CURRENT OPINION
www.current-opinion.com

The top genes: on the distance from transcript to function in yeast glycolysis

Commentary

Dan G Fraenkel

With its abundant components and extensive study, the glycolytic pathway in the yeast *Saccharomyces cerevisiae* would appear ideal to obtain and reconcile the 'omes of transcript, protein, metabolite and flux. But to do so is challenging and, as is often the case, close correlation of gene expression and function is elusive, even in this organism.

Addresses
Department of Microbiology and Molecular Genetics, Harvard Medical School, 200 Longwood Avenue, Boston, Massachusetts 02115, USA
e-mail: fraenkel@hms.harvard.edu

Current Opinion in Microbiology 2003, **6**:198–201

This review comes from a themed issue on
Cell regulation
Edited by Andrée Lazdunski and Carol Gross

1369-5274/03/$ – see front matter

DOI 10.1016/S1369-5274(03)00023-7

Transcripts, proteins, activity and function

In most organisms, 'housekeeping' genes are a relatively small fraction of the total, and among them are those for the glycolytic pathway. *Saccharomyces cerevisiae* has around 20 such genes, including those involved in the steps to ethanol (Figure 1, Table 1). In their 1977 paper [1], widely quoted for its estimate of 15 000 mRNA molecules per yeast cell, Hereford and Roshbash used Crot analysis to suggest that there are three groups of genes with respect to mRNA abundance, roughly 2500 genes contributing one or less molecule each, 400 giving 20 each and 20 giving 200 each. Twenty years later, using their SAGE technique (serial analysis of gene expression) on yeast, Velculescu *et al.* [2] observed a skewed distribution of recovered tags, and seven of the ten top contributing genes (and 13% of total tags) were glycolytics (see column 2 of Table 1; see also Figure 1). Hybridization with gene arrays should provide more complete data; the values for glycolytic genes are averaged from four reports in column 3, and make up 5% of the total signal.

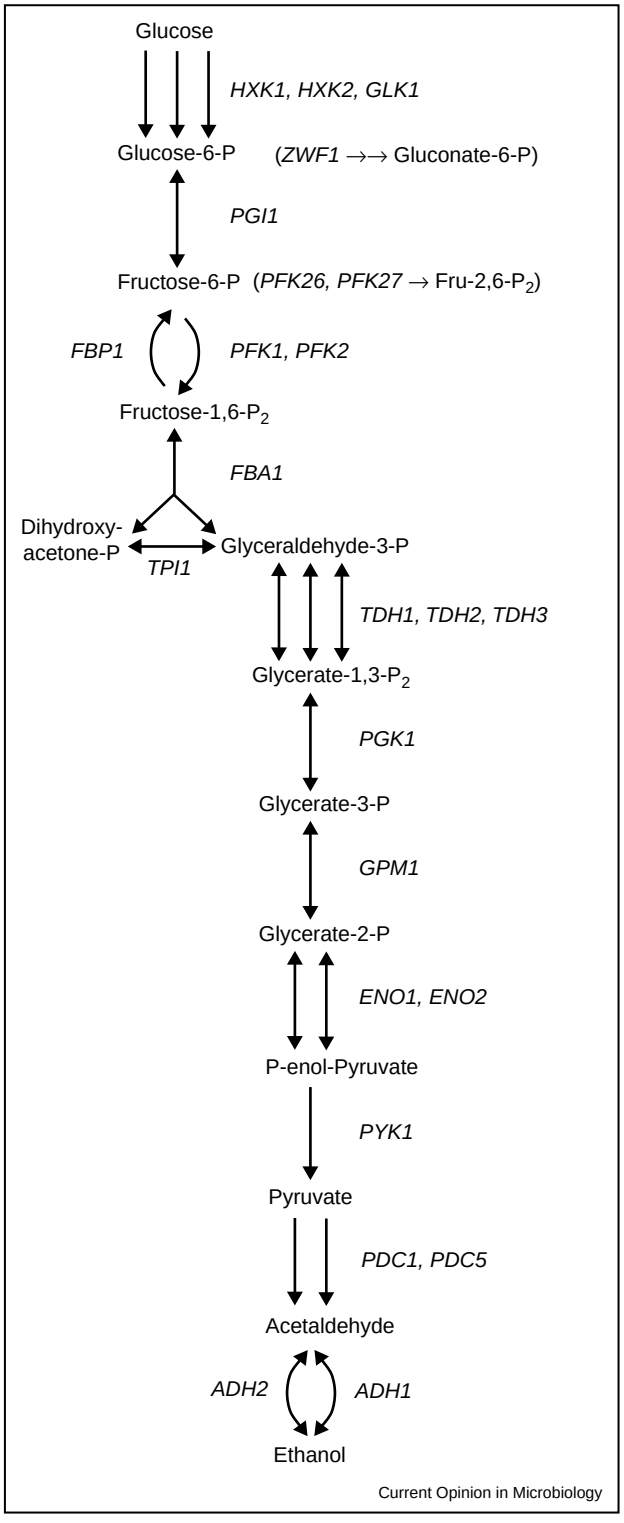
In three reports, seven to nine of the ten highest signals observed in two-dimensional protein gels were identified as being from glycolytic enzymes; these values, expressed per cell, are averaged in column 4 of Table 1, and come to 9% of the total protein subunits. A different type of estimate by Hess in 1965 [3] used specific activity of pure enzyme divided by that in crude extract and, applying knowledge of molecular weight, totalled the glycolytic enzymes to 65% of soluble protein; a revised calculation gives the values in column 5 of Table 1.

