

REVIEW ARTICLE

Opportunities for computational biology uses in enhancement of bioproduct quality: Lessons from olive oil

Karim Ennouri^{1,2*}, Rayda Ben Ayed³, Slim Smaoui²

¹Laboratory of Amelioration and Protection of Olive Genetic Resources, Olive Tree Institute, University of Sfax, Sfax, Tunisia, ²Laboratory of Microorganisms and Biomolecules of the Center of Biotechnology of Sfax-Tunisia, Road of Sidi mansour, Km 6 B.P. 1117, 3018 Sfax-Tunisia, ³Molecular and Cellular Screening Process Laboratory, Center of Biotechnology of Sfax, PB 1177, 3018 Sfax, Tunisia

ABSTRACT

Consistent advance in genetic engineering has offered ascend to various promising fields of research that encouraged the development of twice industrial biotechnology and quality production. The discipline of metabolic engineering, which uses genetic devices, has had an especially solid effect by giving new stages to industrial biotechnology. Late improvements in life sciences guarantee to extend the metabolic engineering tool compartment assist by making novel natural parts for pathway plan. Moreover, synthetic biology represents an innovative concept for learning about and via biomolecules and data, with appliances in fundamental sciences, agriculture and biotechnology. The present review tends to a portion of the current advances in biological sciences and how these can possibly influence bioproduct qualities as olive oil. A few cases of uses for modern sciences as bioinformatics and synthetic biology and metabolic engineering in life sciences and particularly in agronomy are additionally also discussed.

Keywords: Engineering; Synthetic biology; Bioinformatics; Olive oil

INTRODUCTION

Actually, olive oil biotechnology destines to modernize usual manufacturing in the years ahead, mostly owing to the brilliant evolution in our aptitude to engineering of cellular metabolism of plants (Atkinson and Mavituna, 1991). Then, most accomplishments of metabolic engineering have been employed in several industrial fields involved in olive fruit harvest and process. In this view, we highlight new tendencies in metabolic engineering and synthetic biology that illustrate the synthetic potential of molecules and pathway engineering for new product innovation (Purnick and Weiss, 2009). Indeed, the increased employ of bioinformatic devices to mine, arrange and examine biological data; and the raising sensibility and complexity of investigational devices will make the evolution of metabolic engineering and autocatalytic phenomenon. Inspired by these progresses, research teams across the world would take up the defy of secondary metabolite fabrication and modification of olive oil quality with regenerated dynamism,

thus joining to the assortment of products synthesized during transformation and storage via metabolic engineering (Lee et al., 2015; Duthie et al., 2016).

Bioprocess engineering has focused on the optimization of the production of bioproducts through biological transformations and changes. Predominantly, much importance has been positioned on evaluating and estimating the best process mode variables to enhance the production of the desired content (Ramli et al., 2014). More especially, researchers and scientists were obliged to employ detectable process variables, such as the feed supply speed of the bioreactor or as example the modification of oil physic-chemical composition in function of time. Moreover, optimization encloses the strict connexion between yield and productivity (Claes et al., 1999). Raised values of both metrics are required in manufacture and industry; however, the two metrics are in competition. Really, the appearance and the development of synthetic biology and bioinformatics also have permit biologists to expand the

*Corresponding author:

Karim Ennouri, Laboratory of Amelioration and Protection of Olive Genetic Resources, Olive Tree Institute, University of Sfax, Sfax, Tunisia.
E-mail: 1karimennouri1@gmail.com

Received: 17 April 2018; Accepted: 26 July 2018

amelioration of bioprocesses from the visible stage to the infinitesimal stage.

METABOLIC ENGINEERING: DEFINITION AND ROLE

Metabolic engineering is defined as the engineering of biologic systems in the aim to enhance the production of a specific product. Among the raised understanding of metabolic and usual networks, reasonable and logical approaches to develop the production of a desired element have been well developed (Ennouri *et al.*, 2013a). The final product can either be naturally produced by the studied organism, or artificially integrated to the host from another organism via genetic engineering. In this last case, one or more genes required for the reaction pathway are transformed and changed in the host.

