Reconstruction of a generic genome-scale metabolic network for chicken: investigating network connectivity and finding potential biomarkers

Ehsan Salehabadi[1], Ehsan Motamedian[1], Seyed Abbas Shojaosadati[1,*]

[1]Biotechnology Group, Department of Chemical Engineering, Tarbiat Modares University, Tehran, Iran

Correspondence: Seyed Abbas Shojaosadati, Department of Biotechnology, Faculty of Chemical Engineering, Tarbiat Modares University, P.O. Box 14115-143, Tehran, Iran.

E-mail: shoja_sa@modares.ac.ir

Abstract

Chicken is the first sequenced avian that has a crucial role in human life for its meat and egg production. Because of various metabolic disorders, study the metabolism of chicken cell is important. Herein, the first genome-scale metabolic model of a chicken cell named iES1300, consists of 2427 reactions, 2569 metabolites, and 1300 genes, was reconstructed manually based on databases. Interactions of metabolic genes for growth were examined for *E. coli*, *S. cerevisiae*, human, and chicken metabolic models. The results indicated robustness to genetic manipulation for iES1300 similar to the results for human. iES1300 was integrated with transcriptomics data using algorithms and Principal Component Analysis was applied to compare context-specific models of the normal, tumor, lean and fat cell lines. It was found that the normal model has notable metabolic flexibility in the utilization of various metabolic pathways, especially in metabolic pathways of the carbohydrate metabolism, compared to the others. It was also concluded that the fat and tumor models have similar growth metabolisms and the lean chicken model has a more active lipid and carbohydrate metabolism.

Keywords: Genome-scale metabolic model, Chicken cell, Flux balance analysis, Transcriptomics data integration, Single and double gene deletion analyses, PCA

1. Introduction

Metabolism is an important cellular process in a living cell. Thus, a deep understanding of metabolic networks is required [1]. Collected biological data about metabolic pathways has led us to reconstruct a genome-scale metabolic network that can be mathematically represented [2]. Constraint-based metabolic models are known to be structured models that consider a cell a multi-component system and contain detailed intracellular process information; while because of the black box nature of the models for many cases of unsupervised learning methods in deep-learning, further processing may be required to interpret the biological meaning of the model [3]. Therefore, metabolic models will be able to predict the conditions imposed on the cell more reliably [4]. This approach will bring the model prediction as close to reality as possible by considering constraints on the upper bounds of the metabolic reaction fluxes [5]. Besides, metabolic models gain higher predictive power by integration with omics data, which in this respect, their prediction will be more valid than other data-driven models [6] . In recent years, genome-scale metabolic models (GEMs) have been increasingly developed due to the advances in genome sequencing and annotation techniques [7,8]. GEMs build a bridge between genotypic data and phenotypic traits [9].

The importance of GEMs was strengthened when it was reported that manually and automated metabolic models have been submitted for more than 6200 organisms worldwide. More than 200 of these models belong to eukaryotes [10]. Following the reconstruction of the *saccharomyces cerevisiae* GEM in 2003 as the first eukaryotic model [11], the reconstruction of eukaryotic models became prevalent so that in 2007, the first human GEM RECON1 was reconstructed [2], and human GEMs continued to be updated with the expansion of the network [12], improvement of lipid metabolism [13], energy metabolism [14], and structural information [15]. Other studies in this area include mouse model reconstruction, which is

known as the first attempt to reconstruct a mammalian model based on genomic data [16]. The mouse model was also updated in subsequent years using a human model [17]. Chinese Hamster Ovary (CHO) cells are the other interesting eukaryotic models that have been recently reconstructed because of their extensive applications in the biopharmaceutical industries [18,19].While metabolic networks has been reconstructed for most of the important mammalians in the human life, no attempt has been yet made to reconstruct a metabolic network for chickens as an important source of food. Chickens are important eukaryotes because of their large population in animal husbandries and rural life as well as the annually high consumption of their egg and meat [20]. Studies on chicken first began in 1628 with an investigation on the functions of its arteries and veins. Then, chicken genetic research improved when the chicken's first genetic map was constructed in 1936 [21]. Finally, in 2004, the genome sequencing of chicken wherein scientists had estimated 20000 to 23000 genes for chicken was released [22]. Since the last decades, genetic engineering has greatly improved, especially in the field of chickens growth and feed efficiency [23]; however, these genetic improvements have led to intensifying metabolic disorders [23]. In chickens, there are various metabolic disorders including those with environmental origins, such as oxygen or light regime, feeding strategy, as well as growth-related causes such as extraordinary growth. Some of these common disorders are fatal [24]. Thus, the development of a comprehensive metabolic model can be a platform to study the metabolism of chicken. Furthermore, such model can guide us to the treatment and even prevention of various diseases in the chicken.

In this study, for the first time, a comprehensive genome-scale metabolic reconstruction for a chicken cell (named iES1300) was reconstructed. Flux balance analysis (FBA) [25], and single gene deletion as well as double gene deletion analyses were applied to compare robustness of iES1300 and three other important models for growth. Furthermore, transcriptomics data were integrated with iES1300 to construct four types of chicken cell lines, including fat, lean,

normal, and tumor. The models were compared to determine essential metabolic differences for growth.

2. Material and Methods

2.1 The procedure of genome-scale model reconstruction for *gallus gallus*

Fig 1. Schematic representation of (a) step-by-step genome-scale metabolic reconstruction and (b) using transcriptomics data extracted by online gene expression databases (c) to achieve context-specific models by the integration algorithms.

Fig 1 illustrates the overall procedure of iES1300 reconstruction and building specific models. Using the annotation of genome sequencing for *Gallus gallus* [22], a draft model was generated based on the reconstruction protocol [26]. For this purpose, we applied the KEGG database [27] and a pathway-by-pathway analysis of *Gallus gallus* metabolism. The draft consists of all metabolic reactions and their corresponding genes, enzymes, and metabolites collected from the KEGG, BiGG [28] and CHEBI [29] databases. Reaction and metabolite abbreviations were also extracted from BiGG. If there were not any reactions or metabolite names, we added new names. Mass and charge balance as well as the reversibility of each reaction were also performed. based on the literature, the intracellular pH of 7.2 was considered for charge balance [30]. Moreover, gene-to-reaction association information was extracted from the related literature and gene orthology was obtained from close organisms. Subcellular location information was taken from UNIPROT [31]. Information from CELLO [32] and EukmPLoc v2.0 [33] was also used for the prediction of cellular location when the localization information was not available in UNIPROT. These two databases use the amino acid sequence

of the expressed protein in the FASTA format. CELLO can also use gene nucleotide sequences to predict the location of metabolic reactions. To find transport reactions, information from the REACTOME database was used [34]. Further, regarding lack of data for growth-associated maintenance (GAM) and some of the metabolites related to glycogen and lipid contents that appeared in the biomass formation reaction, these data were selected from the CHO model [18]. Furthermore, the amount of non-growth associated maintenance (NGAM) was taken from a mouse cell [35]. Also, for the biomass reaction generation, the existing information for chicken, including amino acid percentage and DNA components percentage was used. In addition, amino acid coefficients information was taken from the High-performance Integrated Virtual Environment (HIVE) database [36] and nucleotide information of the chicken genome data in NCBI [37]. More details about biomass reaction are available in the S2 file. The gap-filling process was also established so that the added reactions made the model capable of growing.

2.2 *In silico* simulations condition

To solve linear programming problems, COBRA toolbox in MATLAB 2017b software and the glpk solver package were utilized [38]. For intracellular reversible reactions, lower and upper bounds were set at -1000 and 1000$\frac{mmol}{gDCW.h}$, respectively. Contrarily, for intracellular irreversible reactions, lower and upper bounds were set at 0 and 1000 $\frac{mmol}{gDCW.h}$, respectively. The upper bound of all the exchange reactions was set at 1000 $\frac{mmol}{gDCW.h}$. An RPMI-like culture medium was also selected for the simulation of the medium, and the lower bounds of exchange reactions were fixed based on [39]. Detailed components of the RPMI-like culture medium are

represented in the S2 file. Besides, the biomass reaction was selected as the objective function in all of the simulations.

