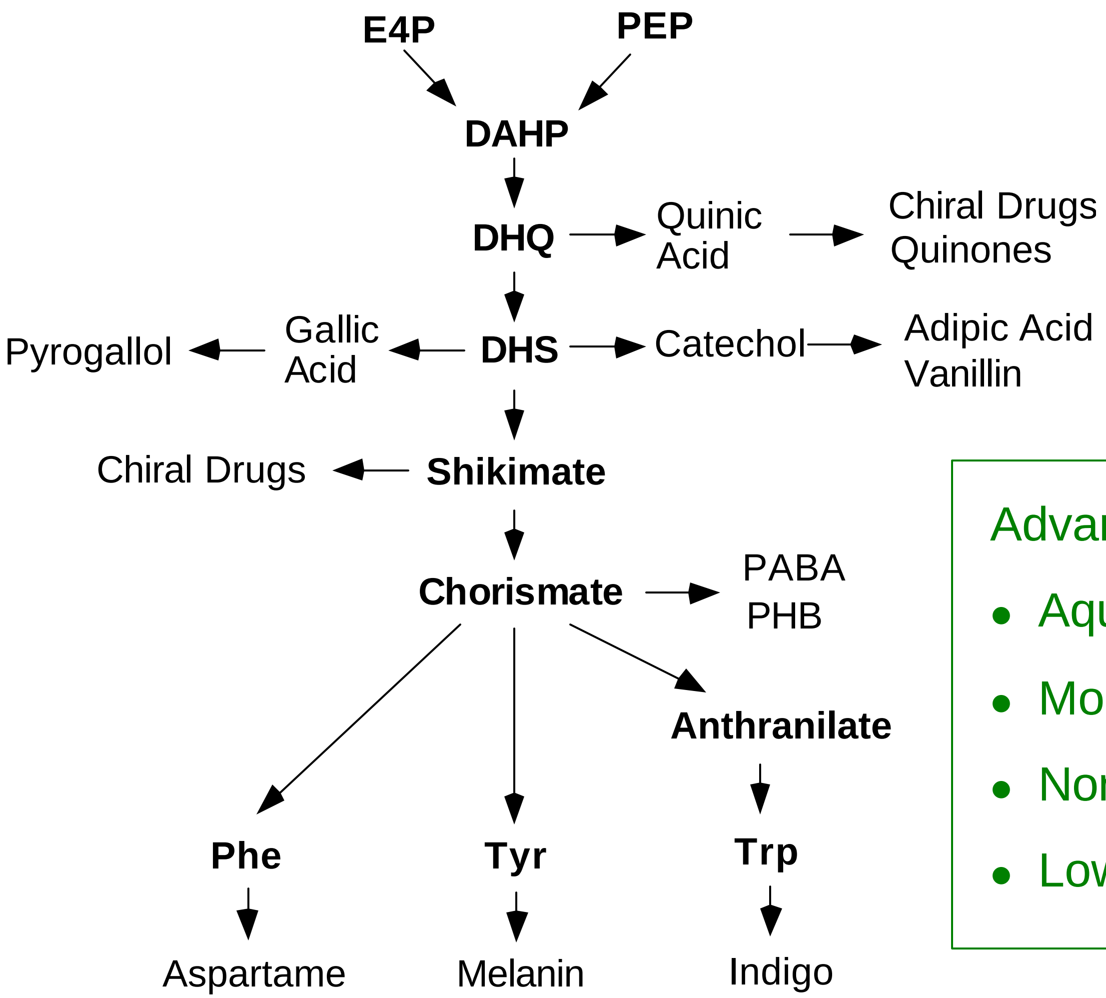
Aromatic Amino Acid Biosynthesis in *Archaeoglobus fulgidus*

H.G. Monbouquette & J.C. Liao, Chemical Engineering, **I. Schröder**,
Microbiology & Molecular Genetics, UCLA, Los Angeles, CA 90095

