Chapter 3
Multi-Omics Data-Driven Systems Biology of *E. coli*

Nobuyoshi Ishii and Masaru Tomita

Contents

Abstract The omics, which means comprehensive analysis of a specific layer in a cellular system, are emerging as essential methodological approaches for molecular biology and systems biology. However, single omics analysis does not always provide enough information to understand the behaviors of a cellular system. Therefore, a combination of multiple omics analyses, the multi-omics approach, is required to acquire a precise picture of living organisms. In this chapter, basic concepts of omics studies, and recent technologies in the omics of metabolism and published multi-omics analyses of *Escherichia coli*, are reviewed. Subsequently, a large-scale multi-omics analysis of *E. coli* K-12, including transcriptomics, proteomics, metabolomics and fluxomics, is presented. This study uncovered the complementary strategies of *E. coli* that result in a metabolic network robust against various types of perturbations, therefore demonstrating the power of a multi-omics, data-driven approach for understanding the functional principles of total cellular systems.

M. Tomita (✉)
Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata 997-0017, Japan
e-mail: tomita@ttck.keio.ac.jp

S.Y. Lee (ed.), *Systems Biology and Biotechnology of Escherichia coli*,
DOI 10.1007/978-1-4020-9394-4_3,

3.1 Overview of Omics and Multi-omics Analysis

3.1.1 A New Approach in Molecular Biology – Omics

The history of molecular biology is defined by a number of innovations. Currently, a new innovative breakthrough is joining the world of the molecular biology – the so-called "omics" (Lee et al. 2005, Yadav 2007). The basic methods of modern molecular biology have been, to simplify the situation somewhat, hypothesis-driven, reductionist, and bottom-up. In most cases, only a few biochemical species are focused on in any one study based on hypotheses formed by the researcher(s) prior to the start of the study. After that, exhaustive investigation is directed towards understanding the properties of the target molecules.

Although such an approach is still valuable for obtaining detailed and precise knowledge of the target molecular species, some inherent problems exist in how the conventional research flows. At the beginning stage of such classical research schemes, the selection of the target strongly depends on the personal experience and intuition of individual researchers. Moreover, information about limited numbers of molecular species does not always provide insight into a biological "system" that consists of networks formed by a number of interacting molecular species (Bruggeman and Westerhoff 2007).

To overcome these weak points in traditional molecular biology, a novel research area, the "omics", is emerging. Omics means a comprehensive analysis of biochemical molecular species or interactions of molecules belonging to a specific layer in a cellular system. For example, "genomics" is defined as the study of whole DNA sequences and the information contained therein. Many different words having the suffix of "omics" have been proposed - transcriptomics, proteomics, lipidomics, glycomics, interactomics, phenomics, and so on. However, all omics approaches can be considered to share two major features in contrast to traditional procedures.

One feature is changing the direction of the flow of analysis. Unlike traditional methods, in omics approaches massive data is first collected with no prior hypothesis, and meaningful targets are searched for within the obtained data set. The second feature of omics is the attempt to understand the target as a total "system" by using information of the relationships between many measured molecular species. From this point of view, the omics can be expected to contribute to the progress of systems biology.

In brief, omics can be said to be a data-driven, holistic, and top-down approach, as opposed to traditional approaches. Rapid advances in the development of high-throughput measuring instruments are inducing dramatic growth in the omics research area. The extreme progression of information technologies, including enhancements of public web-databases of biological knowledge (Caspi et al. 2008, Kanehisa et al. 2008, Teufel et al. 2006, Wittig and De Beuckelaer 2001), also support the expansion of omics studies, which require the handling of hundreds or thousands of measured values.

3.1.2 Omics for Metabolic Systems – Metabolomics and Fluxomics

While many technologies are included in the "omics" family, most activity is found in the following three areas: genomics, transcriptomics, and proteomics. Currently, whole genome sequences of many species have been deciphered by high-performance DNA sequencers. Using the information of a complete genome sequence, most of the products (mRNAs and proteins) coded in the genome can be predicted; nonetheless, post-transcriptional modifications cannot be ignored. Thus, with transcriptomic or proteomic experiments, the near complete detection of biochemical species, i.e., true omics analysis, is possible in principle. However, for other omics studies, it is impossible to define an explicit number of targets.

Although difficult to comprehensively measure, omics analysis of metabolism in cellular systems is highly important (Fiehn 2002). The phenotype of a strain is strongly connected to the profile of metabolite concentrations in the cell. In many cases, adaptations of living cells to environmental changes can be achieved by reconfiguration of enzymatic reaction rates in some metabolic pathways. Therefore, metabolomics (Dettmer et al. 2007, Kell 2004, Mashego et al. 2007, Oldiges et al. 2007, Rabinowitz 2007, Wang et al. 2006), which is the omics study for metabolic compounds (low molecular weight, typically less than 1 kDa), is desired to obtain a more precise overview of life.

Traditionally, large-scale metabolite analysis has been performed by gas chromatography mass spectrometry (GC-MS) (Fiehn et al. 2000), and GC-MS is frequently used in plant metabolomics studies (Sanchez et al. 2008). Other instruments, including liquid chromatography mass spectrometry (LC-MS) (Chen et al. 2007, Tolstikov et al. 2007) and nuclear magnetic resonance (NMR) (Grivet et al. 2003, Jordan and Cheng 2007, Ward et al. 2007), have also been successfully applied to metabolome analyses.

Capillary electrophoresis mass spectrometry (CE-MS) has emerged as a powerful new tool, and various CE-MS methods have been developed for the analysis of charged metabolites (Gaspar et al. 2008, Monton and Soga 2007, Sniehotta et al. 2007, Song et al. 2008). The advantages of CE-MS compared to other separation technologies are that this method exhibits extremely high resolution and that almost any charged species can be infused into MS. (Soga et al. 2003) developed a metabolome analysis method by CE-MS whereby metabolites were first separated by CE based on charge and size and then selectively detected using MS by monitoring over a large range of m/z values. Since hundreds of metabolites can be detected simultaneously by CE-MS, our understanding of the metabolic layer in cellular systems is being greatly expanded. More recently, (Soga et al. 2006) also constructed a coupling of CE and time-of-flight MS (TOFMS), and their CE-TOFMS analysis revealed that serum ophthalmate is a sensitive indicator of hepatic glutathione depletion in mice.