2.3 Comparison of the network

To date, a large number of studies have used different methods of network topology analysis to evaluate metabolic networks in terms of phylogenetic relationships[40]. Therefore, this analysis of iES1300 to see how phylogenetically close it is to its peers was applied. We performed the single- and double-gene deletion analyses to compare iES1300 with the metabolic models of human (RECON1) [2], *Saccharomyces cerevisiae* (iMM904) [41], and *Escherichia coli* (iJO1366) [42] to evaluate all metabolic networks in terms of phylogenetic relationships[40]. These analyses indicated that iES1300 was more flexible and robust compared to other prokaryotic and eukaryotic models. GR ratio (predicted growth rate after gene deletion per growth rate for wild type) was also applied to indicate the efficacy of single- or double-gene deletion. According to the method presented by [43], we determined sick and lethal genes based on the single-gene deletion analysis, while the interactions of genes, synthetic lethal, and synthetic sick genes were specified based on the double-gene deletion analysis. Finally, to compare the four models, the results of the single-gene deletions and the number of interactions were normalized by dividing the number of genes in each model, whereas the results of the double-gene deletions were divided by the square of the number of genes in each model.

2.4 Integration of gene expression data

Considering HCC as an important disorder in the chickens liver [44], and the negative effects of adiposity on the economics of the poultry industries, especially in the case of meat quality [45], gene expression data from four species of chicken, including chicken liver control

samples of hepatocellular carcinoma [46] , normal cell lines [47], and adipose tissue samples of lean and fat [48] were taken from ArrayExpress [49]. In our previous research[50], we indicated that TRFBA[51] and GIMME[52] are successful algorithms, especially for prediction of growth. So, both algorithms were applied to generate high-quality context-specific models. First, GIMME was used to remove reactions supported by genes with low expression levels. Then, TRFBA was employed to constrain the upper bound of the remaining reactions in the model according to the expression level of their supporting genes. In fact, TRFBA first converted all of the reversible reactions of a metabolic model into irreversible and "withoutOR". Next, it added a set of constraints to limit the rate of reactions [51] as follows:

$$\sum i \in K_j \, v_i \leq E_j \times C \quad (1)$$

Where v_i is the reaction flux of i, E_j is the expression of the gene j, K_j is the set of indices of reactions supported by metabolic gene j, and C is a constant parameter that converts the expression levels to the upper bounds of the model reactions. This coefficient indicates the maximum rate supported by one unit of expression level of a gene; thus, the unit for C is mmol $gDCW^{-1}$ h^{-1}.

A threshold of 0.25 and cutoff of 0.9 were also used for GIMME, and the parameter of TRFBA (C) for each cell line was changed in a stepwise approach according to the method presented in the next section.

2.5 Differentiation of cell lines using principal component analysis

After applying GIMME, the stepwise TRFBA was employed by stepwise change in C similar to the method presented by [50]. C was changed from zero to C_{brk} with a step size of 0.1 of C_{brk}, hence, nine flux distributions were constructed for each cell line. C_{brk} is the point at which the growth rate does not change with an increase in the value of C [50]. In fact, TRFBA was used to maximize the growth rate for each cell. On the other hand, to avoid the well-known

degeneracy of solutions, the Manhattan norm of the flux distribution was minimized while the optimal growth rate was given as constraint [53]. Correlated reactions with growth were determined for each cell line by calculating the Pearson correlation coefficient between each reaction flux and growth rate so that reactions with a coefficient more than 0.9 (P-value $\leq$ 0.05) were considered correlated. In the next step, the common growth-correlated reactions for the four cell lines were selected, and PCA was performed to differentiate the cell lines using the selected reactions.

3. Results and Discussions

3.1 Characteristics of the reconstructed model

The reconstructed model contained 2427 biochemical reactions from 95 metabolic subsystems, 1300 genes, and 2569 metabolites. Of these reactions, 1910 reactions were gene-associated reactions and 295 of them were non-gene-associated. The remaining reactions were pseudo reactions, which were mostly the exchange reactions. During the gap-filling process, 67 reactions were added to the model to make it capable of growing. Fig 2a categorizes the reactions of iES1300 into nine main subsystems. As shown in this figure, among these subsystems, lipid and energy metabolism have the largest and the smallest distributions, respectively. Fig 2b, on the other side, determines that iES1300 consists of 10 subcellular locations, named cytosol, mitochondrion, extracellular space, endoplasmic reticulum, Golgi apparatus, lysosome, peroxisome, cytosolic membrane, endosome, and nucleus. This figure also shows that cytosol has the largest metabolite distribution. In Fig 2c, we can see similar as well as different reactions of iES1300 that compared with the two important mammalian metabolic GEMs. Fig 2c indicates that iES1300 has a relatively high similarity with the Recon2v4 and iCHOv1 networks. The new reactions of iES1300 may refer to the HMR reactions existing in the Recon3D reaction list used in iES1300, spontaneous reactions and also

the reactions with different subcellular locations. Moreover, comparison of gene-associated reactions of iES1300 with three GEMs (Fig 2d) indicated that iES1300 had a lower ratio of non-gene-associated reactions to the total number of reactions of each model. Pseudo reactions are also referred to the exchange, demand, sink, biomass, and ATP maintenance reactions.

Fig 2. **Characteristics of the Reconstructed Model.** (a) In iES1300, there are 9 main metabolism categories of which the lipid metabolism is responsible for the largest reaction distribution. (b) Cytosol is considered to have the largest metabolite distribution. (c) Investigating the number of reactions shared by three important mammalian models. The newly added reactions in iES1300 in comparison to the two other models can be discovered. (d) Examination of the gene-to-reaction association's ratios (To the total number of reactions of each model) in the four important mammalian models exhibits that iES1300 has much lower orphan reactions compared to the other models.

3.2 Comparison of iES1300 with the other eukaryotic and prokaryotic models

Three different types of organism cells were chosen to be compared with iES1300 using the single- and double-gene deletion analyses. Table 1 presents that the effect of single- and double-gene deletions on growth of multicellular organisms are much lower than those of unicellular organisms. The lower GR ratio of genes with interactions in two multicellular species also showed that compared with unicellular models, multicellular ones had more robustness and flexibility to genetic perturbations. In addition, the analyses confirmed that the trait of iES1300 was similar to the eukaryotic models.

Table 1. Results of single and double gene deletion analysis on four metabolic models using FBA approach

Cell lines	Number of Genes	Ratios		
		Growth-Related Genes	Growth-Related Double-Genes	Genes with Interaction
iES1300	1300	0.05	0.0007	0.09
RECON1	1905	0.056	0.0002	0.07
iMM904	905	0.16	0.0014	0.016
iJO1366	1367	0.21	0.0013	0.15

3.3 Evaluation of metabolic similarities for cell lines

Growth vs. C/C_{brk} for each model is presented in Fig 3. In this figure, we can see that the growth patterns of fat and tumor cell models are nearly analogous. In addition, metabolisms of lean model is more susceptible to the change of C compared to normal.

Fig 3. The different patterns of the growth sensitivity to the normalized parameter of TRFBA algorithm (C/C_{brk}), are a source for differentiation of the four models.