Although measures of protein, which are based on isotopic labeling (column 4), are probably more correct than those based on enzyme activity (column 5), for function it is usually activity that counts. A key item is the potential activity available, as measured in crude extracts and given in column 6 of Table 1 as the V_{max} for the forward direction. These values might, in turn, be compared with actual use of the reaction *in vivo*, (i.e. the net flux value which, for the 3-carbon intermediates of the pathway, is about 1 (column 7 of Table 1, same units as column 6). With respiration repressed in glucose-replete cells, that value of 1 would be close to the total rate of ATP synthesis and, hence, the sum of all reactions consuming ATP. In contrast, flux in biosynthetic pathways is much less: net polypeptide bond formation would have a value of approximately 0.05 and (an extreme) the net rate of synthesis of the phosphofructokinase activator fructose-2,6-biosphosphate, present at 10 μM, would have a value of approximately 3×10^{-8} (genes *PFK26* and *PFK27*). The abundant glycolytic enzymes indeed catalyze the most massive fluxes in the cell's metabolism. A comparison of columns 6 and 7 shows that, for this pathway, the reactions use enzymes present in activity adequate by assay, many in large excess.

Quality and uses of the data

Eventually, the individual entries in such a table (but obtained for a single strain) should make quantitative sense. But even for this well-known pathway, the data, by nature, are constrained in use and quality. Signals from transcript arrays, with their limitations, are almost always given as factor differences for genes between strains or conditions, and the reports of column 3 are biased to those with nominal abundance values readily available and their high-end values high. There is much variation even in these reports (and for the ten genes of highest apparent mRNA abundance, the sets are quite different; in Lopez and Baker [4] for example, 20% of the total signal comes from *GAG-POL* genes of the family of Ty elements). As quantitative determination of individual mRNAs is an exacting task [5], it is not yet possible to even cite firm values for the yeast glycolytic genes, let alone list the 'top ten'.

Figure 1



Glycolysis in outline. The genes depicted are *HXK1*, *HXK2* and *GLK1* (hexokinases 1 and 2, and glucokinase, respectively); *PGI1* (phosphoglucose isomerase); *PFK1* and *PFK2* (1-phosphofructokinase subunits); *FBA1* (aldolase); *TPI1* (triose-P isomerase); *TDH1-3* (glyceraldehyde-3-P dehydrogenase isozymes); *PGK1* (phosphoglycerate kinase); *GPM1* (phosphoglycerate mutase); *ENO1,2* (enolase isozymes); *PYK1* (pyruvate kinase); *PDC1* and *PDC5* (pyruvate decarboxylase isozymes); and *ADH1* (ethanol dehydrogenase isozyme). Also shown are *ADH2* (repressible ethanol dehydrogenase isozyme); *PFK26* and *PFK27* (2-phosphofructokinase isozymes for the synthesis of cofactor fructose-2,6-P_2); *FBP1* (fructose-1,6-bisphosphatase); and *ZWF1* (glucose-6-P dehydrogenase, the first enzyme of the pentose-P pathway). Not all reactants are shown.