Another new methodology, called fluxomics (Sanford et al. 2002, Sauer et al. 1999, Wiechert et al. 2007), which means detailed metabolic flux analysis (MFA)

(Stephanopoulos et al. 1998) of large-scale metabolic pathways, has joined the omics family for investigating metabolic systems. MFA includes mathematical procedures for the estimation of unmeasurable reaction rates in a metabolic pathway by using measurable data, such as specific consumption rates of substrate and specific production rates of byproduct. MFA itself has relatively a long history – the first MFA is believed to have been conducted by Aiba and Matsuoka in 1979 (Aiba and Matsuoka 1979, Stephanopoulos et al. 1998). However, after the 1990s, the use of stable-isotope labeled substrates has become a common technique, and some advanced algorithms to handle the information of labeled metabolites for calculating metabolic fluxes have been developed (Noh et al. 2006, Sauer 2006, Shimizu 2004, Wiechert 2001). Accordingly, metabolic pathways that have complex topologies can be treated by current MFA technologies, i.e., metabolic fluxes distributed in a wide network can now be estimated. Therefore, fluxomics can be considered as one of the omics methodologies. Figure 3.1 shows a bibliographic search containing the words "metabolomics or metabolome" or "fluxomics or fluxome" using PubMed (http://www.pubmed.gov/). An exponential increase in the number of metabolomics studies and the genesis of fluxomics research can be observed.

A combination of metabolomics and fluxomics has been established by Toya et al. (Toya et al. 2007). They used CE-TOFMS to measure mass distributions of intermediate metabolites in cells cultured by isotope-labeled glucose, and performed flux analysis with the measured mass distribution patterns. Since the pool sizes of intermediate metabolites are generally so small, isotopic pseudo-steady states of intermediate metabolites are immediately achieved (Wiechert and Noh 2005). Accordingly, MFA using CE-TOFMS can be applied to metabolic systems under drastic dynamical change, such as in a batch culture, which is practically important in fermentation industries. Other methods of MFA using LC-MS to determine labeling patterns of intermediate metabolites have also been reported (Costenoble et al. 2007, Noh et al. 2007, Schaub et al. 2008, van Winden et al. 2005). Further

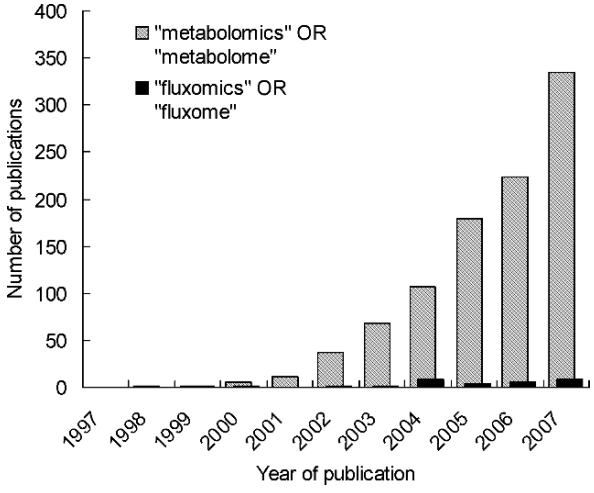

Fig. 3.1 Bibliographic searches in PubMed (http://www.pubmed.gov/) containing the keywords "metabolome OR metabolomics" and "fluxome OR fluxomics" (as of May 20, 2008)

collaborations of metabolomics and fluxomics are expected to be developed and to be employed in investigations of complex and large-scale metabolic systems.

3.1.3 Integration of Various Omics Analyses – Multi-omics

These days various omics analyses are frequently employed in many experimental studies. However, it has been gradually realized that obtaining useful biological knowledge from a single type of omics data (for example, DNA microarray only) is no easy task. One reason is that single omics analysis provides us with information about only one layer of a cellular system. Obviously, multiple functional cellular layers, including the mRNA, protein, and metabolite layers, are interacting with each other; thus the response of a total cellular system to given perturbations cannot be fully captured from a single layer. Figure 3.2 shows a schematic diagram of the functional layers and their interactions in a cellular system.

In conclusion, not just one omics analysis, but multiple omics analyses are required for deep and precise understanding of a cellular system. This recognition seems to be shared by many researchers (Andersen et al. 2008, De Keersmaecker et al. 2006, Joyce and Palsson 2006, Steinfath et al. 2007). Toyoda et al. proposed the concept of the "omic space", which consists of multi-layered state variables, and suggested a data integration framework and graphic presentation method of multiple omics data (Toyoda et al. 2007, Toyoda and Wada 2004). Figure 3.3 displays a conceptual diagram of the "omic space". (Lee et al. 2005) indicated the essen-

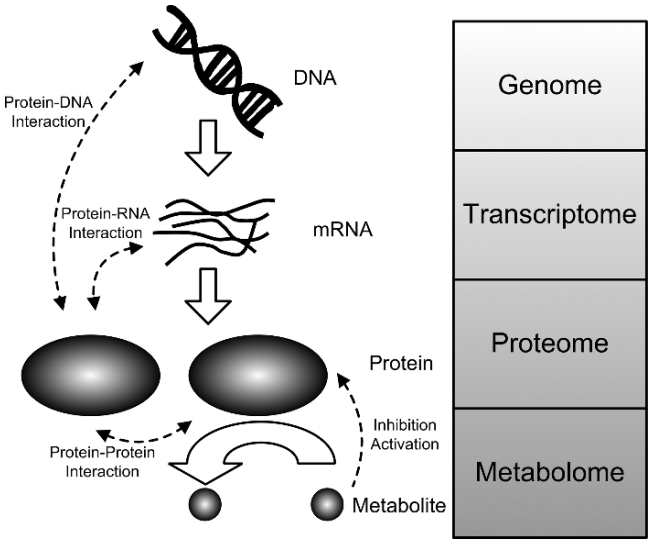

Fig. 3.2 Schematic diagram of interactions among various functional layers in a cellular system. *Blank arrow*, flow of biological information; *dashed line*, possible interaction between various biomolecular species