- Discover novel enzymes for metabolic engineering of aromatic synthesis pathways
- Conduct first analysis of aromatic biosynthesis in Archaea
- Clone, express, and characterize the thermostable aromatic biosynthesis enzymes of *A. fulgidus*
- Implement LC/MS and DNA microarrays as complementary tools for functional proteomics

Aromatic Biosynthesis Pathways Lead to Industrial Products



E4P **PEP**

DAHP

DHQ → Quinic Acid → Chiral Drugs, Quinones

Pyrogallol ← Gallic Acid ← **DHS** → Catechol → Adipic Acid, Vanillin

Chiral Drugs ← **Shikimate**

Chorismate → PABA, PHB

Anthranilate

Phe **Tyr** **Trp**

Aspartame Melanin Indigo

Advantages

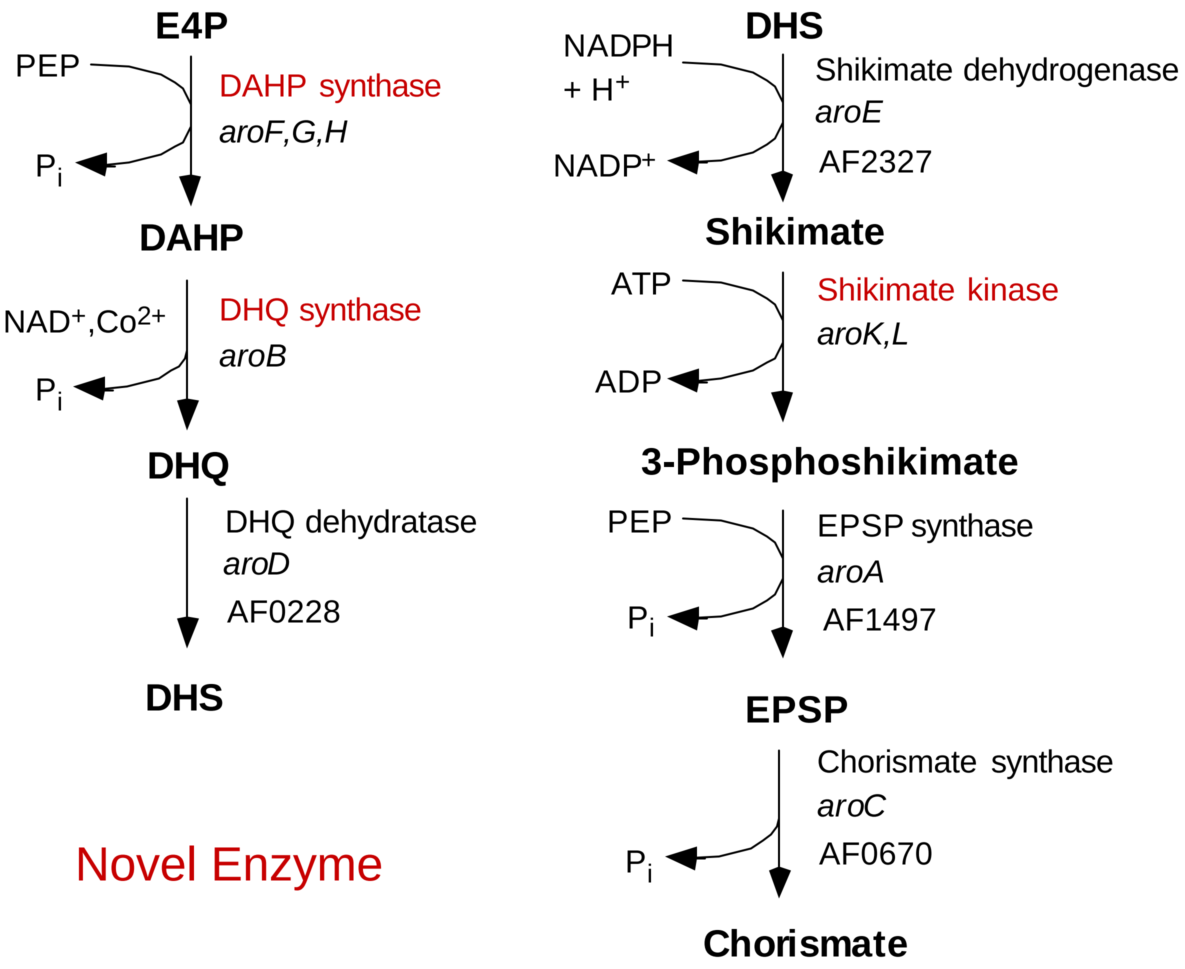
- Aqueous solvent
- Moderate T, p
- Non-toxic intermediates
- Lower raw material cost