In metabolic engineering, biological networks are deeply managed to enhance the production of a desired product. The most essential principle concerns varying the metabolic pathways in the aim to assure that the resources and metabolic mediators are provided at the suitable levels to boost product yield. Products naturally generated by living cells and comparatively basic products necessitate a few manipulations to increase yield (Ennouri *et al.*, 2013a, 2016). In most cases, removal of specific pathways or over expression of particular proteins is sufficient to attain yield targets. More complex products entail amelioration and junction of more pathways. If major cellular antecedents are consumed or toxic intermediators enhance, the consequence can be harmful and dangerous to the cells. The multifaceted character of metabolic networks implies that precise management of biological function is indispensable for system-level improvement.

Apart from yield, efficiency has significance on the economical possibility of large-scale manufacture. For limited quantities, high-cost ingredients, the emphasis is on enhancing the product yield. In contrast, the regulation between productivity and yield is crucial and decisive for the possibility of a higher volume, and low cost (bulk) production procedure. Productivity becomes a key feature in these cases since the profit division is restricted.

Metabolic engineering has advantaged from progress and advances in synthetic biology, and enlarged the development and evolution of bioprocesses at the experimental stage. Synthetic biology is a device that joins theories from engineering and biology, and makes simple the re-engineering of metabolic networks. The purpose of synthetic biology consists to increase well-characterized biological parts that can be created for an organism, and carry-out an unusual task. The raising quantity of accessible

parts has enlarged the tool set for control and efficiency of metabolic engineering requests and applications.

SYNTHETIC BIOLOGY: ROLE AND IMPORTANCE

Metabolic networks have fundamental and elementary mechanisms that strongly organize gene expression. While these mechanisms have been known, the development of engineered, dynamic controllers or managers has been estimated. Synthetic biology, supported by mathematical conceptions of gene expression, has helped the model-based design of abundant dynamic artificial plans.

Moreover, progresses in applied science are making huge data sets in numerous scientific domains, such as combinatorial chemistry, physics and bioinformatics. In order to find helpful information in these data sets, researchers and engineers are orienting to data mining methods. In fact, data mining is an coherent method designed to inspect data (known as “big data”) in investigation of dependable patterns and/or usual associations between variables, and hence to corroborate the conclusions by applying the identified samples to novel data subsets. The purpose of data mining is to decide composition inside unstructured data, extract importance from fuzzy data, establish patterns in apparently random data, and use all these informations to better understand predispositions, correlations, and ultimately predict result performance and challenge trends. Furthermore, the ending purpose of data mining is predict - and extrapolative data mining is the most common form of data mining and one that has the most shortest production importance. The technique of data mining comprises of three actions: (1) the initial assessment, (2) model building, and (3) utilization.

In addition, a mathematical model can be described as a quantitative design of a system (Ennouri *et al.*, 2013b; Ben Ayed *et al.*, 2017a). A mechanistic model is established on natural instructions, and can be used to assist knowledge or replicate the activity and performance of the system (Ingalls, 2013). Indeed, model analysis can offer insights into the performance of a system below varying conditions. Then, model analysis and simulation can be a useful support to experimental design as well (Saltelli *et al.*, 2008; Ennouri *et al.*, 2017).

As example in animal breeding, simulation permits to recognize, characterise and forecast the associations between the animal (cow, sheep, goat) and his environment in order to enlarge apparatus that boost the effectiveness of use of available resources whilst optimizing behaviour, adaptive capacity and animal performance, structure and energy metabolism. The persistent aim is: the description and quantification of the processes by which

animals achieve, metabolize and divide nutrients between production and other life tasks. These features are essential for predicting the consequences of varied strategies in the organization and genetic selection of the robustness of the systems and animal. To this end, the engineering organizes modelling and phenotyping approaches, as well as learning of most important mechanisms. The phenotyping effort allows symbolizing difference between individuals in their aptitude to manage with environmental perturbations. The modelling effort expands the structure for forecast of animal performance across environments and physiological steps. Simultaneously, modelling and phenotyping supply synergistically to the growth of decision support devices mainly in the field of life sciences.