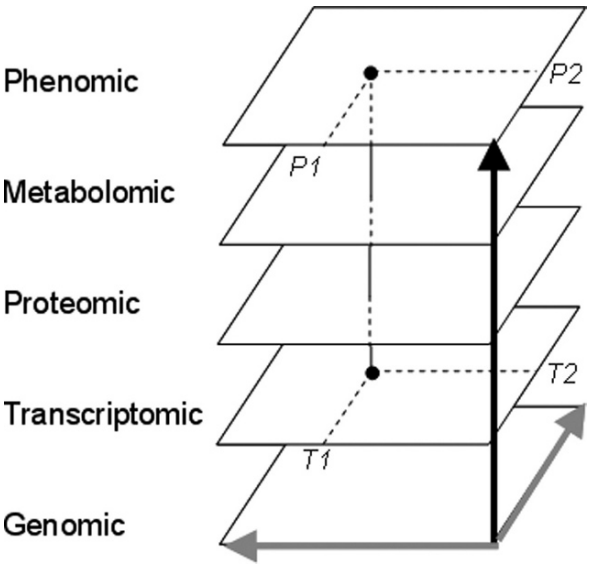

Fig. 3.3 Conceptual projection view of omic space (Toyoda and Wada 2004). *Black* arrow, direction of ascending order in transcriptomic, proteomic, metabolomic and phenomic planes. *Gray* arrow genomic-coordinate axes. The epistatic P1–P2 interaction on the phenomic plane corresponds to T1 and T2 genes interacting on the transcriptomic planes

tiality of the combination of multiple omics analyses for strain improvements in fermentation industries. (Paley and Karp 2006) developed the "Omics viewer" that can show different types of data sets (for example, measurements of gene expression and metabolite concentrations) simultaneously on a metabolic pathway map. (Arakawa et al. 2005) also developed a mapping tool to display complex omics data together.

Excellent studies using a combination of multiple omics methods have begun to be reported. Confining examples to studies of *Escherichia coli*, the following works can be found: Yoon et al. (2003) carried out combined transcriptomic (DNA microarray) and proteomic (two-dimensional gel electrophoresis; 2-DE) analyses of *E. coli* during high cell density cultivation, which is required for higher productivity of recombinant proteins. They showed that patterns of gene expression were mostly similar to patterns of protein expression, except for several discrepancies observed for a few genes (Fig. 3.4). Fong et al. (2006) investigated transcriptomics (DNA microarray) and fluxomics (^{13}C-labeled glucose was used as a substrate, and label patterns of amino acids of hydrolyzed cells were measured by GC-MS) of *E. coli* to reveal the mechanisms of adaptive mutations of some gene-disrupted strains. They found that activation of latent pathways and flux changes in the tricarboxylic acid (TCA) cycle in the adaptive evolved strains correlate well with changes in the transcriptome. Bore et al. (2007) performed transcriptomics (quantitative reverse-transcription polymerase chain reaction; qRT-PCR) and proteomics (peptide mass fingerprinting) to study *E. coli* adaptation to benzalkonium chloride, which is a commonly used disinfectant and preservative. Their analysis indicated that benzalkonium chloride treatment might result in superoxide stress in *E. coli*. Wittmann et al. (2007) studied the fluxome (GC-MS analysis of labeled proteinogenic amino acids) and metabolome (enzymatic analyses) of *E. coli* during temperature-induced

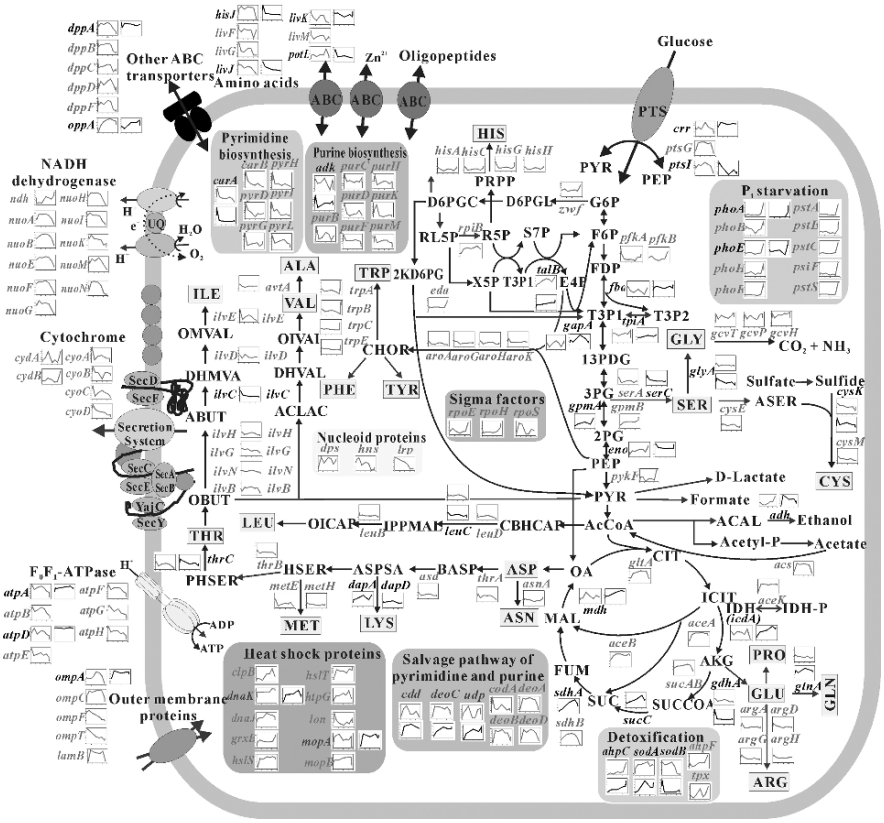

Fig. 3.4 Transcriptome and proteome analysis of *E. coli* during high cell density culture. (From Yoon SH, Han MJ, Lee SY, Jeong KJ, Yoo JS (2003) Combined transcriptome and proteome analysis of *Escherichia coli* during high cell density culture. Biotechnol Bioeng. 81(7):753–767. Copyright © 2003 by Wiley Periodicals, Inc. Reprinted with permission of Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc.). X axis, cell concentration (g DCW/L); Y axis, expression level in log2 scale for transcriptome (*gray*) and in absolute value of volume % for proteome (*black*); *gray-colored* gene name, only mRNA level was detected; *black-colored* gene name, both mRNA and protein level were detected

recombinant production of human fibroblast growth factor. Their analysis showed a relationship between the adenylate energy charge drop and an increase in the glycolytic flux. Other regulations in central carbon metabolism were also estimated. (Durrschmid et al. 2008) performed transcriptomics (DNA microarray) and proteomics (two-dimensional difference gel electrophoresis (Marouga et al. 2005); 2D-DIGE) analyses of *E. coli* stress response mechanisms towards recombinant protein expression. Their investigation of the expression of two model proteins demonstrated that there is a distinct impact of recombinant proteins, particularly on levels of known stress regulatory genes and proteins, as well as on the response