Furthermore, by applying PCA for the common growth-related reactions, it was found that the normal and lean models were significantly different from tumor and fat models (Fig 4a). The PCA results also indicated that the normal chicken cell was the most different cell line from the other cell lines. Fig 4b illustrates that by using the first principal component, 85% of the difference between normal cell metabolism and other cell lines can be explained. We further presented the difference of lean chicken cell with other cell lines by the second principal

component with approximately 15% variance. It can be seen in Fig 4a that the fat model has metabolic similarities to tumor model. Therefore, we can state that the normal model had key metabolic differences from lean, fat, and tumor chick cell metabolisms. Focusing on these differences and targeting them can prevent such metabolic disorders in chickens. Results of PCA also verified that a considerable number of reactions that shifted along the PC2 were orthogonal to the reactions that appear in PC1, indicating that the lean and the normal models have presumably separate growth mechanisms.

Fig 4. PCA for extracting the reactions that play the role of biomarkers. (a) PCA shows that the 36 flux distributions created by the combination of two integration algorithms, GIMME[52], and stepwise TRFBA[51] are successfully categorized into four different groups so that each group belongs to one cell line. (b) The variance explanation chart indicates that the 1st principal component is responsible for about 85% of the variance. The other 15% can be explained by the 2nd principal component. (c) The main differentiated reactions in PCA. 28 reactions are shown to have participated in the differentiation of the fat, lean, normal and tumor cells. Other reactions with lower absolute PC values were not investigated. The complete list of the reaction names is available in the S2 file.

Fig 4c also shows the 28 main reactions in the differentiation of fat, lean, tumor, and normal chicken cell lines. Reactions with low values of absolute principal components were not considered because their differentiated flux distributions were insignificant.

Besides, it is worth mentioning that the fat and tumor chicken models had strong resemblances, mostly because of their glucose, nucleotide, and lipid metabolism activities and many essential amino acid exchange reactions of these two models. These results are presented in the S2 file, where all the differentiated reactions in PCA are explained. Research works, especially in the field of obesity and Hepatocellular Carcinoma (HCC), indicated that in obesity, fat accumulation leads to liver malfunction, and consequently the liver cannot send out more triglycerides by very-low-density lipoprotein (VLDL) than that are synthesized. This

intrahepatic triglyceride increasing would, in turn, result in fatty liver, and consequently liver failure and HCC [54,55]. In the case of iES1300, lots of relationships and similarities were observed, especially in three pathways of lipid metabolism, including Fatty Acid Synthesis (FAS), sphingolipid, and glycerophospholipid Metabolisms. Similarities in the flux patterns of FAS could lead both tumor and fat models to equally produce Palmitoyl-COA, which is a key metabolite in the progression of many other lipid pathways. Sphingolipid metabolism is one of the metabolic pathways, which is affected by FAS. This metabolic pathway is known for having some bioactive metabolites involved in the regulation of cell growth [56]. Therefore, it can be one of the primary sources of similarity in both models. Tracking sphingolipid and glycerophospholipid metabolisms has also revealed that some of their major metabolites play a key role in the biomass objective function reaction. Given that the nucleotides are widely used in various functions of all cells, and because of their relation with cell proliferation to DNA replication and RNA production [57], the more balance in the nucleotide metabolism activity could result in more equivalency of nucleotide production used for biomass generation. Moreover, it has been proved that obesity can systemically impact glucose metabolism by elevating glucose and insulin level, which favors cancer cell progression [58]. It can be presumed that the similarity in these three main metabolisms (glucose, nucleotide, and lipid metabolisms) controlling cell growth and proliferation affects the similarity between the amounts of many biomass reactants produced in tumor and fat models. This similarity could lead to the identical cell growth patterns in the tumor and fat models rather than the normal and lean models.

PCA revealed that in the 1st Principal Component, which differentiated the normal chicken model from other models, especially tumor model, there were numerous distinguished reactions in three metabolisms, including nucleotide, carbohydrate, and lipid metabolism. Although number of these reactions were higher in the tumor cell metabolism than in the

normal cells, the results demonstrated that glycolysis and pentose phosphate pathways were the most critical metabolic pathways in diversities of normal and tumor chicken models. It is also noteworthy that PC1 results of the 28 leading differentiated reactions indicated that the glycolysis pathway in tumor cell metabolism had higher activity toward the production of 3-phosphoglycerate, which is a key metabolite in the development of glycine, serine, and threonine metabolism. On this account, the tumor cells are expected to be more active than a normal cell in this metabolic pathway. This assertion could be justified by knowing that serine and glycine provide the main precursors for tumor cell metabolism [59]. It has been also observed that despite the fact that most reactions happen in the tumor cell, the normal cell could produce significant amounts of ribose-5-phosphate. Further investigations have shown that in the next step, this metabolite is converted to 5-phospho-ribose 1-diphosphate, which plays an influential role in the progression of nucleotide metabolism. It is important to note that nucleotides can be synthesizable from two main metabolisms of *de novo* synthesis and salvage pathways [57]. The proliferating cells such as cancer cells are more eager to synthesize their required nucleotides through *de novo* synthesis [60]. Similarly, in the present study, iES1300 demonstrated that the preference of normal chicken cell to use recycling of its nucleosides and nucleobases through salvage pathways is much higher than tumor cell.

On the other side, for the 2nd Principal Component, which differentiated the lean cell line model from the others, especially fat model metabolism, various differentially flux distributions were observed. Results illustrated that in many metabolisms, because of several positive shifts along the PC2 axis, the metabolic activity of lean chicken can be more than that of fat chicken. However, a few of them were identified to be significant. Intriguingly, we perceived that the activity of enzyme phosphoglucomutase in the production of glucose-6-phosphate and the way it is used can be the dominant source of differences between lean and fat models. In the lean model, a considerable amount of glucose-6-phosphate heads towards inositol phosphate as well

as ascorbate and aldarate metabolisms to produce significant amounts of uridine diphosphate glucose (UDPG). Since UDPG is a key precursor in starch and sucrose metabolism, it was expected that the lean model had much more activity on this pathway. Additional investigations indicated that the increase in activity of lean model was not due to glycogen production and storage, but it resulted from the production of glucose by 4-alpha-glucanotransferase. As a result, the starch and sucrose metabolism was considered a key metabolic pathway, and the principal source of cell metabolism for glucose generation and utilization. Glucose is consumed in the lean model at a significantly higher rate, even in adipose tissue. Previous studies on the relationship between glycogen and lipid oxidation in the liver and muscles have also shown that the glycogen storage reduction could increase lipid oxidation by stimulation of cellular energy state [61,62]. Likewise, in obese adipocyte cell metabolism, the lower gene expression of fatty acid pathways because of defection in mitochondrial function resulting from a decrease in the mitochondrial acetyl-CoA concentration has already been established [13]. In this research work, iES1300 represented a significant diversity in the flux distributions of the lean and fat models, especially flux of three important reactions in cholesterol metabolism, fatty acid oxidation, and glyoxylate and dicarboxylate metabolism are significantly different. These reactions contributed to the production of mitochondrial acetyl-CoA and subsequently significant discrepancy in fatty acid oxidation metabolism was observed.

Conclusion

Chickens are the animals most associated with humans in rural life and animal husbandries, chickens are the most associated animals with humans to produce meat and egg. As a result, a metabolic model of the chicken cell will help us understand more about this organism's behavior and prevent disease outbreaks. Accordingly, to reconstruct the first genome-scale metabolic model of the chicken cell as well as the chicken biological and genomic data were

manually collected in the form of a draft from different available bioinformatics databases and used after performing the gap-filling process. The final model consisted of 2427 reactions, 2569 metabolites, and 1300 genes. The chicken model was compared with three other important models to evaluate the interaction of metabolic gene networks. This comparison demonstrated the relative similarity of the chicken's gene network to human. After model reconstruction, the transcriptomics data of the four cell types of lean, fat, normal, and tumor were integrated using the two algorithms of GIMME and TRFBA. Finally, by implementing PCA, we concluded that PCA, in addition to properly differentiating cell types from each other, has recommended important biomarkers. These biomarkers mostly participates in different metabolic pathways such as carbohydrate metabolism, to distinguish normal cells from three other cell lines.