Other entries are also insecure. Column 5 depends on activity measurements, with the usual uncertainties about incomplete extraction and alterations of activity or structure during purification. For column 6, in addition, the assay needs to be relevant to the milieu of the enzyme in the cells so as to compare with the *in vivo* flux values of column 7. To make sense of columns 6 and 7, kinetic parameters of the enzymes, obtained *in vitro*, might be matched with the *in vivo* flux and metabolite concentrations. However, such an exercise needs far more information (see for example, [6]), with issues of rapid metabolite turnover and of availability of methods [7]. Indeed, techniques for ready assignment of metabolite to cell compartments are yet to be devised.

In comparison with many other pathways, the glycolytic pathway is one of the more tractable; it is unbranched, has high flux, relatively high concentrations of components and adequate stability of many intermediates, and some of the enzymes can even be bought. And yet, the fitting of a single yeast glycolytic step to its *in vivo* context can be uncertain [8]. As the Amsterdam group concluded in its recent modeling of the entire pathway, only ''half of the enzymes matched their predicted flux *in vivo* within a factor of 2'' [9]. Such work shows both how much and how little is known, a general question being whether the limited fit is a matter of not getting the details quite right, or might instead reflect larger gaps in knowledge. Of course, as succinctly reviewed in the last paper written by the late Paul Srere [10], the cell has long been known as not just a bag of enzymes. But how that qualification applies to the glycolytic pathway in yeast is still intriguingly uncertain (see, for example, [11,12]), a century after the Buchners showed that fermentation can occur in cell extracts.

Constitutive expression

In growth on respiratory substrates such as lactate or ethanol (the latter of which occurs after glucose use in batch culture), the reversible steps of the glycolytic pathway and the one-way fructose-1,6-bisphosphatase reaction (gene *FBP1*) operate in the 'gluconeogenic' direction with net flux values of 0.05 or less. The known large derepression factor for *FBP1* is shown in Table 1 (mRNA, column 9), as is that for *ADH2* (protein, column 8). However, for glycolytic genes, the degree of control of expression has been less clear with, for example, reports of large [13] or small [14] factor differences for enzyme activity, glycolysis versus gluconeogenesis, and individual

Table 1

Transcript, protein and flux.

1		2	3	4	5	6	7	8	9
Gene		mRNA	mRNA	Protein	Protein	V_{max} in extract	Flux *in vivo*	Protein, derepn	mRNA, derepn
	(Method)	SAGE	Arrays	2D-gels	Purifn	Assay various	Glucose use	2D-gels	Arrays
	(Reference)	[2]	[4,20–22]	[23–25]			[8]	[25]	[4,26]
	(Units)	mol/cell	mol/cell	subunits/ cell, $\times 10^{-3}$	subunits/ cell, $\times 10^{-3}$	U/mg prot	U/mg prot	[ratio]	[ratio]
	(Medium)	Glucose	Glucose	Glucose	various	Glucose	Glucose	Ethanol/ Glucose	Derepn/ Glucose
1 HXK1,2 & GLK1		23	22	31	120	2	0.5	na	1.1, 3.2
2 PGI1		19	11	160	124	12	0.5	0.75	0.9, 0.7
3 PFK1,2		14	9	75	113	0.5	0.5	0.71	0.6, 1.0
4 FBA1		179	92	582	1,200	2	0.5	0.95	0.5, 1.2
5 TPI1		na	41	na	182	1	0.5	na	0.6, 0.6
6 TDH1-3		557	170	884	1,330	10	1	1.3	0.8, 1.5
7 PGK1		161	61	299	1,490	96	1	na	0.8, 0.7
8 GPM1		164	50	279	764	10	1	na	0.4, 1.5
9 ENO1,2		283	98	873	1,560	8	1	1.1	0.5, 0.4
10 PYK1		102	37	265	1,250	11	1	na	0.2, 1.0
11 PDC1,5		195	74	259	2,360	3	1	0.73	0.3, 0.4
12 ADH1		(252)	48	730	1,850	11	1	0.8	(0.6), 1.9
	(sum)	13%	5%	9%	25%				
13 ADH2		(252)	1	na	na	na	na	>20	(0.6), 3.0
14 PFK26,27		2	3	na	3.4 × 10e-2	1 × 10e-5	3 × 10e-8	na	3.9, 3.3
15 FBP1		na	1	na	0 (133)	0 (0.1)	0 (<0.05)	na	17, 14
16 ZWF1		1	1	na	16	0.2	na	na	1.7, 1.2