Table 3.1 Reported multi-omics analyses of *E. coli*

	Transcriptome	Proteome	Metabolome	Fluxome
Yoon et al. 2003	○	○		
Fong et al. 2006	○			○
Bore et al. 2007	○	○		
Wittmann et al. 2007			○	○
Durrschmid et al. 2008	○	○		

associated with ArcA and *psp*. Table 3.1 summarizes these multi-omics research studies targeting *E. coli*.

In 2001, the Institute for Advanced Biosciences (IAB) of Keio University was founded at Tsuruoka City, Yamagata, Japan. The purpose of the IAB is to actualize the crossover association of different research fields, including genomics, proteomics, metabolomics and informatics, for the establishment of "integrative systems biology" to obtain a more complete picture of living organisms. *E. coli* was selected as the primary target of the IAB, and a multi-omics approach was applied to reveal the basic principles of cellular responses of *E. coli* to genetic or environmental perturbations (Ishii et al. 2007). In the following section, a large-scale multi-omics study performed in the IAB is presented.

3.2 Multi-omics Analysis of *E. coli*

3.2.1 Chemostat Cultures of the Keio Collection

The Keio collection (Baba et al. 2006), which is the complete collection of all single-gene disruptants of *E. coli* K-12, was used for this study. From the Keio collection, 24 single-gene disrupted strains were selected. These strains are disruptants of genes in glycolysis or pentose phosphate pathway metabolism. These metabolic pathways are parts of the "central carbon metabolism", which functions to supply energy and synthesize essential precursors used for cellular components. Since central carbon metabolism is crucial for living cells, the disruption of genes involved in this metabolism was expected to result in dramatic changes in the cellular system. A uniform dilution rate of $0.2\,h^{-1}$ was applied to the chemostat cultures of the Keio collection strains.

Gene disruption can be thought of as an "internal" perturbation to the cell. On the other hand, "external" perturbation can be added by changing environmental factors. In this study, we chose substrate concentration change as the external perturbation. This was carried out by changing the dilution rate of the chemostat culture. Table 3.2 summarizes the strains and culture conditions used, and Fig. 3.5 shows the pathway map of central carbon metabolism of *E. coli* and the positions of disrupted single genes examined in this study.

Table 3.2 Strain and culture conditions

Strain	*E. coli* BW25113 *galM, glk, pgm, pgi, pfkA, pfkB, fbp, fbaB, gapC, gpmA, gpmB, pykA, pykF, ppsA, zwf, pgl, gnd, rpe, rpiA, rpiB, tktA, tktB, talA, talB*
Medium	Modified M9
Carbon source	Glucose
Oxygen supply	Aerobic
Temperature	37 °C
pH	7.0
Dilution rate	0.2 h^{-1} (for single-gene disruptants) 0.1, 0.2, 0.4, 0.5, 0.7 h^{-1} (for wild-type)

3.2.2 Transcriptome, Proteome, Metabolome, and Fluxome Analysis

The performed multi-omics analysis included layers closest to the genome and those closest to the phenotype, i.e., including transcriptomics, proteomics, metabolomics and fluxomics. Both cell-wide semi-quantitative analysis and targeted quantitative methods were employed in the transcriptome and proteome analyses. The transcriptome analysis was performed by DNA microarray for 4213 genes and qRT-PCR for 85 genes involved in central carbon metabolism. The proteome analysis was carried out with 2D-DIGE (approximately 2000 proteins were detected) and quantitative methods using liquid chromatography-mass spectrometry/mass spectrometry (LC-MS/MS) for 57 proteins involved in central carbon metabolism. The metabolome analysis was performed by CE-TOFMS for 579 metabolites. To perform the fluxome analysis, ^{13}C-labeled glucose was used as a substrate and mass distributions of proteinogenic amino acids of cultured cells were measured by GC-MS. The metabolic fluxes were calculated from the information of the obtained mass distributions. Table 3.3 summarizes the omics technologies employed in this study. All measurement data is published on our website (http://ecoli.keio.ac.jp/).

The obtained data set was used to analyze the response of the cellular system to the perturbations. For this purpose, two-step normalizations were applied to the measurement values (Ishii et al. 2007), and final converted values are named as "expression index" (EI).

3.2.3 Observed Robustness in E. coli Metabolic System

Figure 3.6 shows EIs of quantitative measurements (qRT-PCR for mRNAs, LC-MS/MS applied methods for proteins, and CE-TOFMS for metabolites) for all samples. Upon first glance of this figure, mRNAs and proteins seem to vary with the change of specific growth rate (equal to the dilution rate in a chemostat culture). Surprisingly, no clear changes of mRNAs and proteins were found for most single-gene disruptants, even when the disrupted gene concerns crucial central carbon

