aspergillus niger*



Paul Schäpe, Timothy C Cairns, Vera Meyer

19.02.2020

Technische Universität Berlin
Chair of Applied and Molecular Microbiology

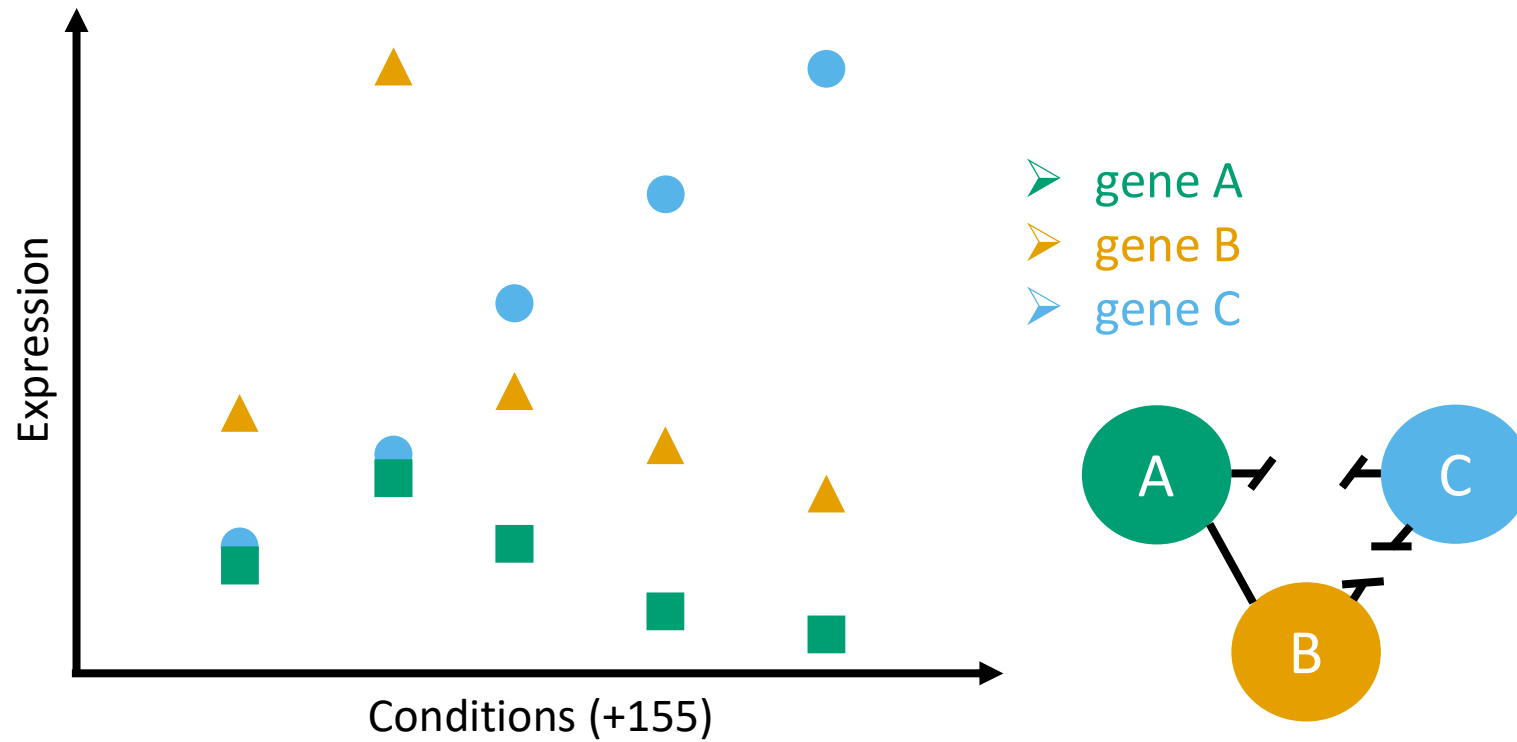






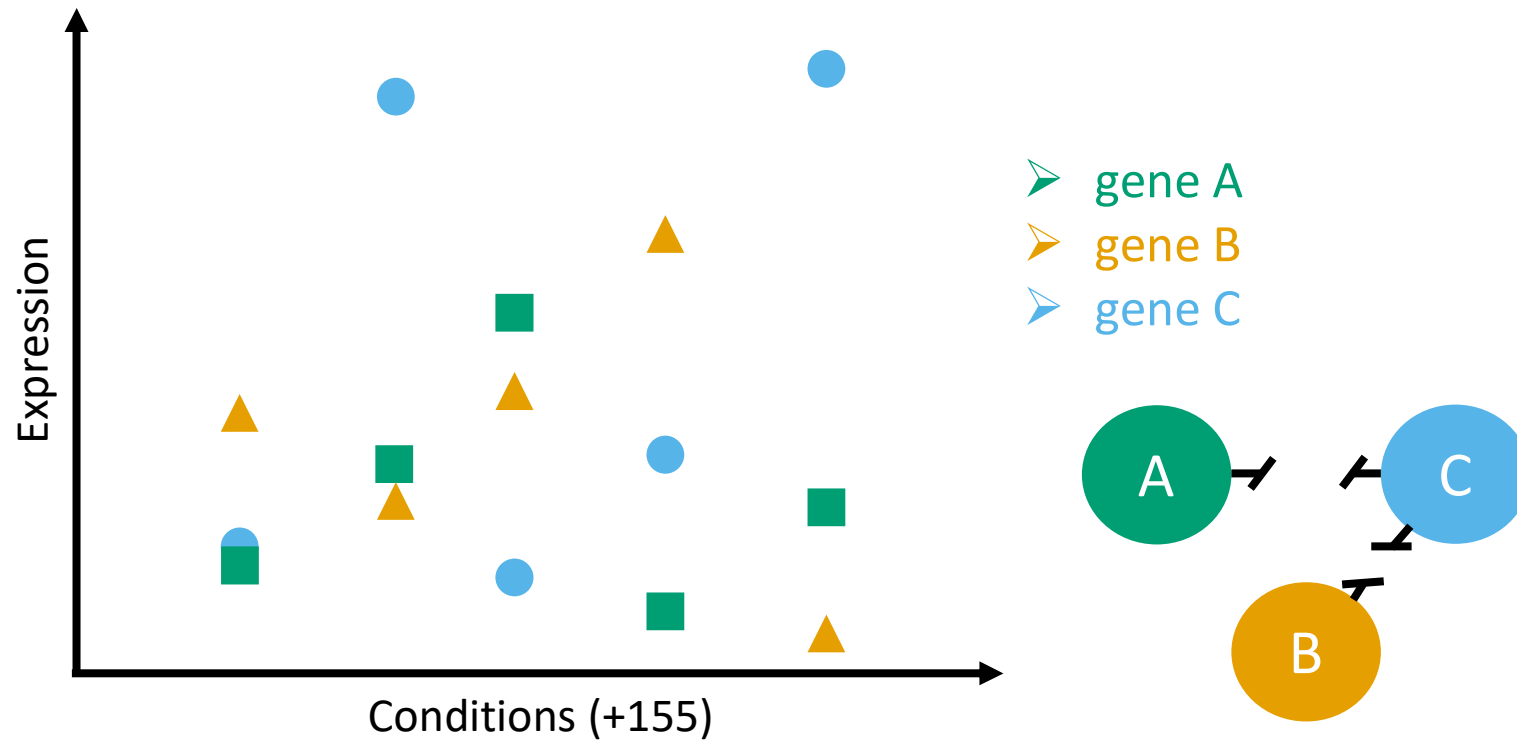
Expression Correlation

➤ Cut-off determination



Expression Correlation

➤ Cut-off determination





Understanding Biology: SM

- Investigate regulation of secondary metabolites
- Apply coexpression-network on predicted transcription factors
- Two novel transcription factors regulating secondary metabolism were found