Genetic simulation studies play an increasingly influential function in assisting scientists to understanding evolutionary and hereditary results of complex processes whose relations cannot be designed or forecasted by analytical approaches. Using simulation in genetics is now in the scene and has gained high popularity. In human genetics, simulation modeling was widely used to discover genetic factors underlying complex diseases, which are caused by multiple genetic parameters (Yang *et al.*, 2003, Miller *et al.*, 2009). In fact, several researcher teams identified and characterized SNP markers by using Genome-wide association studies (GWAS) methodology to more understand the genomic variation affecting dangerous human illnesses like diabetes (Type 2) and also breast cancerous tumors (Peng and Kimmel, 2007). Genetic design is besides employed to assess the accuracy of biostatistical methods, particularly, analyzing molecular marker data to select for example tag single-nucleotide polymorphisms (SNPs) and to design haplotype blocks in order to recognize disease-combined SNP interactions (Miller *et al.*, 2009).

In an agricultural field, plant modeling can simulate crop growth, productivity, and agronomic properties and provide clear answers to farmers in selecting suitable genotypes by taking into account the effect of climatic variation scenarios, edaphic conditions and planting method. Indeed, Crop simulation models coupled with plant physiology, environmental parameters, and molecular genetics have been applied in diverse ways to support plant propagation, specifically in modeling ideotypes, as example 'model plants', for diverse crops and vegetal production environments (Slafer, 2003). In fact, Shi *et al.* (2013) simulated a process-based model to predict high temperature-induced sterility interaction and help to understand the flowering characteristics of rice crop under temperature fluctuation. In addition, Krishnan *et al.* (2007) have developed crop growth-based models to better understand the complex interactions among different environmental variables that influence growth and yield of

crop. Moreover and more recently, Ben Ayed *et al.* (2017a), identified two SNP markers closely associated with fatty acid composition in Tunisian olive oil and designed haplotype blocks, eventually responsible to fatty acid-associated SNP interactions by using statistical genetic simulation. Besides, computer simulation modeling has been used to study the relationships between plant varieties. In fact, Ben Ayed *et al.* (2012, 2016, 2017b) combined genetic simulation approach and molecular marker data to provide a more precise description of olive genetic structure and to simulate the correlation between productivity and morphological, biological and physicochemical parameters of main table olive varieties exploited throughout the globe in order to select the suitable table olive genotype.

Highly sophisticated and scalable Data Mining devices for such applications represent one of the most dynamic research frontiers in Data Mining. Actually, an unprecedented overflow of data forces biologists to modify their approach to scientific breakthrough. In fact, large scale data-collection methods have appeared for a number of data sources restricted by output, of the quantity of obtainable data. Examples of the data comprise: systematic DNA genome sequencing of organisms; improved output determination of small molecule configurations, as well as huge macromolecular constructions (such as DNA, RNA, and proteins); big scale measures of molecular connections; and simultaneous measure expression stage of all genes (hundreds to tens of hundreds) in a group of cells.

Indeed, the availability of this data obliges the biologists to invent system for storing, organizing, and broadcasting it, thus causing a need for criterion terminologies and evolution of standards of switch and notes. Moreover, because of the obvious occasions for automated learning from the data sets, a field for robust machine learning has appeared to take benefit of previous cognition without being overly slanted in the seek for novel knowledge.