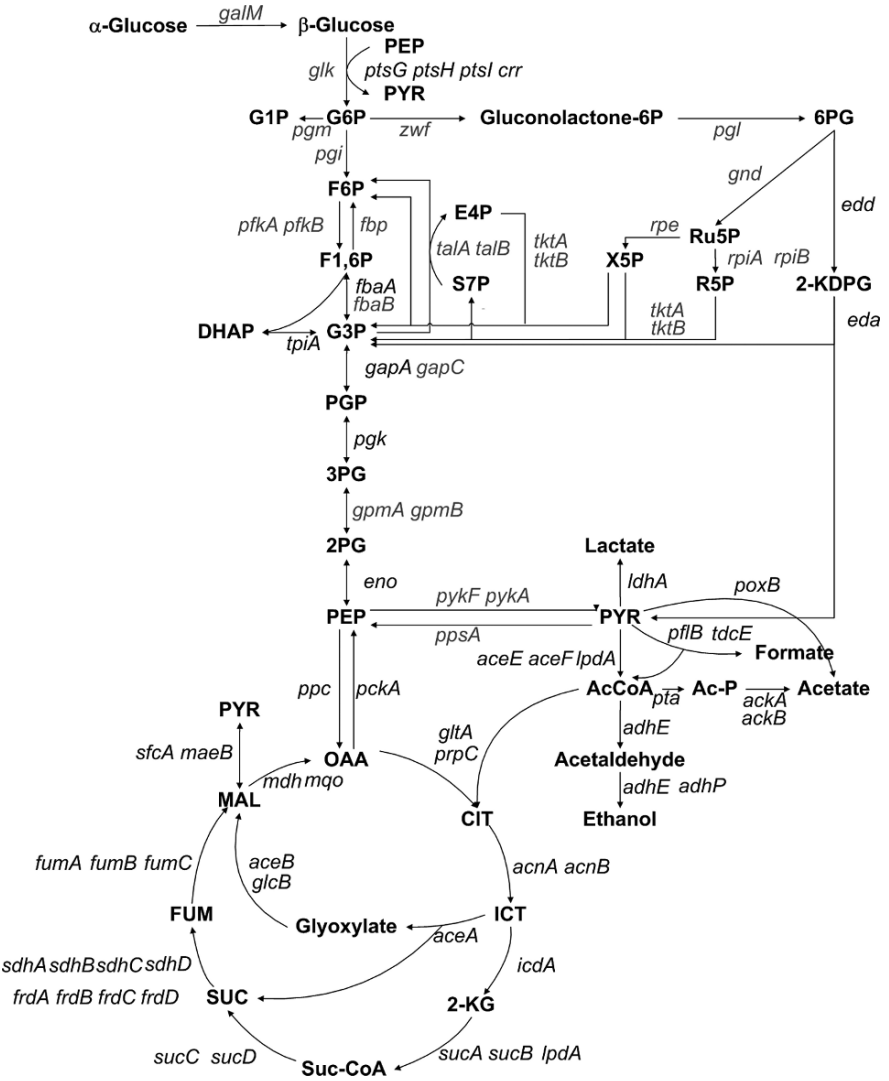

Fig. 3.5 Map of *E. coli* K-12 central carbon metabolism. (Modified from Ishii et al. 2007). Bold font, metabolites; italics, genes. *Gray* character genes are examined single-gene disruptions

metabolism. Moreover, no significant or regular change was observed for metabolites in both growth rate changed samples and single-gene disrupted samples. Some nucleotides in single-gene disruptants showed relatively large variances, but this is probably because of instability and/or low extraction efficiency of the nucleotide compounds. To authenticate these findings, averages of absolute values of the EI included in a specific category (i.e., mRNAs, proteins, or metabolites) were calculated and referred to as the average expression index (AEI). Figure 3.7 shows the AEIs of each category.

Table 3.3 Employed technologies in the multi-omics study of *E. coli*

	Used technology	Number of measured chemical species
Transcriptomics	DNA microarray	4213
	qRT-PCR	85
Proteomics	2D-DIGE	2000 (approximately)
	LC-MS/MS	57
Metabolomics	CE-TOFMS	579
Fluxomics	GC-MS	104 (isotopomers of fragment from proteinogenic amino acids)

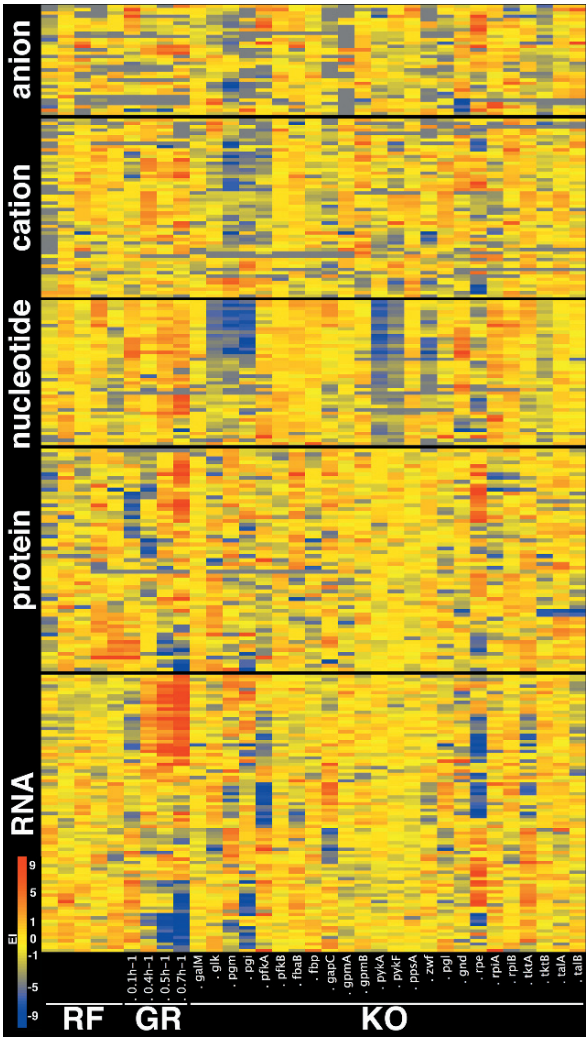

Fig. 3.6 Heatmap of the EI values of intracellular components. (Modified from Ishii et al. 2007). The heatmap shows the EI values of intracellular components that were detected in more than half the samples. RF, reference sample (wild-type cells cultured at a specific growth rate of $0.2\,h^{-1}$); GR, wild-type cells cultured at the indicated specific growth rates; KO, single-gene knockout mutants cultured at a specific growth rate of $0.2\,h^{-1}$