Acknowledgements

The authors thank Tarbiat Modares University Research and Technology Unit for supporting this study.

References

1. Çakir T, Khatibipour MJ. Metabolic network discovery by top-down and bottom-up approaches and paths for reconciliation. Front Bioeng Biotechnol. 2014;2. doi:10.3389/fbioe.2014.00062

2. Duarte NC, Becker SA, Jamshidi N, Thiele I, Mo ML, Vo TD, et al. Global reconstruction of the human metabolic network based on genomic and bibliomic data. Proc Natl Acad Sci U S A. 2007;104: 1777–1782. doi:10.1073/pnas.0610772104

3. Xu C, Jackson SA. Machine learning and complex biological data. 2019; 1–4.

4. Karimian E, Motamedian E. ACBM: An Integrated Agent and Constraint Based Modeling Framework for Simulation of Microbial Communities. Sci Rep. 2020;10: 1–10. doi:10.1038/s41598-020-65659-w

5. Bordbar A, Monk JM, King ZA, Palsson BO. Constraint-based models predict metabolic and associated cellular functions. Nat Rev Genet. 2014;15: 107–120. doi:10.1038/nrg3643

6. Id GZ, Id SV, Yaneske E, Id CA. Machine and deep learning meet genome- scale metabolic modeling. 2019; 1–24.

7. Monk J, Nogales J, Palsson BO. Optimizing genome-scale network reconstructions. Nat Biotechnol. 2014;32: 447–452. doi:10.1038/nbt.2870

8. Saha R, Chowdhury A, Maranas CD. Recent advances in the reconstruction of metabolic models and integration of omics data. Curr Opin Biotechnol. 2014;29: 39–45. doi:10.1016/j.copbio.2014.02.011

9. Oberhardt MA, Palsson B, Papin JA. Applications of genome-scale metabolic

reconstructions. Mol Syst Biol. 2009;5. doi:10.1038/msb.2009.77

10. Gu C, Kim GB, Kim WJ, Kim HU, Lee SY. Current status and applications of genome-scale metabolic models. Genome Biol. 2019;20. doi:10.1186/s13059-019-1730-3

11. Förster J, Famili I, Fu P, Palsson B, Nielsen J. Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. Genome Res. 2003. doi:10.1101/gr.234503

12. Thiele I, Swainston N, Fleming RMT, Hoppe A, Sahoo S, Aurich MK, et al. A community-driven global reconstruction of human metabolism. Nat Biotechnol. 2013;31: 419–425. doi:10.1038/nbt.2488

13. Mardinoglu A, Agren R, Kampf C, Asplund A, Nookaew I, Jacobson P, et al. Integration of clinical data with a genome-scale metabolic model of the human adipocyte. Mol Syst Biol. 2013;9: 10. doi:10.1038/msb.2013.5

14. Swainston N, Smallbone K, Hefzi H, Dobson PD, Brewer J, Hanscho M, et al. Recon 2.2: from reconstruction to model of human metabolism. Metabolomics. 2016;12. doi:10.1007/s11306-016-1051-4

15. Brunk E, Sahoo S, Zielinski DC, Altunkaya A, Dräger A, Mih N, et al. Recon3D enables a three-dimensional view of gene variation in human metabolism. Nat Biotechnol. 2018;36: 272–281. doi:10.1038/nbt.4072

16. Sheikh K, Förster J, Nielsen LK. Modeling hybridoma cell metabolism using a generic genome-scale metabolic model of Mus musculus. Biotechnol Prog. 2005. doi:10.1021/bp0498138

17. Sigurdsson MI, Jamshidi N, Steingrimsson E, Thiele I, Palsson BT. A detailed genome-wide reconstruction of mouse metabolism based on human Recon 1. BMC Syst Biol.

2010;4. doi:10.1186/1752-0509-4-140

18. Hefzi H, Ang KS, Hanscho M, Bordbar A, Ruckerbauer D, Lakshmanan M, et al. A Consensus Genome-scale Reconstruction of Chinese Hamster Ovary Cell Metabolism. Cell Syst. 2016;3: 434-443.e8. doi:10.1016/j.cels.2016.10.020

19. Calmels C, McCann A, Malphettes L, Andersen MR. Application of a curated genome-scale metabolic model of CHO DG44 to an industrial fed-batch process. Metab Eng. 2019;51: 9–19. doi:10.1016/j.ymben.2018.09.009

20. Thornton PK. Livestock production: Recent trends, future prospects. Philos Trans R Soc B Biol Sci. 2010;365: 2853–2867. doi:10.1098/rstb.2010.0134

21. Stern CD. The chick: A great model system becomes even greater. Dev Cell. 2005;8: 9–17. doi:10.1016/S1534-5807(04)00425-3

22. Hillier LW, Miller W, Birney E, Warren W, Hardison RC, Ponting CP, et al. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature. 2004;432: 695–716. doi:10.1038/nature03154

23. Angel R. Metabolic disorders: Limitations to growth of and mineral deposition into the broiler skeleton after hatch and potential implications for leg problems. J Appl Poult Res. 2007;16: 138–149. doi:10.1093/japr/16.1.138

24. Julian RJ. Production and growth related disorders and other metabolic diseases of poultry - A review. Vet J. 2005;169: 350–369. doi:10.1016/j.tvjl.2004.04.015

25. Orth JD, Thiele I, Palsson BO. What is flux balance analysis? Nat Biotechnol. 2010;28: 245–248. doi:10.1038/nbt.1614

26. Thiele I, Palsson B. A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat Protoc. 2010;5: 93–121. doi:10.1038/nprot.2009.203

27. Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M. New approach for understanding genome variations in KEGG. Nucleic Acids Res. 2019;47: D590–D595. doi:10.1093/nar/gky962

28. King ZA, Lu J, Dräger A, Miller P, Federowicz S, Lerman JA, et al. BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. Nucleic Acids Res. 2016;44: D515–D522. doi:10.1093/nar/gkv1049

29. Hastings J, De Matos P, Dekker A, Ennis M, Harsha B, Kale N, et al. The ChEBI reference database and ontology for biologically relevant chemistry: Enhancements for 2013. Nucleic Acids Res. 2013;41. doi:10.1093/nar/gks1146

30. Freudenrich CC, Murphy E, Levy LA, London RE, Lieberman M. Intracellular pH modulates cytosolic free magnesium in cultured chicken heart cells. Am J Physiol - Cell Physiol. 1992;262. doi:10.1152/ajpcell.1992.262.4.c1024

31. Bateman A. UniProt: A worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47: D506–D515. doi:10.1093/nar/gky1049

32. Yu CS, Chen YC, Lu CH, Hwang JK. Prediction of protein subcellular localization. Proteins Struct Funct Genet. 2006;64: 643–651. doi:10.1002/prot.21018

33. Chou KC, Shen H Bin. A new method for predicting the subcellular localization of eukaryotic proteins with both single and multiple sites: Euk-mPLoc 2.0. PLoS One. 2010. doi:10.1371/journal.pone.0009931

34. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The Reactome Pathway Knowledgebase. Nucleic Acids Res. 2018;46: D649–D655. doi:10.1093/nar/gkx1132

35. Sinclair R. Response of mammalian cells to controlled growth rates in steady-state

continuous culture. In Vitro. 1974. doi:10.1007/BF02615311

36. Simonyan V, Chumakov K, Dingerdissen H, Faison W, Goldweber S, Golikov A, et al. High-performance Integrated Virtual Environment (HIVE): A robust infrastructure for next-generation sequence data analysis. Database. 2016;2016. doi:10.1093/database/baw022

37. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016. doi:10.1093/nar/gkv1189