Not all metabolites (bold) are shown. **E4P**, D-erythrose-4-phosphate; **PEP**, phosphoenolpyruvate; **DAHP**, 3-deoxy-D-*arabino*-heptulosonate-7-phosphate; **DHQ**, 3-dehydroquinate; **DHS**, 3-dehydroshikimate; **PABA**, *p*-aminobenzoic acid; **PHB**, *p*-hydroxybenzoic acid; **Phe**, phenylalanine; **Tyr**, tyrosine; **Trp**, tryptophan.

Archaeoglobus fulgidus

- Marine Archaeon
- Isolated from hydrothermal vents, oil wells
- Optimal growth at ~85 °C
- Genome is sequenced: 2,436 ORFs, 1,290 with no assigned biological role
- Energy: lactate + $1.5SO_4^{2-} + 4H^+ \longrightarrow 3CO_2 + 1.5H_2S + 3\,H_2O$

Shikimate Pathway Likely Contains Three Novel Enzymes



E4P

PEP

DAHP synthase

aroF,G,H

P_i

DAHP

NAD^+, Co^{2+}

DHQ synthase

aroB

P_i

DHQ

DHQ dehydratase

aroD

AF0228

DHS

DHS

NADPH + H^+

Shikimate dehydrogenase

aroE

$NADP^+$

AF2327

Shikimate

ATP

Shikimate kinase

aroK,L

ADP

3-Phosphoshikimate

PEP

EPSP synthase

aroA

P_i

AF1497

EPSP

Chorismate synthase

aroC

P_i

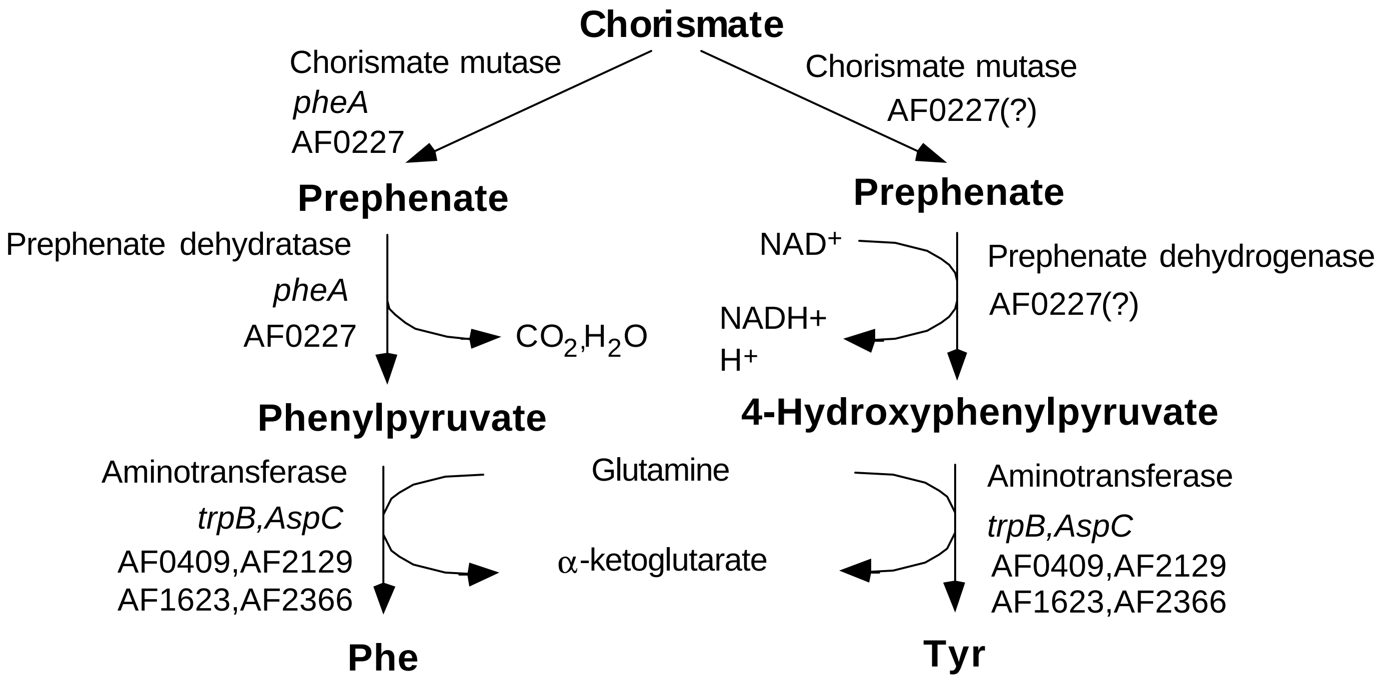
AF0670

Chorismate

Novel Enzyme

A. fulgidus ORFs (AF#) where identified. **E4P**, D-erythrose-4-phosphate; **PEP**, phosphoenolpyruvate; **DAHP**, 3-deoxy-D-*arabino*-heptulosonate-7-phosphate; **DHQ**, 3-dehydroquinate; **DHS**, 3-dehydroshikimate; **EPSP**, 5-enolpyruvoylshikimate-3-phosphate

A Novel Trifunctional Chorismate Mutase/Prephenate Dehydratase/Prephenate Dehydrogenase?



Chorismate

Chorismate mutase
pheA
AF0227

Chorismate mutase
AF0227(?)

Prephenate

Prephenate

Prephenate dehydratase
pheA
AF0227

CO_2,H_2O

NAD^+

NADH+
H^+

Prephenate dehydrogenase
AF0227(?)