Data mining methods are mostly used in several life science sectors and they progressively are playing an essential role in environment and agriculture-related areas. Consequently, it is imperative to supply researchers in agriculture and environmental-related areas with the most sophisticated knowledge innovation methods. Data mining is the procedure of obtaining main and practical information from huge sets of data. This information can be transformed into helpful knowledge that could facilitate to better recognize the difficulty in study and to better forecast future developments.

Simulation of computational and mathematical models of natural networks is an important device for biologists aiming to know the dynamic compartment and behaviour

of complex systems (Ben Ayed et al., 2017b). In the cases of Systems and Synthetic Biology, repetitive cycles of model-driven postulation generation, contested or validated by lab experimentation, lead to advanced quantitative and extrapolative models. *In silico* testing with these models is cheaper, quicker, and more repetitive than its physical complement.

BIOINFORMATICS: APPLICATION IN LIFE SCIENCES AND AGRICULTURE

Bioinformatics is a multidisciplinary discipline appearing from connections of computer, biology, statistics and calculations to study genome organization and compositions, biological sequence data, forecast the configuration and role of macromolecules that employ in clarifying and decrypting the genome of plants (Edwards et al., 2009). The enormous quantities of data produced in life sciences obtained from the evolution and expansion of bioinformatic sciences. Omics, computational devices and bioinformatics are extremely fundamental to comprehend genomics and the molecular systems that trigger numerous vegetal functions (Xue et al., 2007). A range of novel omic levels such as genome; epigenome analysis and metabolome have appeared via scientific progresses. Such incorporation of information allows and assists the recognition of gene expression which aids to understand the affiliation among both genotype and phenotype, hence accepting from genome to phenome system-extensive analysis (Iquebal et al., 2015). Use of different biostatistical devices and databases permit analysis, visualization, annotation and recovery of results to helps improved discernment in living system investigation.

In other hand, numerous analytical methods can be used in the aim to ameliorate result interpretation. Among these techniques, we can cite correspondence Analysis (CA), this method is known as a multivariate graphical procedure designed to investigate associations among categorical variables. In fact, Semmar et al. (2016) proved chemical discrimination involving Chetoui and Oueslati by employing correspondence analysis (CA) method on Fatty Acids. In addition, Correspondence analysis is known as an exploratory and descriptive method proposed to establish simple two-way and multi-way tables comprising some value of correspondence between the columns and rows. The results give information which is analogous in nature to those produced by Factor Analysis methods, and they permit to investigate the configuration of categorical variables integrated in the table. Besides, Ben Hlima et al. (2017) studied variances between Tunisian olive oils belonging to diverse environmental positions and searched trait parameters as well as fatty acids percentage

and phenolic acid concentrations of diverse virgin olive oil experiments via Principal Component Analysis (PCA) and HCA methods. Principal Component Analysis is practical for reducing and understanding huge multivariate data sets with original linear constructions, and for observing in the past unsuspected associations (Begum et al., 2013; Bo et al., 2017). The Hierarchical Cluster Analysis (HCA) procedure attempts to recognize moderately harmonized classes of cases (or variables) based on particular features, via an algorithm that begins with each case (or variable) in a separate group and joins clusters until only one is missing (Sofalian et al., 2012; Giuffrè, 2017). As example, Hierarchical Cluster Analysis was employed on phenol, sterol and triacylglycerol compositions to statistically underline the central function of these metabolites in particular segregation of studied olive cultivars (Manai Jebali et al., 2012).

In addition, we note also the existence of a useful method called discriminant analysis and known as a classification problem, where two or more groups or populations are known *a priori* and one or more new comments are classified into one of the identified populations based on the calculated characteristics (McGarigal et al., 2000; Aziz and Al-Hur, 2013). As example, Petrakis et al. (2008) employed discriminant analysis for experiment identification of diverse greek extra virgin olive oils (EVOO).