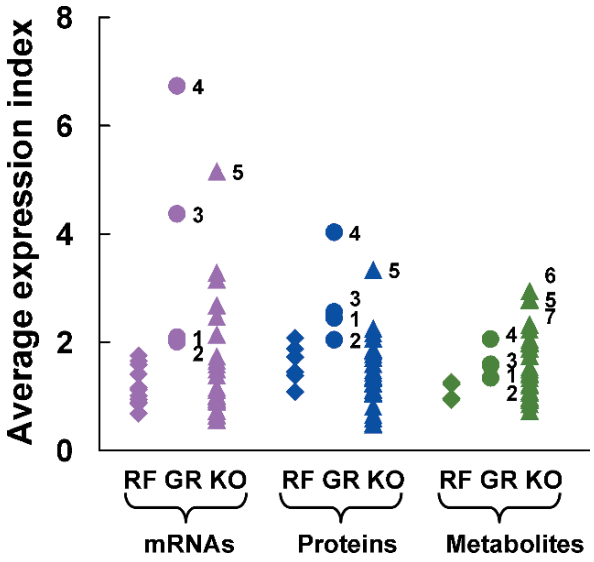

Fig. 3.7 AEI values for quantitative measurements obtained by targeted analysis. (Modified from Ishii et al. 2007). RF, reference sample (wild-type cells cultured at a specific growth rate of 0.2 h^{-1}); GR, wild-type cells cultured at the indicated specific growth rates; KO, single gene knockout mutants cultured at a specific growth rate of 0.2 h^{-1}. Numbers 1, 2, 3 and 4, correspond to specific growth rates of 0.1, 0.4, 0.5 and 0.7 h^{-1}, and numbers 5, 6 and 7 correspond to *rpe*, *pgi* and *pgm* disruptants, respectively

The AEIs for mRNAs and proteins gradually increased at higher growth rates. This suggests that *E. coli* actively regulates global gene and protein levels to meet increasing metabolic demands. Meanwhile, the AEI values for metabolites did not change significantly with the growth rate. This relative stability in metabolite level may be a consequence of the active regulation of enzyme expression. Focusing on local pathways, large changes of expression levels of proteins related to energy supply under aerobic condition were observed accompanying an increase in the specific growth rate (Ishii et al. 2007).

In contrast to the changes observed in wild-type cells cultured at various growth rates, the AEIs for both mRNAs and proteins in most gene-disruptants showed small changes, which fell within the range of variation observed in wild-type samples at the same specific growth rate (i.e., reference samples). In comparisons of targeted analyses of mRNAs (qRT-PCR) and proteins (LC-MS/MS), the AEI values in all disruptants were smaller than the AEI values observed for wild-type cells at a specific growth rate of 0.7 h^{-1}. Similar results were obtained for the AEI values representing the global analysis of expression of mRNAs (DNA microarray) and proteins (2D-DIGE) (Ishii et al. 2007). An overview of the changes in AEIs explained above is displayed in Table 3.4.

Table 3.4 Changes in AEIs

	Growth rate change (wild-type)	Most of examined single-gene disruptants
mRNAs	+	−
Proteins	+	−
Metabolites	−	−

+: Variation among samples was large. −: Variation among samples was small.

These findings suggest that *E. coli* does not appreciably respond to the loss of a single enzyme in central carbon metabolism by regulating the abundance of other compensatory enzymes. Actually, in most single-gene disruptants, the expression level of isozyme coding genes was almost same as the level in the wild-type strain (Ishii et al. 2007). In single-gene disruptants, a stable metabolic state is maintained by using remaining isozymes or by rerouting metabolic fluxes. For example, in the *zwf*-disruptant, some fluxes flow in a countercurrent direction compared to the wild-type, as reported in a previous study (Zhao et al. 2004)

Two exceptions were the *pfkA*-disruptant and *rpiA*-disruptant (Ishii et al. 2007). In these strains, potential mutations in genes other than the disrupted gene were checked, and various mutations enhancing the expression level of compensatory isozymes of the disrupted gene (*pfkB* in *pfkA*-disruptant and *rpiB* in *rpiA*-disruptant) were found, as reported in previous studies (Daldal 1983, Skinner and Cooper 1974).

3.3 Concluding Remarks

Changes in the dilution rate of a chemostat culture correspond to changes in the concentration of growth-limiting substrates, and thus various settings of the dilution rate can be regarded as an environmental perturbation for *E. coli*. On the other hand, the disruption of a gene can be thought of as an intracellular perturbation. Our multi-omics analysis demonstrates that the metabolic network of *E. coli* is markedly robust against both types of perturbations. *E. coli* can actively respond to changes in the concentration of growth-limiting substrates by regulating the level of enzyme expression to maximize growth rate, which is reflected in the observed stability of metabolite levels. However, this strategy may come at a high cost, because the cell must prepare additional systems (such as sensor proteins, signal mediators, and transcriptional regulators) to detect and react appropriately to each specific perturbation. This strategy contrasts with the finding that *E. coli* does not appear to reconfigure mRNA or protein levels actively when most single metabolic genes are disrupted. In this case, structural redundancy in the metabolic network itself provides the necessary robustness. As a result, the levels of most metabolites remain at wild-type levels, although some localized perturbations may occur. This strategy seems to save more energy than the active regulation of mRNA or proteins, because it requires no specific molecular machinery for detecting each mutation. Even if this strategy appears insufficient in the face of some mutations, *E. coli* may survive by accumulating additional mutations, as observed for *pfkA* and *rpiA* disruptants. Using multiple strategies may thus enable *E. coli* to maintain a stable metabolic state when exposed to various types of perturbations.

Biological robustness is one of the central subjects in systems biology (Kitano 2004), and conceptual descriptions or analyses with mathematical models have been attempted to explain how robustness is achieved (Kitano 2007). Furthermore, some omics or multi-omics analyses to study robustness in real cells have also been reported. For example, (Becker et al. 2006) performed a proteomics analysis

of *Salmonella*, and found extensive metabolic redundancies and access to diverse host nutrients. (Gibon et al. 2006) measured the transcriptomes, proteomes, and metabolomes of *Arabidopsis rosettes* wild-type and *pgm*-disruptant, and demonstrated that the amplitudes of diurnal changes in metabolite levels in *pgm* were similar or smaller than those in the wild-type. The above mentioned multi-omics analysis of *E. coli* also supports the existence of robustness as a common principle to ensure survival in the face of countless accidents.

The next challenge of the multi-omics data-driven systems biology of *E. coli* is to construct a mathematical model incorporating the obtained multi-omics data to elucidate a tangible mechanism of metabolic robustness in *E. coli*. Analyses using a mathematical model will suggest methods for breaking cellular robustness to enhance the productivity of useful metabolic compounds. Finally, and needless to say, the integrative multi-omics approach can be applied to many organisms, not just microorganisms, and thus expanding applications of this approach can be expected in the future.