38. Schellenberger J, Que R, Fleming RMT, Thiele I, Orth JD, Feist AM, et al. Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox v2.0. Nat Protoc. 2011;6: 1290–1307. doi:10.1038/nprot.2011.308

39. Folger O, Jerby L, Frezza C, Gottlieb E, Ruppin E, Shlomi T. Predicting selective drug targets in cancer through metabolic networks. Mol Syst Biol. 2011;7. doi:10.1038/msb.2011.35

40. Machicao J, Filho HA, Lahr DJG, Buckeridge M, Bruno OM. Topological assessment of metabolic networks reveals evolutionary information. Sci Rep. 2018;8: 1–11. doi:10.1038/s41598-018-34163-7

41. Mo ML, Palsson B, Herrgård MJ. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. BMC Syst Biol. 2009. doi:10.1186/1752-0509-3-37

42. Orth JD, Conrad TM, Na J, Lerman JA, Nam H, Feist AM, et al. A comprehensive genome-scale reconstruction of Escherichia coli metabolism-2011. Mol Syst Biol. 2011;7. doi:10.1038/msb.2011.65

43. Motamedian E, Ghavami G, Sardari S. Investigation on metabolism of cisplatin resistant

ovarian cancer using a genome scale metabolic model and microarray data. Iran J Basic Med Sci. 2015;18: 267–276. doi:10.22038/ijbms.2015.4131

44. Suksaweang S, Lin CM, Jiang TX, Hughes MW, Widelitz RB, Chuong CM. Morphogenesis of chicken liver: Identification of localized growth zones and the role of β-catenin/Wnt in size regulation. Dev Biol. 2004;266: 109–122. doi:10.1016/j.ydbio.2003.10.010

45. Peng M, Han J, Li L, Ma H. Suppression of fat deposition in broiler chickens by (-)-hydroxycitric acid supplementation: A proteomics perspective. Sci Rep. 2016;6. doi:10.1038/srep32580

46. Sun L, Lamont SJ, Cooksey AM, McCarthy F, Tudor CO, Vijay-Shanker K, et al. Transcriptome response to heat stress in a chicken hepatocellular carcinoma cell line. Cell Stress Chaperones. 2015;20: 939–950. doi:10.1007/s12192-015-0621-0

47. Coble DJ, Fleming D, Persia ME, Ashwell CM, Rothschild MF, Schmidt CJ, et al. RNA-seq analysis of broiler liver transcriptome reveals novel responses to high ambient temperature. BMC Genomics. 2014;15. doi:10.1186/1471-2164-15-1084

48. Wang HB, Li H, Wang QG, Zhang XY, Wang SZ, Wang YX, et al. Profiling of chicken adipose tissue gene expression by genome array. BMC Genomics. 2007;8. doi:10.1186/1471-2164-8-193

49. Kolesnikov N, Hastings E, Keays M, Melnichuk O, Tang YA, Williams E, et al. ArrayExpress update-simplifying data submissions. Nucleic Acids Res. 2015;43: D1113–D1116. doi:10.1093/nar/gku1057

50. Jamialahmadi O, Hashemi-Najafabadi S, Motamedian E, Romeo S, Bagheri F. A benchmark-driven approach to reconstruct metabolic networks for studying cancer

metabolism. PLoS Comput Biol. 2019;15. doi:10.1371/journal.pcbi.1006936

51. Motamedian E, Mohammadi M, Shojaosadati SA, Heydari M. TRFBA: An algorithm to integrate genome-scale metabolic and transcriptional regulatory networks with incorporation of expression data. Bioinformatics. 2017. doi:10.1093/bioinformatics/btw772

52. Becker SA, Palsson BO. Context-specific metabolic networks are consistent with experiments. PLoS Comput Biol. 2008;4. doi:10.1371/journal.pcbi.1000082

53. Lewis NE, Hixson KK, Conrad TM, Lerman JA, Charusanti P, Polpitiya AD, et al. Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models. Mol Syst Biol. 2010;6. doi:10.1038/msb.2010.47

54. Sun B, Karin M. Obesity, inflammation, and liver cancer. J Hepatol. 2012;56: 704–713. doi:10.1016/j.jhep.2011.09.020

55. Fabbrini E, Sullivan S, Klein S. Obesity and nonalcoholic fatty liver disease: Biochemical, metabolic, and clinical implications. Hepatology. 2010;51: 679–689. doi:10.1002/hep.23280

56. Bartke N, Hannun YA. Bioactive sphingolipids: Metabolism and function. J Lipid Res. 2009;50. doi:10.1194/jlr.R800080-JLR200

57. Lane AN, Fan TWM. Regulation of mammalian nucleotide metabolism and biosynthesis. Nucleic Acids Res. 2015;43: 2466–2485. doi:10.1093/nar/gkv047

58. Font-Burgada J, Sun B, Karin M. Obesity and Cancer: The Oil that Feeds the Flame. Cell Metab. 2016;23: 48–62. doi:10.1016/j.cmet.2015.12.015

59. Amelio I, Cutruzzolá F, Antonov A, Agostini M, Melino G. Serine and glycine metabolism in cancer. Trends Biochem Sci. 2014;39: 191–198.

512 doi:10.1016/j.tibs.2014.02.004

513 60. Villa E, Ali ES, Sahu U, Ben-Sahra I. Cancer cells tune the signaling pathways to
514 empower de novo synthesis of nucleotides. Cancers. 2019.
515 doi:10.3390/cancers11050688

516 61. Hardie DG, Scott JW, Pan DA, Hudson ER. Management of cellular energy by the
517 AMP-activated protein kinase system. FEBS Lett. 2003;546: 113–120.
518 doi:10.1016/S0014-5793(03)00560-X

519 62. Galgani J, Ravussin E. Energy metabolism, fuel selection and body weight regulation.
520 Int J Obes. 2008;32. doi:10.1038/ijo.2008.246

521

522

Supporting Information

S1 File. The Excel and SBML version of the reconstructed model. In this supplementary file, a zip file containing the Excel and SBML version of the reconstructed chicken cell model is attached.

S2 File. Detailed information of the model. In this supplementary file, detailed information in the form of Excel sheets that has been mentioned in the paper including protein and DNA components coefficients, simulated growth medium, all and the main differentiated reactions from PCA results are available.