In another study, Partial least squares (PLS) was shared with linear regression (PLSR) to forecast fractions of extra virgin olive oils (EVOO) and Palm oil in some vegetable oil combinations composed by sunflower, corn and soybean oils correspondingly (Jiang et al., 2015). Partial least squares (PLS) is defined as a technique for building extrapolative models when there are many extremely collinear features. On other hand, multiple linear regression is known to be a statistical way used to analyse possible connection among one answer variable (defined as a dependent variable) with two or more variables (defined as independent variables). If the association between the dependent and other independent variables could be established by multiple linear regression, an enhanced manage approach could be wanted.

Additionally, Support Vector Machines (SVM) are based on the notion of decision planes that describe assessment limits. A decision plane is one that divides between a set of objects having diverse category memberships. SVMs were used to Italian olive oils described by mid- and near-infrared spectrum in the aim to differentiate among diverse localities (Devos et al., 2014).

Also, a method called K-nearest neighbour (KNN) can be carried-out for both categorization and regression analytical

problems. Nevertheless, it is more commonly used in categorization problems in the manufacturing. KNN algorithm fails across all factors of concern. It is usually used for its simple of explanation and low computation period. Bajoub et al. (2017) employed K-nearest neighbour technique to forecast several olive accessions cultured in North African region (Morocco).

As well, Artificial Neural networks are appropriate in virtually every state in which a connection between the forecaster variables (inputs) and forecasted variables (outputs) subsists, even when that connection is very difficult and not simple to articulate in the typical terms of “correlations” or “differences between groups.” Copied from biological neural network structures, artificial neural networks are defined as assemblages of basic processing elements named artificial neurons attached together to constitute a directed graph (Haykin, 2009). Nodes of the diagram symbolize biological neurons and associations between them symbolize synapses. In opposition with biological neural networks, relations between artificial neurons are not generally included or eliminated after the system was constructed. As an alternative, associations are tested and the weights are adjusted by use of learning algorithm. Input signal circulates into the network in the way of connections until it attains output of the arrangement. In supervised learning, algorithm adjusts the weights in the aim to reduce the variation between the network output and wished output given by the user (Ennouri et al., 2017; Winiczenko et al., 2018).

Zupan et al. (1994) applied ANNs effectively for categorization of more than 500 Italian olive oils from fatty acid profiles (input variables). Besides, Funes et al. (2017) employed artificial neural models to forecast the types of the extra virgin olive oil designed as a quality purpose.

Ordukaya and Karlik (2017) have presented numerous approaches for categorization of olive oils for quality manages via different machine learning algorithm. Naïve Bayesian, K-NN, LDA, Decision Tree, ANN, and SVM were carried out to evaluate their precision results. Beiki et al. (2012) used Support Vector Machines and Naïve Bayes in order to forecast the correct category of olive cultivars having elevated precision. Then, they concluded that their obtained findings revealed that data mining methods can be successfully employed to differentiate among plant cultivars and suggested machine learning used by systems in their work can forecast new olive cultivars through the greatest achievable precision.

Bioinformatic science is considered as the combination of natural science and computational sciences, in the aim to discover explanation to recent biological interrogations

(Emmerich and Wolf, 2004). Nevertheless, there is irresistible bioinformatics reasoning, with various file structures, interfaces, advances and proposals that in training bound otherwise still block the employ of similar devices via beginners generally and genome scientists particularly. Consequently, diverse algorithms, devices and databases have been expanded via unsuited programs and materiel resources, employed in favour of analogous goals although in a someway disorganized approximation, creating a need of principles which encumbers the evolution of skills and knowledge. Its central objective is to locate and identify Single Nucleotide Polymorphisms (SNP) among a genomic dataset corresponding to diverse cultivars, varieties, strains or breeds. For illustration, we have employed comparable device for quality check and to recognize untrustworthy olive oil (Hernandez et al., 2001).