Rows 1–12 are for the glycolytic steps and 13–16 for some other reactions (see Figure 1). Entries for isozymes (other than ethanol dehydrogenases) are grouped together and not reported individually. Columns 2 and 3 (mRNA [SAGE method] and mRNA [array method], use a normalization value of 15,000 total mRNA molecules per cell [1]. Columns 3 and 4 average data from the references cited. Entries in column 4 and 5 use a protein content of 4×10^{-12} g and 50×10^{6} total subunits, per cell [25]. In column 6, V_{max} values are for the forward (glycolytic) direction. In column 9, the two entries are from [26] (post-/pre-glucose exhaustion) and [4] (lactate growth/glucose growth), respectively. Parenthetical entries indicate, in rows 12 and 13 (ADH1 and ADH2, respectively), that the two genes were not discriminated or, in row 15 (FBP1), values for gluconeogenic growth. Derepn, derepression; na, not available or applicable, purifn, purification.

assessment of mRNAs pointing to the smaller effects (reviewed in [15]). Proteome data shows small factors for most glycolytic proteins (column 8), whereas for gene arrays, a few of the entries for mRNA are somewhat larger (column 9). Overall, the glycolytic pathway in yeast appears to be 'fairly constitutive'.

The term is used advisedly, as there are significant differences in expression of certain glycolytic enzymes at the level of mRNA or protein; a few more would be apparent if Table 1 were less compressed. And even a factor of two for a highly expressed gene can be a lot of gene product. However, secondary effects related to growth rate and overall cell composition also contribute. One reason that differences in enzyme levels, gluconeogenesis versus glycolysis, are not as large as might be expected from the different flux values is that, for most of the two-way enzymes, their V_{max} values are also less in the gluconeogenic direction. However, several enzymes (genes *PFK1*, *PFK2*, *PYK1*, *PDC1* and *PDC5*) that are one-way in the glycolytic direction are also relatively highly expressed during gluconeogenesis. Other considerations might include: the need to avoid metabolite imbalance; the potential to rapidly alter flux if conditions change; the few situations or strains tested (selected for baking or brewing); starvation as the natural state of microbes; unknown functions of the pathway; gluconeogenesis preceding glycolysis in evolution; or just selective pressures having not always made for exquisite sensitivity of gene expression.

After all, in metabolism generally, pathway usage is not usually governed in the short term by gene expression; and it is widely observed (see, for example, [16]) that experimental changing of enzyme levels (up or down) need not have much effect on flux either (although metabolites might be perturbed [17]).

Indeed, in an even broader context, Gleaver *et al.* [18] have observed that the inducibility of many genes in yeast poorly correlates with the fitness defect of their deletion. To investigate these matters, it would help to have more ways to arrange small factor differences in chosen enzyme levels. Another wish [19] is to change those levels but avoid compensating expression of other genes.