Acknowledgments This study was carried out with the collaboration of following members of the IAB: Ishii N*, Nakahigashi K*, Baba T*, Robert M*, Soga T*, Kanai A*, Hirasawa T*, Naba M, Hirai K, Hoque A, Ho PY, Kakazu Y, Sugawara K, Igarashi S, Harada S, Masuda T, Sugiyama N, Togashi T, Hasegawa M, Takai Y, Yugi K, Arakawa K, Iwata N, Toya Y, Nakayama Y, Nishioka T, Shimizu K, Mori H and Tomita M. (*These members contributed equally to this work.) This study was supported by grants from the New Energy and Industrial Technology Development Organization (NEDO) of the Ministry of Economy, Trade, and Industry of Japan (METI) (Development of a Technological Infrastructure for Industrial Bioprocess Project), the Core Research for Evolutional Science and Technology (CREST) of the Japan Science and Technology Agency (JST) (Development of Modeling/Simulation Environment for Systems Biology), the Japan Society for the Promotion of Science (JSPS), and a "Grant-in-Aid" from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) for the 21st Century Center of Excellence (COE) Program (Understanding and Control of Life's Function via Systems Biology), Scientific Research on Priority Areas "Systems Genomics," and Scientific Research on Priority Areas "Lifesurveyor", as well as funds from the Yamagata Prefectural Government and Tsuruoka City. We thank the faculty and students of the Systems Biology Program at Keio University for useful discussions, as well as members of the International *E. coli* Alliance (IECA).

References

Aiba S, Matsuoka M (1979) Identification of metabolic model: Citrate production from glucose by *Candida lipolytica*. Biotechnol Bioeng 21(8):1373–86

Andersen MR, Nielsen ML, Nielsen J (2008) Metabolic model integration of the bibliome, genome, metabolome and reactome of *Aspergillus niger*. Mol Syst Biol 4:178

Arakawa K, Kono N, Yamada Y et al. (2005) KEGG-based pathway visualization tool for complex omics data. In Silico Biol 5(4):419–23

Baba T, Ara T, Hasegawa M et al. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2:2006 0008 doi:10.1038/msb4100050

Becker D, Selbach M, Rollenhagen C et al. (2006) Robust *Salmonella* metabolism limits possibilities for new antimicrobials. Nature 440(7082):303–7

Bore E, Hebraud M, Chafsey I et al. (2007) Adapted tolerance to benzalkonium chloride in *Escherichia coli* K-12 studied by transcriptome and proteome analyses. Microbiology 153(Pt 4):935–46

Bruggeman FJ, Westerhoff HV (2007) The nature of systems biology. Trends Microbiol 15(1): 45–50

Caspi R, Foerster H, Fulcher CA et al. (2008) The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. Nucleic Acids Res 36(Database issue):D623–31

Chen C, Gonzalez FJ, Idle JR (2007) LC-MS-based metabolomics in drug metabolism. Drug Metab Rev 39(2–3):581–97

Costenoble R, Muller D, Barl T et al. (2007) ^{13}C-Labeled metabolic flux analysis of a fed-batch culture of elutriated *Saccharomyces cerevisiae*. FEMS Yeast Res 7(4):511–26

Daldal F (1983) Molecular cloning of the gene for phosphofructokinase-2 of *Escherichia coli* and the nature of a mutation, *pfkB1*, causing a high level of the enzyme. J Mol Biol 168(2):285–305

De Keersmaecker SC, Thijs IM, Vanderleyden J et al. (2006) Integration of omics data: how well does it work for bacteria? Mol Microbiol 62(5):1239–50

Dettmer K, Aronov PA, Hammock BD (2007) Mass spectrometry-based metabolomics. Mass Spectrom Rev 26(1):51–78

Durrschmid K, Reischer H, Schmidt-Heck W et al. (2008) Monitoring of transcriptome and proteome profiles to investigate the cellular response of *E. coli* towards recombinant protein expression under defined chemostat conditions. J Biotechnol 135(1):34–44

Fiehn O (2002) Metabolomics–the link between genotypes and phenotypes. Plant Mol Biol 48(1–2):155–71

Fiehn O, Kopka J, Dormann P et al. (2000) Metabolite profiling for plant functional genomics. Nat Biotechnol 18(11):1157–61

Fong SS, Nanchen A, Palsson BO et al. (2006) Latent pathway activation and increased pathway capacity enable *Escherichia coli* adaptation to loss of key metabolic enzymes. J Biol Chem 281(12):8024–33

Gaspar A, Englmann M, Fekete A et al. (2008) Trends in CE-MS 2005–2006. Electrophoresis 29(1):66–79

Gibon Y, Usadel B, Blaesing OE et al. (2006) Integration of metabolite with transcript and enzyme activity profiling during diurnal cycles in *Arabidopsis rosettes*. Genome Biol 7(8):R76

Grivet JP, Delort AM, Portais JC (2003) NMR and microbiology: from physiology to metabolomics. Biochimie 85(9):823–40

Ishii N, Nakahigashi K, Baba T et al. (2007) Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. Science 316(5824):593–97

Jordan KW, Cheng LL (2007) NMR-based metabolomics approach to target biomarkers for human prostate cancer. Expert Rev Proteomics 4(3):389–400

Joyce AR, Palsson BO (2006) The model organism as a system: integrating 'omics' data sets. Nat Rev Mol Cell Biol 7(3):198–210

Kanehisa M, Araki M, Goto S et al. (2008) KEGG for linking genomes to life and the environment. Nucleic Acids Res 36(Database issue):D480-84

Kell DB (2004) Metabolomics and systems biology: making sense of the soup. Curr Opin Microbiol 7(3):296–307

Kitano H (2004) Biological robustness. Nat Rev Genet 5(11):826–37

Kitano H (2007) Towards a theory of biological robustness. Mol Syst Biol 3:137

Lee SY, Lee DY, Kim TY (2005) Systems biotechnology for strain improvement. Trends Biotechnol 23(7):349–58

Marouga R, David S, Hawkins E (2005) The development of the DIGE system: 2D fluorescence difference gel analysis technology. Anal Bioanal Chem 382(3):669–78