CONCLUSION

Synthetic biology entails discernment between the basic biochemical fundamentals of biological systems. Firstly, a biological view must be arranged of the target system, to be decrypted into a defined design explanation that holds to a particular design appropriate for simulation. Scientists in the areas of artificial biology and systems employ computational methods and techniques to compute, interpret and comprehend complicated organic compositions, such as example the olive oil composition. Simulation is a main device for computational theory testing that is typically performed by estimating ordinary differential equation models. The discipline of artificial biology aims to generate massive biological systems from recognized parameters.

CONFLICT OF INTEREST

All authors affirm that they have no conflict of interest.

REFERENCES

- Atkinson, B. and F. Mavituna. 1991. *Biochemical Engineering and Biotechnology Handbook*. Stockton Press, Michigan.
- Aziz, M. M. A. and F. S. Al-Hur. 2013. Differentiation between three Saudi goat types using size-free canonical discriminant Analysis. *Emir. J. Food. Agric.* 25: 723-735.
- Bajoub, A., S. Medina-Rodriguez, M. Gomez-Romeo, E. A. Ajal, M. G. Bagur-Gonzalez, A. Fernandez-Gutierrez and A. Carrasco-Poncorbo. 2017. Assessing the varietal origin of extra-virgin olive oil using liquid chromatography fingerprints of phenolic compound, data fusion and chemometrics. *Food Chem.* 215: 245-255.
- Begum, F., A. K. M. Aminul Islam, M. Golam Rasul, M. A. K. Mian and M. Mofazzal Hossain. 2013. Morphological diversity of eggplant (*Solanum melongena*) in Bangladesh. *Emir. J. Food. Agric.* 25:45-51.