Phenylpyruvate

4-Hydroxyphenylpyruvate

Aminotransferase
trpB,AspC
AF0409,AF2129
AF1623,AF2366

Glutamine

α-ketoglutarate

Aminotransferase
trpB,AspC
AF0409,AF2129
AF1623,AF2366

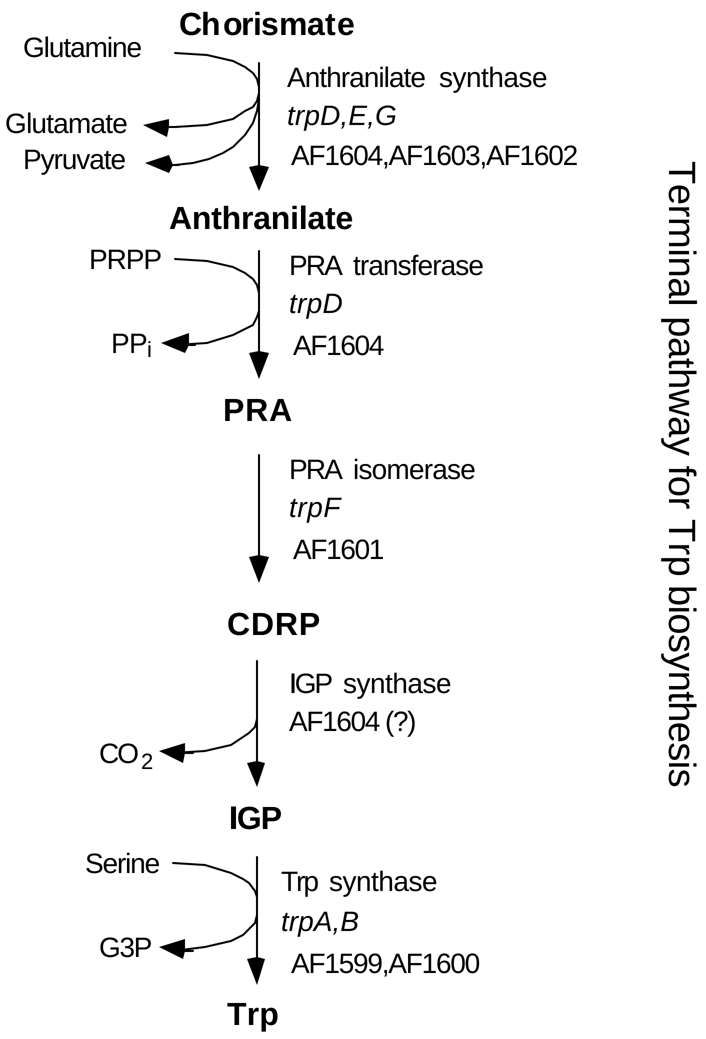
Phe

Tyr

Terminal pathways for Phe & Tyr Biosynthesis

No aromatic aminotransferase genes (*tyrB*) genes identified on the *A. fulgidus* genome

A Novel PRA Transferase/IGP Synthase?



Chorismate

Glutamine

Glutamate

Pyruvate

Anthranilate synthase

trpD,E,G

AF1604,AF1603,AF1602

Anthranilate

PRPP

PP_i

PRA transferase

trpD

AF1604

PRA

PRA isomerase

trpF

AF1601

CDRP

IGP synthase

AF1604 (?)

CO_2

IGP

Serine

G3P

Trp synthase

trpA,B

AF1599,AF1600

Trp

Terminal pathway for Trp biosynthesis

A. fulgidus ORFs (AF#) where identified. **PRA**, phosphoribosyl anthranilate; **CDRP**, 1-(*o*-carboxyphenylamino)-1-deoxyribulose-5-phosphate; **IGP**, indoleglycerol phosphate

All Identified Genes Cloned, Many Soluble Overexpression Products

AF#	**Enzyme**	**Gene**	**Size (bp)**	**AA's**	**MW_{calc}**	**Solubility**	**Comments**
AF0228	DHQ dehydratase	*aroD*	588	196	22206.99	Soluble	
AF2327	Shikimate dehydrogenase	*aroE*	807	269	29136.83	Soluble	
AF1497	EPSP synthase	*aroA*	1248	416	45085.29	Insoluble	monomer