References and recommended reading

1. Hereford LM, Rosbash M: **Number and distribution of polyadenylated RNA sequences in yeast**. *Cell* 1977, **10**:453-462.
2. Velculescu VE, Zhang L, Zhou W, Vogelstein J, Basrai MA, Bassett DE Jr, Hieter P, Vogelstein B, Kinzler KW: **Characterization of the yeast transcriptome**. *Cell* 1997, **88**:243-251.
3. Boiteux A, Hess B: **Design of glycolysis**. *Phil Trans Roy Soc Lond B* 1981, **293**:5-22.
4. Lopez MC, Baker HV: **Understanding the growth phenotype of the yeast *gcr1* mutant in terms of global genomic expression patterns**. *J Bacteriol* 2000, **182**:4970-4978.
5. Iyer V, Struhl K: **Absolute mRNA levels and transcription initiation rates in *Saccharomyces cerevisiae***. *Proc Natl Acad Sci USA* 1996, **93**:5208-5212.
6. Kashiwaya Y, Sato K, Tsuchiya N, Thomas S, Fell DA, Veech RL, Passonneau JV: **Control of glucose utilization in working rat heart**. *J Biol Chem* 1994, **269**:25502-25514.
7. Passonneau JV, Lowry OH: *Enzymatic Analysis: a Practical Guide*. Towata, New Jersey: Humana Press;1993.
8. Uemura H, Fraenkel DG: **Glucose metabolism in *gcr* mutants of *Saccharomyces cerevisiae***. *J Bacteriol* 1999, **181**:4719-4723.
9. Teusinck B, Passarge J, Reijenga CA, Escalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV *et al.*: **Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry**. *Eur J Biochem* 2000, **267**:5313-5329.
10. Srere PA: **Macromolecular interactions: tracing the roots**. *TIBS* 2000, **25**:150-153.
11. Gotz R, Schluter E, Shoham G, Zimmermann FK: **A potential role of the cytoskeleton of *Saccharomyces cerevisiae* in a functional organization of glycolytic enzymes**. *Yeast* 1999, **15**:1619-1629.
12. Gavin A-C, Bosche M, Krause R, Grandi P, Marzloch M, Bauer A, Schultz J, Rick JM, Michon A-M, Cruciat C-M *et al.*: **Functional organization of the yeast proteome by systematic analysis of protein complexes**. *Nature* 2002, **415**:141-147.
13. Maitra PK, Lobo Z: **A kinetic study of glycolytic enzyme synthesis in yeast**. *J Biol Chem* 1971, **246**:475-488.
14. Baker HV: **Glycolytic gene expression in *Saccharomyces cerevisiae*. Nucleotide sequence of *GCR1*, null mutants, and evidence for expression**. *Mol Cell Biol* 1986, **6**:3774-3784.
15. Chambers A, Packham EA, Graham IR: **Control of glycolytic gene expression in the budding yeast (*Saccharomyces cerevisiae*)**. *Curr Genet* 1995, **29**:1-9.
16. Fell D: *Understanding the Control of Metabolism*. London: Portland Press; 1997.
17. Raamsdonk LM, Teusink B, Broadhurst D, Zhang N, Hayes A, Walsh MC, Borden JA, Brindle KM, Kell DB, Rowland JJ *et al.*: **A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations**. *Nat Biotechnol* 2001, **19**:45-50.
18. Gleaver G, Chu AM, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, Arkin AP *et al.*: **Functional profiling of the *Saccharomyces cerevisiae* genome**. *Nature* 2002, **418**:387-391.
19. Heidrich K, Fraenkel DG: **Setting of graded levels of a protein in yeast by a t-degron technique as applied to phosphoglycerate mutase**. *BMC Genet* 2002, **3**:13.
20. Holstege FCP, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA: **Dissecting the regulatory circuitry of a eukaryotic genome**. *Cell* 1998, **95**:717-728.
21. Jelinsky SA, Ester P, Church GM, Samson LD: **Regulatory networks revealed by transcriptional profiling of damaged *Saccharomyces cerevisiae*: Rpn4 links base excision repair with proteosomes**. *Mol Cell Biol* 2000, **20**:8157-8167.
22. Roth FP, Hughes JD, Estep PW, Church GM: **Finding DNA regulatory motifs within unaligned noncoding sequences clustered by whole-genome mRNA quantitation**. *Nat Biotechnol* 1998, **16**:939-945.
23. Norbeck J, Blomberg A: **Two-dimensional electrophoretic separation of yeast proteins using a non-linear wide range (pH 3-10) immobilized pH gradient in the first dimension: reproducibility and evidence for isoelectric focussing of alkaline (pI>7) proteins**. *Yeast* 1997, **13**:1519-1534.
24. Gygi SP, Rochan Y, Franza BR, Aebersold R: **Correlation between protein and mRNA abundance in yeast**. *Mol Cell Biol* 1999, **19**:1720-1730.
25. Futcher B, Latter HI, Monardo P, McLaughlin CS, Garrels JI: **A sampling of the yeast proteome**. *Mol Cell Biol* 1999, **19**:7357-7368.
26. DeRisi JL, Iyer VR, Brown PO: **Exploring the metabolic and genetic control of gene expression on a genomic scale**. *Science* 1997, **278**:680-696.