- Beiki, A. H., S. Saboor and M. Ebrahimi. 2012. A new avenue for classification and prediction of olive cultivars using supervised and unsupervised algorithms. *PLoS ONE*. 7(9): e44164.
- Ben Ayed, R., H. Ben Hassen, K. Ennouri and A. Rebai. 2016. Genetic markers analyses and bioinformatic approaches to distinguish between olive trees (*Olea europaea* L.) cultivars. *Interdiscip. Sci.* 8: 366-373.
- Ben Ayed, R., K. Ennouri, F. Ben Amar, F. Moreau, M. A. Triki and A. Rebai. 2017a. Bayesian and Phylogenetic Approaches for Studying Relationships among Table Olive Cultivars. *Biochem. Genet.* 55: 300-313.
- Ben Ayed, R., K. Ennouri, H. Ben Hlima, S. Smaoui, M. Hanana, M. Mzid, S. Ercisli and A. Rebai. 2017b. Identification and characterization of single nucleotide polymorphism markers in FADS2 gene associated with olive oil fatty acids composition. *Lipids Health Dis.* 16: 138.
- Ben Ayed, R., N. Grati-Kamoun, C. Sans-Grout, F. Moreau and A. Rebai. 2012. Characterization and authenticity of virgin olive oil (*Olea europaea* L.) cultivars by microsatellite markers. *Eur. Food Res. Technol.* 234: 263-271.
- Ben Hlima, H., R. Ben Ayed, K. Ennouri and S. Smaoui. 2017. Geographical discrimination of virgin olive oils from the Tunisian coasts by combining fatty acids and phenolic acids profiles within a multivariate analysis. *J. Oleo Sci.* 66: 963-971.
- Bo, W., B. Fu, G. Qin, G. Xing and Y. Wang. 2017. Evaluation of drought resistance in *Iris germanica* L. based on subordination function and principal component analysis. *Emir. J. Food. Agric.* 29: 770-778.
- Claes, J. E., A. H. Geeraerd and J. F. Van Impe. 1999. Heuristic feed rate profiles for optimal yield and productivity of fed-batch bioprocesses. *Chem. Eng. Commun.* 172: 189-216.
- Devos, O., G. Downey and L. Duponchel. 2014. Simultaneous data re-processing and SVM classification model selection based on a parallel genetic algorithm applied to spectroscopic data of olive oils. *Food Chem.* 148: 124-130.
- Duthie, G. G., P. T. Gardner, P. C. Morrice and D. B. McPhail. 2016. The contribution of α -tocopherol and γ -tocopherol to the antioxidant capacity of several edible plant oils. *Nat. Sci.* 8: 41-48.
- Edwards, D., J. Stajich and D. Hansen. 2009. *Bioinformatics: Tools and applications*. Springer Science and Business Media.
- Emmerich, W and A. L. Wolf. 2004. *Component Deployment*. CD 2004. Lecture Notes in Computer Science, 3083. Springer, Heidelberg.
- Ennouri, K., H. Ben Hassen and N. Zouari. 2013b. Optimization of bioinsecticides overproduction by *Bacillus thuringiensis* subsp. *Kurstaki* using linear regression. *Pol. J. Microbiol.* 62: 287-293.
- Ennouri, K., H. Ben Hassen, S. Ben Khedher and N. Zouari. 2013a. Concomitant production of proteases and delta-endotoxins of *Bacillus thuringiensis* subsp *kurstaki* in a low-cost medium: Effect of medium components. *Acta Biol. Szeged.* 57: 13-19.
- Ennouri, K., R. Ben Ayed, M. A. Triki, E. Ottaviani, M. Mazzarello, F. Hertelli and N. Zouari. 2017. Multiple linear regression and artificial neural networks for delta-endotoxin and protease yields modelling of *Bacillus thuringiensis*. *3 Biotech.* 7: 187.
- Ennouri, K., R. Ben Ayed, M. Mazzarello, E. Ottaviani, F. Hertelli and H. Azzouz. 2016. Classical and Bayesian predictions applied to *Bacillus* toxin production. *3 Biotech.* 6: 206.
- Funes, E., Y. Allouche, G. Beltrán, M. Paz Aguilera and A. Jiménez. 2017. A predictive artificial neural network model as a simulator of the extra virgin olive oil elaboration process. *J. Near Infrared Spectrosc.* 25: 278-285.
- Giuffrè, A.M. 2017. Biometric evaluation of twelve olive cultivars under rainfed conditions in the region of Calabria, South Italy. *Emir. J. Food. Agric.* 29: 696-709.
- Haykin, S. S. 2009. *Neural networks and learning machines*. Pearson, Upper Saddle River, N.J.
- Hernandez, P., R. de la Rosa, L. Rallo, A. Martín and G. Dorado. 2001. First evidence of a retrotransposon-like element in olive (*Olea europaea* L.): Implications in plant variety identification by SCAR-marker development. *Theor. Appl. Genet.* 102: 1082-1087.
- Ingalls, B. 2013. *Mathematical Modelling in Systems Biology: An Introduction*. Applied Mathematics University of Waterloo.
- Iqbal, M. A., S. Jaiswal, C. S. Mukhopadhyay, C. Sarkar, A. Rai and D. Kumar. 2015. Applications of bioinformatics in plant and agriculture. In D. Barh, M.S. Khan and E. Davies (Eds.), *PlantOmics: The Omics of Plant Science*. Springer, India, pp. 755-789.
- Jiang, L., H. Zheng and H. Lu. 2015. Application of UV spectrometry and chemometric models for detecting olive oil-vegetable oil blends adulteration. *J. Food Sci. Technol.* 52: 479-485.
- Krishnan, P., D. K. Swain, B. C. Bhaskar, S. K. Nayak and R. N. Dash. 2007. Impact of elevated CO₂ and temperature on rice yield and methods of adaptation as evaluated by crop simulation studies. *Agric. Ecosyst. Environ.* 122: 233-242.
- Lee, K. R., G. Q. Chen and H. U. Kim. 2015. Current progress towards the metabolic engineering of plant seed oil for hydroxy fatty acids production. *Plant Cell Rep.* 34: 603-615.
- Manai-Djebali, H., D. Krichène, Y. Ouni, L. Gallardo, J. Sanchez, E. Osorio, D. Daoud, F. Guido and M. Zarrouk. 2012. Chemical profiles of five minor olive oil varieties grown in central Tunisia. *J. Food Compos. Anal.* 27: 109-119.
- McGarigal, K., S. Cushman and S. Stafford. 2000. *Multivariate Statistics for Wildlife and Ecology Research*. Springer Verlag, New York.
- Miller, D. J., Y. Zhang, G. Yu, Y. Liu, L. Chen, C. D. Langefeld, D. Herrington and Y. Wang. 2009. An algorithm for learning maximum entropy probability models of disease risk that efficiently searches and sparingly encodes multilocus genomic interactions. *Bioinformatics.* 25: 2478-2485.
- Ordukaya E. and B. Karlik. 2017. Quality control of olive oils using machine learning and electronic nose. *J. Food Qual.* 2017: 1-7.
- Peng, B. and M. Kimmel. 2007. Simulations provide support for the common disease-common variant hypothesis. *Genetics.* 175: 763-776.
- Petrakis, P. V., A. Agiomirgianaki, S. Christophoridou, A. Spyros and P. Dais. 2008. Geographical characterization of Greek virgin olive oils (Cv. *Koroneiki*) using ¹H and ³¹P NMR fingerprinting with canonical discriminant analysis and classification binary trees. *J. Agric. Food Chem.* 56: 3200-3207.
- Purnick, P. E. and R. Weiss. 2009. The second wave of synthetic biology: From modules to systems. *Nat. Rev. Mol. Cell. Biol.* 10: 410-422.
- Ramli, U. S., M. Tang, P. A. Quant, I. A. Guschina, T. Fawcett and J. L. Harwood. 2014. Informed metabolic engineering of oil crops using control analysis. *Biocatal. Agric. Biotechnol.* 3: 49- 52.
- Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana and S. Tarantola. 2008. *Global sensitivity analysis: The primer*. Wiley.
- Semmar, N., S. Laroussi-Mezghani, N. Grati-Kamoun, M. Hammami and J. Artaud. 2016. A new simplex chemometric approach to identify olive oil blends with potentially high traceability. *Food Chem.* 208: 150-160.
- Shi, W., R. Muthurajan, H. Rahman, J. Selvam, S. Peng, Y. Zou and