AF0670	Chorismate synthase	*aroC*	1080	360	39262.66	Soluble	homotrimer
AF0227	Chor/preph/preph	*pheA*	1860	620	70946.38	Soluble	trifunctional
AF0409	Asp aminotransferase	*aspB-4*	1158	386	43140.16		homodimer
AF2129	Asp aminotransferase	*aspB-2*	1137	379	42698.89	Soluble	homodimer
AF1623	Asp aminotransferase	*aspB-3*	1170	390	43615.06	Soluble	homodimer
AF2366	Asp aminotransferase	*aspB-1*	1119	373	41703.09	Insoluble	homodimer
AF1603	Anthranilate synthase I	*trpE*	1233	411	46345.54	Insoluble	
AF1602	Anthranilate synthase II	*trpG*	534	178	19592.77		
AF1604	PRA transf/IGP synth	*trpD*	1638	546	59327.61	Insoluble	bifunctional
AF1601	PRA isomerase	*trpF*	597	199	22017.2	Soluble	monomer
AF1599	Trp synthase α subunit	*trpA*	744	248	27299.98	Soluble	
AF1600	Trp synthase β subunit	*trpB-2*	1191	397	43744.69	Soluble	

Cloning & Enzyme Characterization Results

- Putative gene for bifunctional, PRA transferase/IGP synthase (AF1604) contains a stop codon between functionalities
- Product of gene for hypothetical trifunctional chorismate mutase/prephenate dehydratase/prephenate dehydrogenase (AF0227) shows dehydrogenase activity
- Shikimate dehydrogenase partially purified and characterized
- Shikimate dehydrogenase and trifunctional enzyme clones sent to collaborators at NIST for crystal structure determination