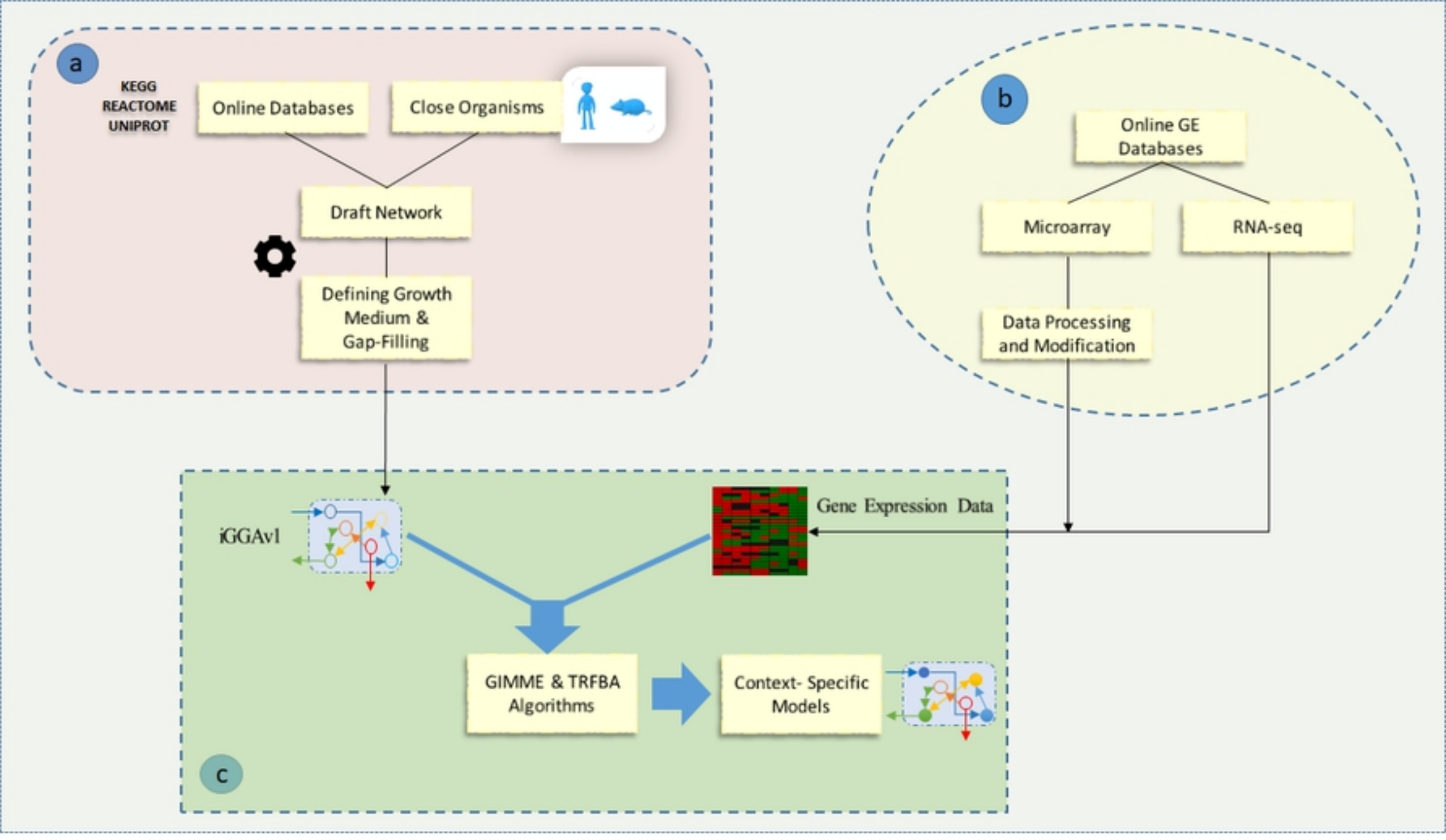
a
KEGG
REACTOME
UNIPROT
Online Databases
Close Organisms
Draft Network
Defining Growth Medium & Gap-Filling
b
Online GE Databases
Microarray
RNA-seq
Data Processing and Modification
iGGAv1
Gene Expression Data
GIMME & TRFBA Algorithms
Context- Specific Models
c

Figure 1

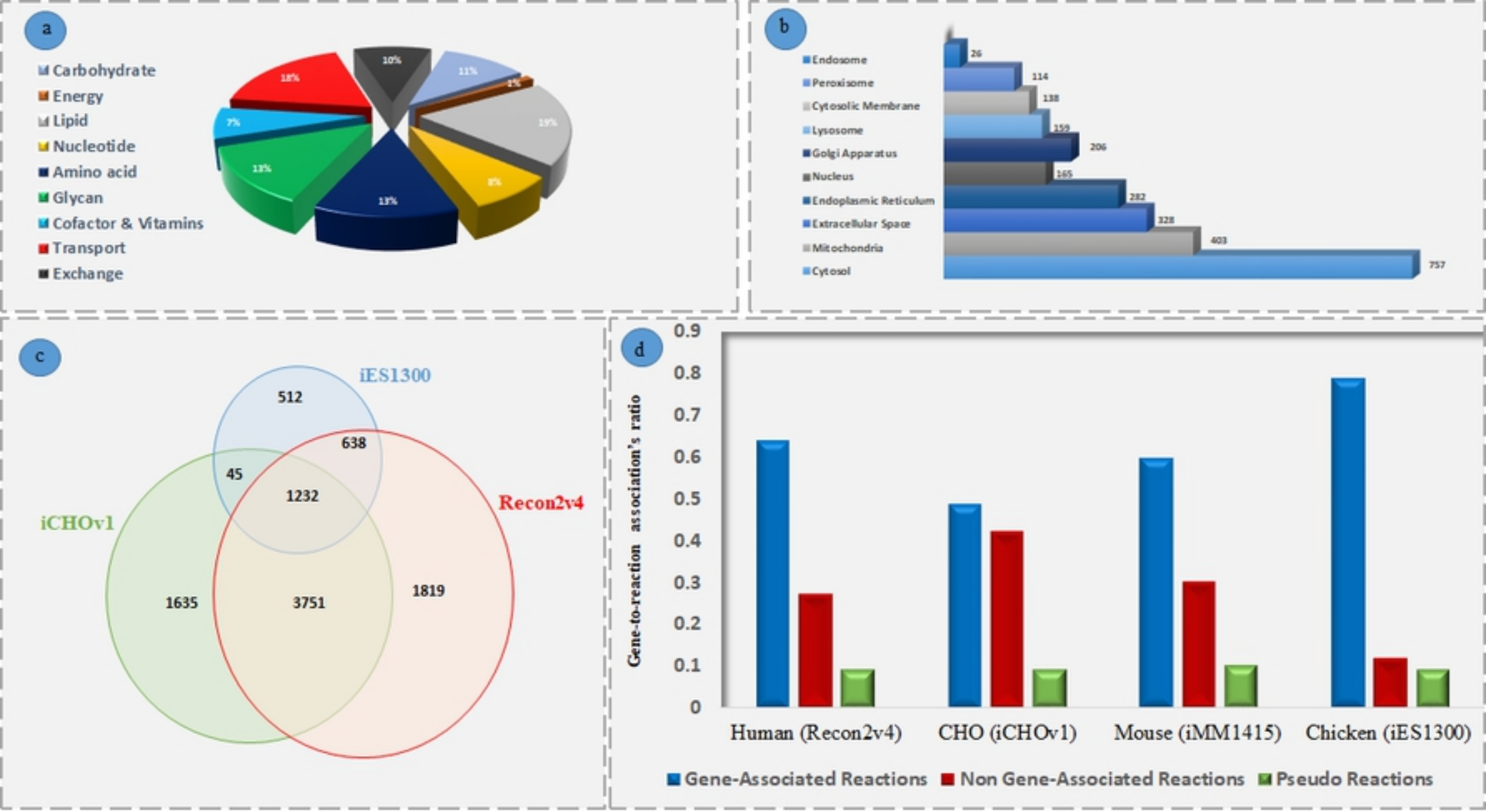
a
Carbohydrate
Energy
Lipid
Nucleotide
Amino acid
Glycan
Cofactor & Vitamins
Transport
Exchange
10%
11%
1%
19%
8%
18%
13%
7%
18%
b
Endosome
Peroxisome
Cytosolic Membrane
Lysosome
Golgi Apparatus
Nucleus
Endoplasmic Reticulum
Extracellular Space
Mitochondria
Cytosol
26
114
138
159
206
165
282
328
403
757
c
iES1300
iCHOv1
Recon2v4
512
638
45
1232
1635
3751
1819
d
Gene-to-reaction association's ratio
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Human (Recon2v4)
CHO (iCHOv1)
Mouse (iMM1415)
Chicken (iES1300)
Gene-Associated Reactions
Non Gene-Associated Reactions
Pseudo Reactions