- K. S. Jagadish. 2013. Source-sink dynamics and proteomic reprogramming under elevated night temperature and their impact on rice yield and grain quality. *New Phytol.* 197: 825-837.
- Slafer, G. A. 2003. Genetic basis of yield as viewed from a crop physiologist's perspective. *Ann. Appl. Biol.* 142: 117-128.
- Sofalian, O., A. Asgari, M. Shokrpour and R. A. Zakarya. 2012. Relationship between storage protein banding pattern and growth type of some bread wheat varieties. *Emir. J. Food. Agric.* 24: 288-292.
- Winiczenko, R., K. Górnicki, A. Kaleta, M. Janaszek-Mańkowska and J. Trajer. 2018. Multi-objective optimization of the apple drying and rehydration processes parameters. *Emir. J. Food. Agric.* 30: 1-9.
- Xue, J., S. Zhao, Y. Liang, C. Hou and J. Wang. 2007. Bioinformatics and its Applications in Agriculture. In D. Li (Eds), *Computer And Computing Technologies In Agriculture, Volume II*. CCTA 2007. The International Federation for Information Processing, Springer, Boston. pp. 977-982.
- Yang, Q., M. J. Khoury, L. Botto, J. M. Friedman and W. D. Flanders. 2003. Improving the prediction of complex diseases by testing for multiple disease-susceptibility genes. *Am. J. Hum. Genet.* 72: 636-649.
- Zupan, J., M. Novic, X. Li and J. Gasteiger. 1994. Classification of multicomponent analytical data of olive oils using different neural networks. *Anal. Chim. Acta.* 292: 219-234.