Properties of *A. fulgidus* Shikimate Dehydrogenase (Preliminary)

- M_r ~ 32 kDa; M_{calc} = 29.1 kDa (*E. coli* M_{calc} = 29.4 kDa)
- $K_{m,\ NADP}$(87 °C, pH 7.5) = 120 ± 10 μM (*E. coli*, pH 8.5, 31 μM)
- $K_{m,\ NAD}$(87 °C, pH 7.5) = 8.5 ± 0.8 mM
- $K_{m,\ shik}$(87 °C, pH 7.5) = 160 ± 10 μM (*E. coli*, pH 8.5, 55 μM)
- T_{opt} ~ 90 °C; pH_{opt} ~ 7.5 (*E. coli* pH 8.5)
- Spec. Act. >300 U/mg; Retains ~20% of activity at 60 °C
- $t_{1/2}$(87 °C) ~ 1 hr; $t_{1/2}$(60 °C) ~ 6.5 hr

Zymograms for Novel Enzyme Separation and Activity-Based Detection

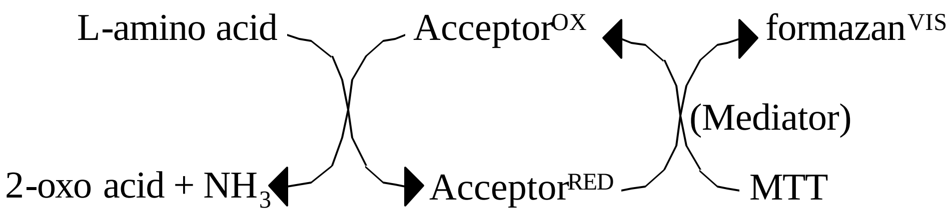
General Procedure

- Separate native proteins by electrophoresis
- Incubate gel with detection components at 85 °C

Comments

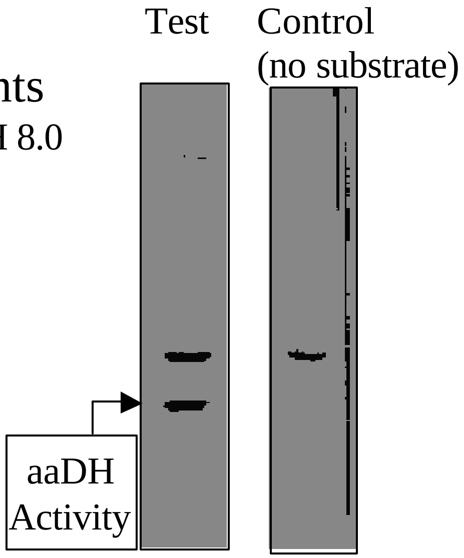
- Sensitive
- Separates proteins
- Detects isozymes
- Requires clever assay design
- Must purify protein for N-terminal sequence