Figure 2

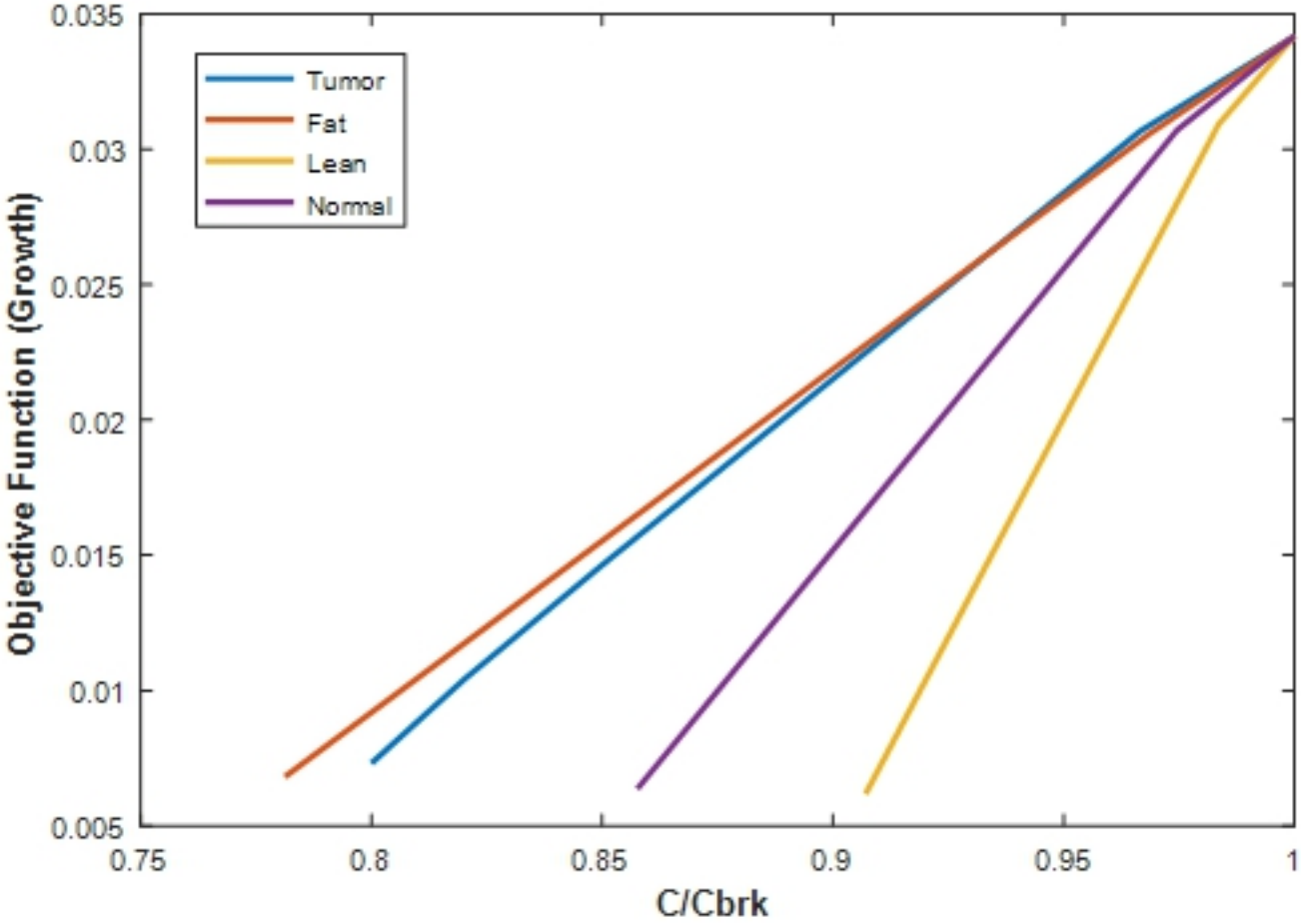
Tumor
Fat
Lean
Normal
Objective Function (Growth)
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0.75
0.8
0.85
0.9
0.95
1
C/Cbrk

Figure 3

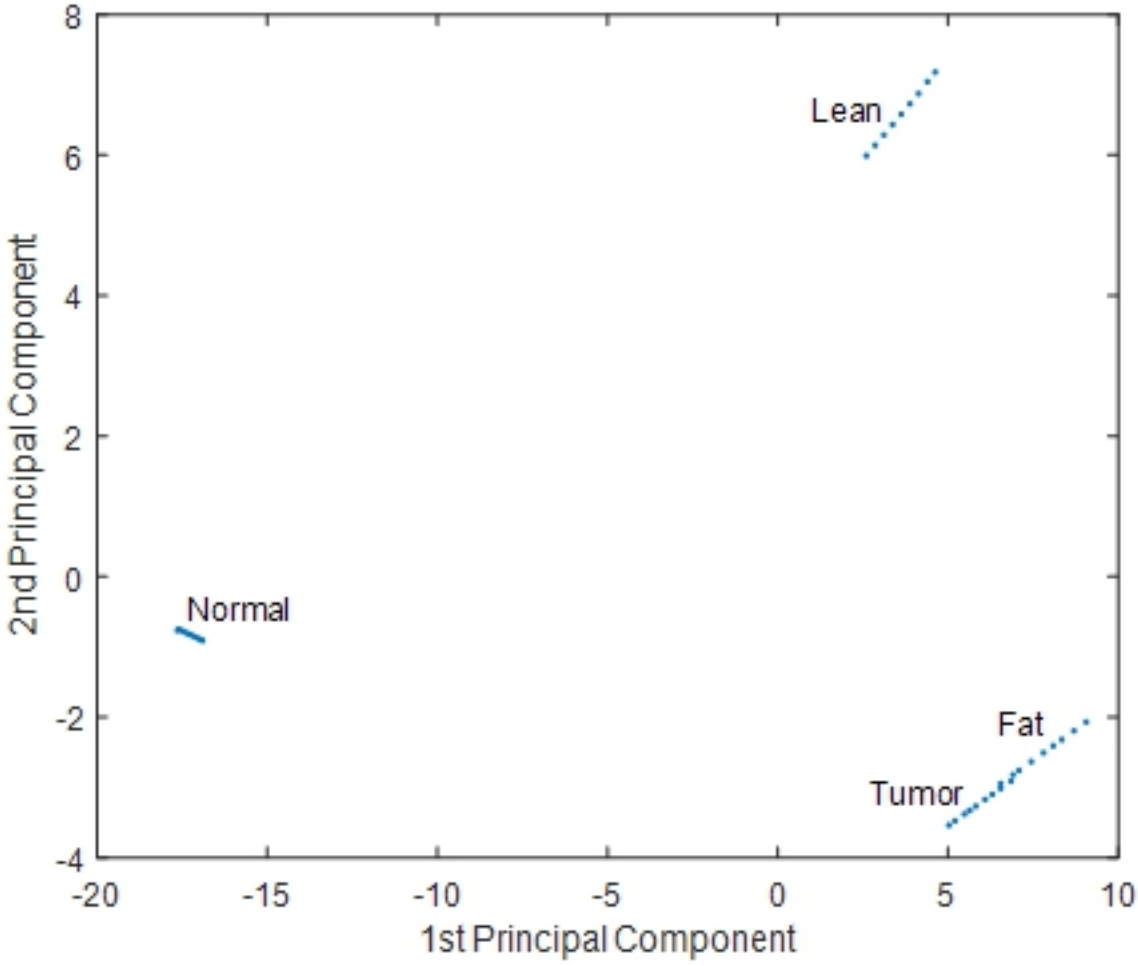
Lean
Normal
Fat
Tumor
2nd Principal Component
1st Principal Component
8
6
4
2
0
-2
-4
-20
-15
-10
-5
0
5
10