Mashego MR, Rumbold K, De Mey M et al. (2007) Microbial metabolomics: past, present and future methodologies. Biotechnol Lett 29(1):1–16

Monton MR, Soga T (2007) Metabolome analysis by capillary electrophoresis-mass spectrometry. J Chromatogr A 1168(1–2):237–46; discussion 236

Noh K, Gronke K, Luo B et al. (2007) Metabolic flux analysis at ultra short time scale: Isotopically non-stationary ^{13}C labeling experiments. J Biotechnol 129(2):249–67

Noh K, Wahl A, Wiechert W (2006) Computational tools for isotopically instationary ^{13}C labeling experiments under metabolic steady state conditions. Metab Eng 8(6):554–77

Oldiges M, Lutz S, Pflug S et al. (2007) Metabolomics: current state and evolving methodologies and tools. Appl Microbiol Biotechnol 76(3):495–511

Paley SM, Karp PD (2006) The Pathway Tools cellular overview diagram and Omics Viewer. Nucleic Acids Res 34(13):3771–8

Rabinowitz JD (2007) Cellular metabolomics of *Escherchia coli*. Expert Rev Proteomics 4(2): 187–98

Sanchez DH, Siahpoosh MR, Roessner U et al. (2008) Plant metabolomics reveals conserved and divergent metabolic responses to salinity. Physiol Plant 132(2):209–19

Sanford K, Soucaille P, Whited G et al. (2002) Genomics to fluxomics and physiomics - pathway engineering. Curr Opin Microbiol 5(3):318–22

Sauer U (2006) Metabolic networks in motion: ^{13}C-based flux analysis. Mol Syst Biol 2:62

Sauer U, Lasko DR, Fiaux J et al. (1999) Metabolic flux ratio analysis of genetic and environmental modulations of *Escherichia coli* central carbon metabolism. J Bacteriol 181(21):6679–88

Schaub J, Mauch K, Reuss M (2008) Metabolic flux analysis in *Escherichia coli* by integrating isotopic dynamic and isotopic stationary ^{13}C labeling data. Biotechnol Bioeng 99(5):1170–85

Shimizu K (2004) Metabolic flux analysis based on ^{13}C-labeling experiments and integration of the information with gene and protein expression patterns. Adv Biochem Eng Biotechnol 91:1–49

Skinner AJ, Cooper RA (1974) Genetic studies on ribose 5-phosphate isomerase mutants of *Escherichia coli* K-12. J Bacteriol 118(3):1183–85

Sniehotta M, Schiffer E, Zurbig P et al. (2007) CE - a multifunctional application for clinical diagnosis. Electrophoresis 28(9):1407–17

Soga T, Baran R, Suematsu M et al. (2006) Differential metabolomics reveals ophthalmic acid as an oxidative stress biomarker indicating hepatic glutathione consumption. J Biol Chem 281(24):16768–76

Soga T, Ohashi Y, Ueno Y et al. (2003) Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. J Proteome Res 2(5):488–94

Song EJ, Babar SM, Oh E et al. (2008) CE at the omics level: towards systems biology–an update. Electrophoresis 29(1):129–42

Steinfath M, Repsilber D, Scholz M et al. (2007) Integrated data analysis for genome-wide research. Exs 97:309–29

Stephanopoulos GN, Nielsen J, Aristidou A (1998) Metabolic Engineering: Principles and Methodologies. Academic Press, San Diego

Teufel A, Krupp M, Weinmann A et al. (2006) Current bioinformatics tools in genomic biomedical research (Review). Int J Mol Med 17(6):967–73

Tolstikov VV, Fiehn O, Tanaka N (2007) Application of liquid chromatography-mass spectrometry analysis in metabolomics: reversed-phase monolithic capillary chromatography and hydrophilic chromatography coupled to electrospray ionization-mass spectrometry. Methods Mol Biol 358:141–55

Toya Y, Ishii N, Hirasawa T et al. (2007) Direct measurement of isotopomer of intracellular metabolites using capillary electrophoresis time-of-flight mass spectrometry for efficient metabolic flux analysis. J Chromatogr A 1159(1–2):134–41

Toyoda T, Mochizuki Y, Player K et al. (2007) OmicBrowse: a browser of multidimensional omics annotations. Bioinformatics 23(4):524–6

Toyoda T, Wada A (2004) Omic space: coordinate-based integration and analysis of genomic phenomic interactions. Bioinformatics 20(11):1759–65

van Winden WA, van Dam JC, Ras C et al. (2005) Metabolic-flux analysis of *Saccharomyces cerevisiae* CEN.PK113–7D based on mass isotopomer measurements of ^{13}C-labeled primary metabolites. FEMS Yeast Res 5(6–7):559–68

Wang QZ, Wu CY, Chen T et al. (2006) Integrating metabolomics into a systems biology framework to exploit metabolic complexity: strategies and applications in microorganisms. Appl Microbiol Biotechnol 70(2):151–61

Ward JL, Baker JM, Beale MH (2007) Recent applications of NMR spectroscopy in plant metabolomics. Febs J 274(5):1126–31

Wiechert W (2001) ^{13}C metabolic flux analysis. Metab Eng 3(3):195–206

Wiechert W, Noh K (2005) From stationary to instationary metabolic flux analysis. Adv Biochem Eng Biotechnol 92:145–72

Wiechert W, Schweissgut O, Takanaga H et al. (2007) Fluxomics: mass spectrometry versus quantitative imaging. Curr Opin Plant Biol 10(3):323–30

Wittig U, De Beuckelaer A (2001) Analysis and comparison of metabolic pathway databases. Brief Bioinform 2(2):126–42

Wittmann C, Weber J, Betiku E et al. (2007) Response of fluxome and metabolome to temperature-induced recombinant protein synthesis in *Escherichia coli*. J Biotechnol 132(4):375–84

Yadav SP (2007) The wholeness in suffix -omics, -omes, and the word om. J Biomol Tech 18(5):277

Yoon SH, Han MJ, Lee SY et al. (2003) Combined transcriptome and proteome analysis of *Escherichia coli* during high cell density culture. Biotechnol Bioeng 81(7):753–67

Zhao J, Baba T, Mori H et al. (2004) Effect of *zwf* gene knockout on the metabolism of *Escherichia coli* grown on glucose or acetate. Metab Eng 6(2):164–74