Example: Amino Acid Dehydrogenase

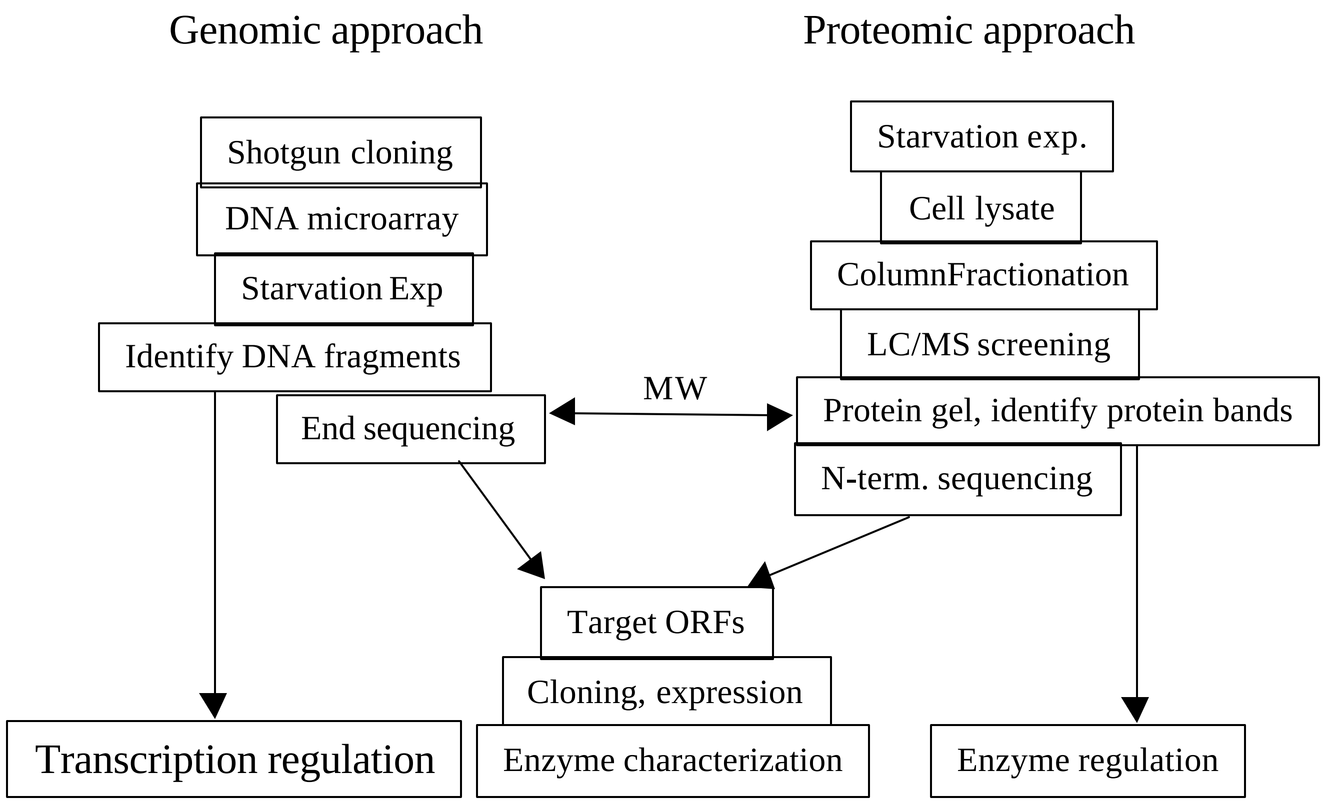


Detection Components

100 mM Tricene /KOH pH 8.0
1.0 mM L-amino acid
0.05 mg/ml MTT
0.01 mM PMS
10 mM NAD
10 mM NADP



Coordinated Use of DNA Microarrays and Activity-Based LC/MS Enzyme Assays



Genomic approach

Shotgun cloning

DNA microarray

Starvation Exp

Identify DNA fragments

End sequencing

Transcription regulation

Proteomic approach

Starvation exp.