Figure 4a

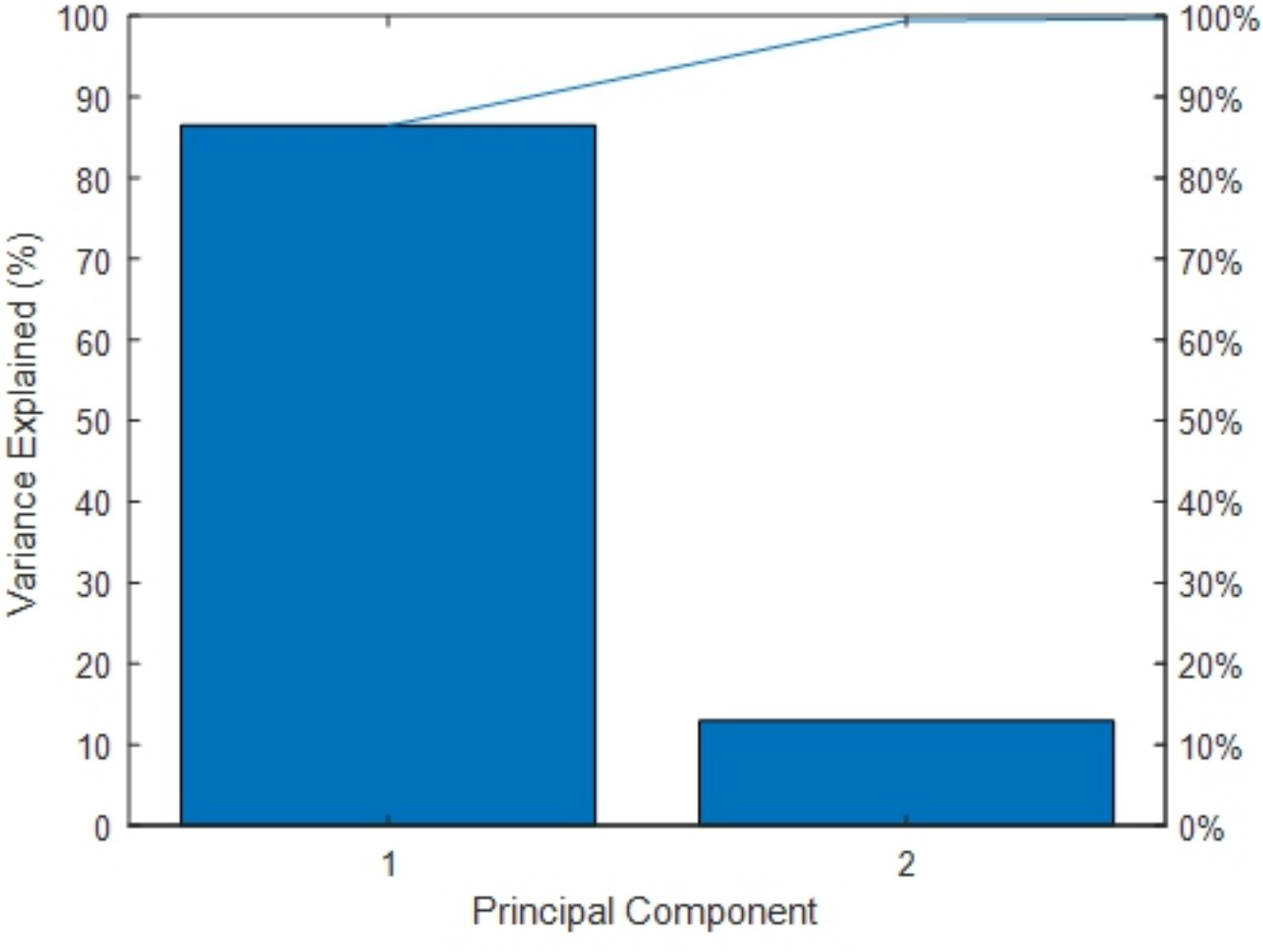
100
90
80
70
60
50
40
30
20
10
0
Variance Explained (%)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1
2
Principal Component

Figure 4b

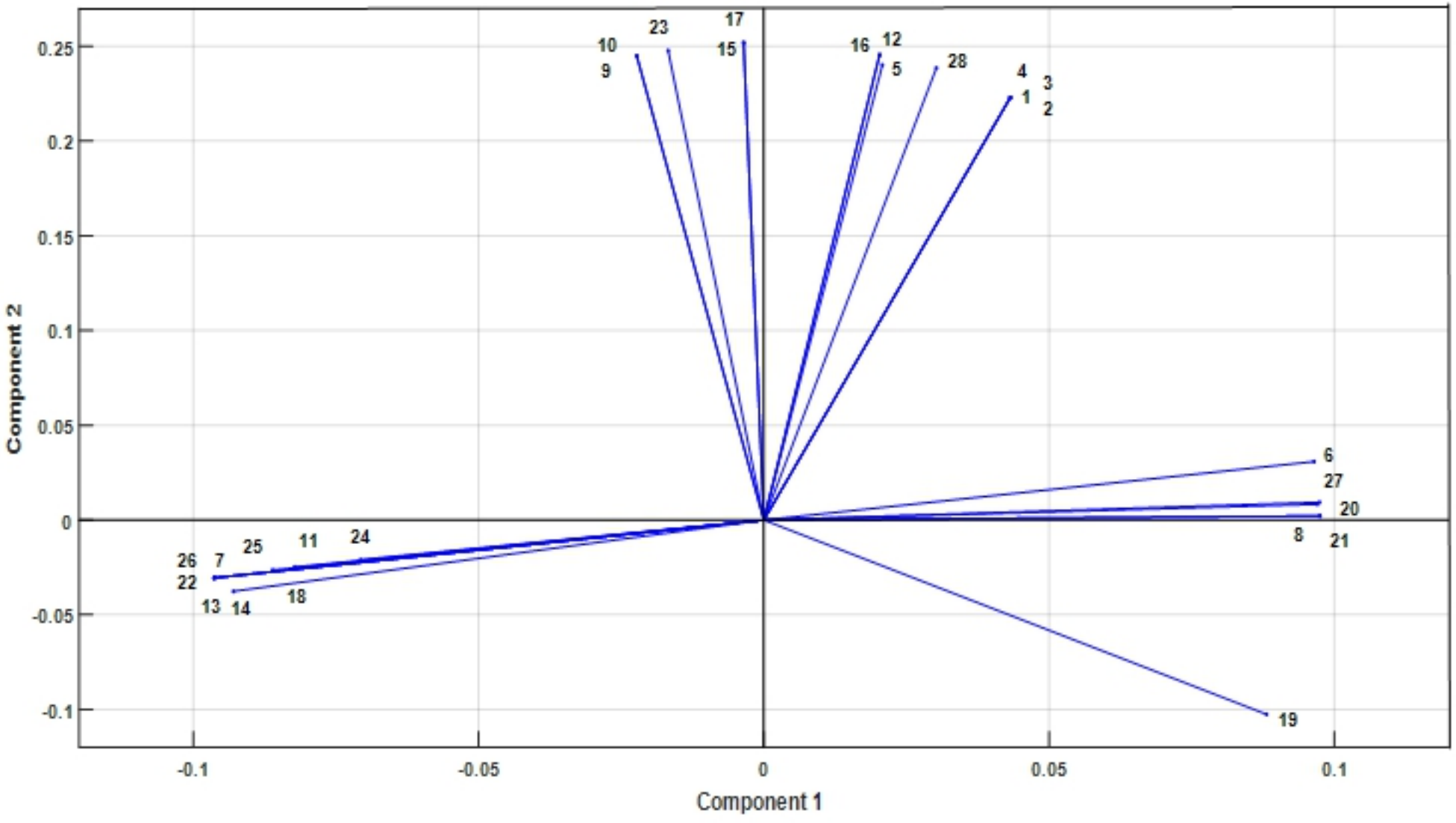
0.25
0.2
0.15
0.1
0.05
0
-0.05
-0.1
Component 2
-0.1
-0.05
0
0.05
0.1
Component 1
10
9
23
17
15
16
12
5
28
4
1
3
2
6
27
20
8
21
19
26
22
7
25
11
24
13
14
18

Figure 4c