Cell lysate

ColumnFractionation

LC/MS screening

Protein gel, identify protein bands

N-term. sequencing

Enzyme regulation

MW

Target ORFs

Cloning, expression

Enzyme characterization

95% Confidence Intervals for DNA Microarray Data

Hierarchical Linear Model

$$\log\left[\frac{(Cy5)}{f(Cy3)}\right]_{seg}$$

$$x_{seg} \sim N(m_{eg}, t_g^{\ 2})$$

Slide-effect variance

$$t_g^{\ 2} \sim k\tilde{t}_g^{\ 2} \big/ c_k^{\ 2}$$

$$m_{eg} \sim N(q_g, s_g^{\ 2})$$

Experimental effect variance

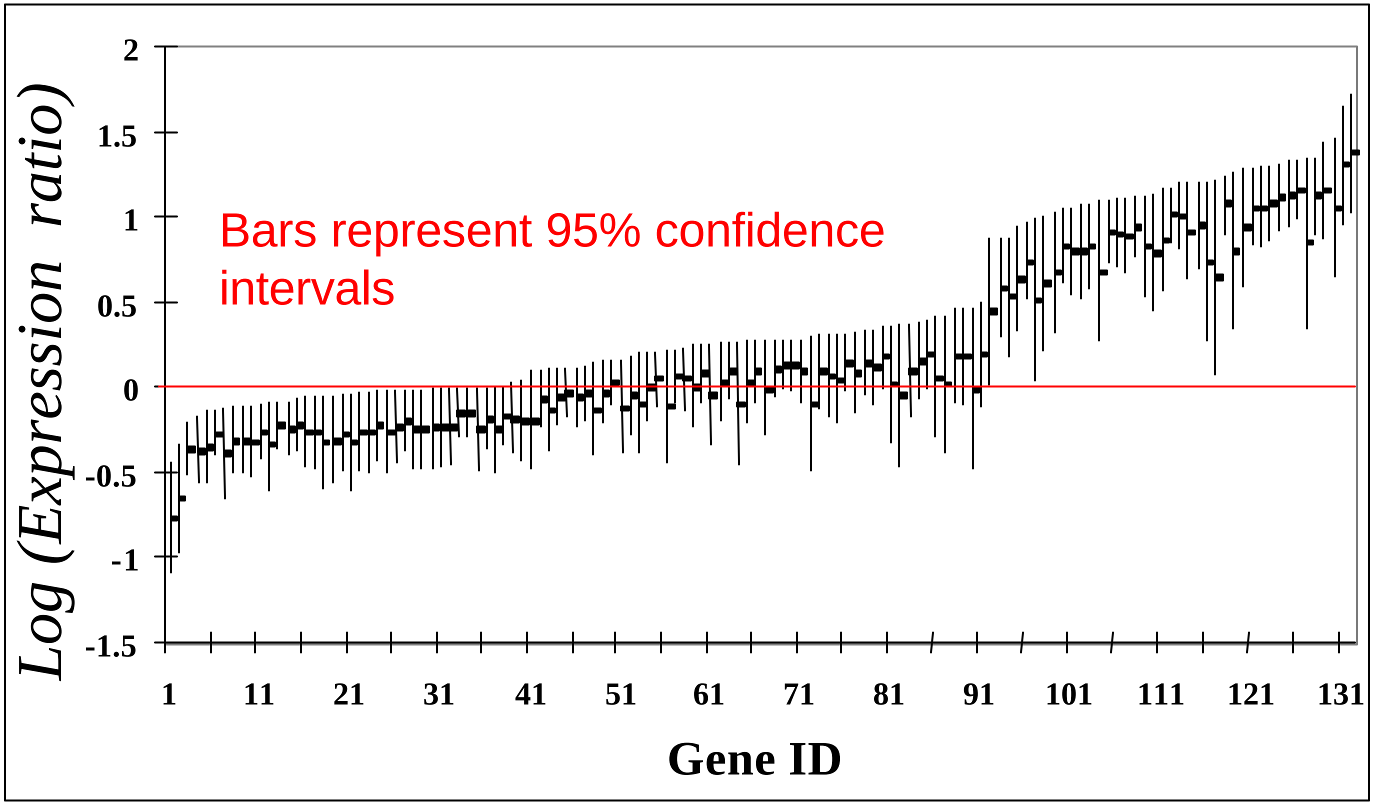
$$s_g^{\ 2} \sim h\tilde{s}^2 \big/ c_h^{\ 2}$$

Markov Chain Monte Carlo (MCMC) Method

(N=4000)

Tseng *et al.* (2001) *Nucleic Acid Res.*, **29**, 2549-2557

Expression Level Changes for *E. coli* Growth on Acetate vs. Glucose



Bars represent 95% confidence intervals

Concluding Remarks and Goals

- Early success cloning and expressing *A. fulgidus* genes in *E. coli*
- Continued characterization of cloned gene products
- DNA microarrays can resolve <2-fold expression levels
- Further improve standardization, normalization, and statistical analysis of DNA microarray data
- Coordinate application of DNA microarrays and LC/MS for enzyme/pathway discovery
- Focus on identification of novel genes/enzymes in shikimate pathway

Acknowledgements

- NSF (BES-9911718)
- NIST (70NANBOH0064)