Briefings in Bioinformatics, 00(00), 2020, 1–12

doi: 10.1093/bib/bbaa204
Method Review

Deep learning meets metabolomics: a methodological perspective

Partho Sen, Santosh Lamichhane, Vivek B. Mathema, Aidan McGlinchey, Alex M. Dickens, Sakda Khoomrung and Matej Orešič

Corresponding author: Partho Sen, Turku Bioscience Centre, University of Turku and Åbo Akademi University, FI-20520 Turku, Finland. Tel.: +358 469608145; E-mail: partho.sen@utu.fi

Abstract

Deep learning (DL), an emerging area of investigation in the fields of machine learning and artificial intelligence, has markedly advanced over the past years. DL techniques are being applied to assist medical professionals and researchers in improving clinical diagnosis, disease prediction and drug discovery. It is expected that DL will help to provide actionable knowledge from a variety of 'big data', including metabolomics data. In this review, we discuss the applicability of DL to metabolomics, while presenting and discussing several examples from recent research. We emphasize the use of DL in tackling bottlenecks in metabolomics data acquisition, processing, metabolite identification, as well as in metabolic phenotyping and biomarker discovery. Finally, we discuss how DL is used in genome-scale metabolic modelling and in interpretation of metabolomics data. The DL-based approaches discussed here may assist computational biologists with the integration, prediction and drawing of statistical inference about biological outcomes, based on metabolomics data.

Key words: artificial intelligence; genome-scale metabolic modelling; lipidomics; machine learning; deep learning; metabolism; metabolomics

Introduction

Metabolomics is a post-genomic approach used to identify and interrogate the repertoire of metabolites (the metabolome) in biological systems [1, 2]. Metabolites are defined as small molecules (molecular weight ≤ 1500 Daltons), which can be produced endogenously as downstream products of cellular metabolism or via environmental sources such as the diet or commensal/invading microbes [1, 3]. Therefore, metabolites serve as a direct functional readout of the physiological state of an organism. Metabolic profiling of cells, tissue, organs and

Partho Sen is a postdoctoral researcher with expertise in systems biology and bioinformatics. He obtained his PhD in biological sciences from University of Montpellier 2, France. He is keen to study human metabolism in clinical health and disorders.
Santosh Lamichhane currently works as a postdoctoral researcher at Turku Bioscience Centre. Lamichhane completed his PhD in food science (Metabolomics) from Aarhus University, Aarhus, Denmark. His research expertise and interest involves lipidomics/metabolomics, the gut metabolome and integrative data analysis.
Vivek B. Mathema has a PhD in bioclinical science and graduate studies in biotechnology. Research expertise and interest include oncology, bioinformatics, parasitology and molecular epidemiology.
Aiden McGlinchey is a postdoctoral research bioinformatician and biologist specializing in regenerative medicine/ageing and human disease, working currently in understanding and modelling the metabolomics of human metabolic diseases.
Alex M. Dickens main research interest is in the use of metabolomics to understand neurological diseases. As such, he has developed both targeted and untargeted assays for the quantification and detection of a wide range of metabolites.
Sakda Khoomrung received a PhD in chemistry from Karl-Franzens University, Austria in 2011. He is currently an instructor at the Department of Biochemistry, Faculty of Medicine Siriraj Hospital. His research focus is to develop and apply tools in metabolomics and systems biology to achieve comprehensive understanding of cellular metabolism and diseases.
Matej Orešič's main areas of research are metabolomics applications in biomedical research and systems medicine. He also initiated the MZmine project, leading to popular open source software for metabolomics data processing.
Submitted: 4 May 2020; **Received (in revised form):** 8 August 2020

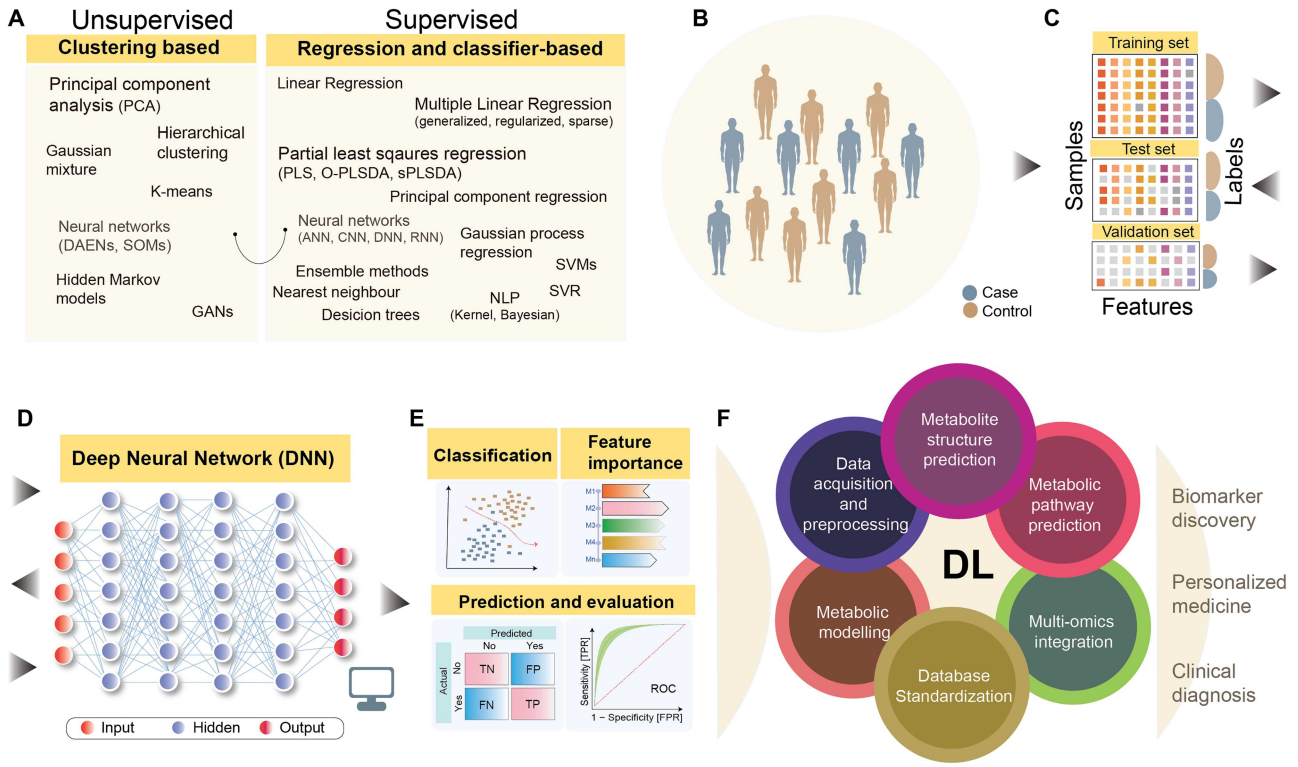

Figure 1. (A) Supervised and unsupervised ML approaches commonly used in metabolomics data processing. (B–C) Partitioning of labelled and heterogeneous data from large cohort(s) into training, test and validation sets. (D) Architecture of DNN, including input, hidden (fully connected) and output layers. Other architectures of DNN, CNN and RNN are reviewed elsewhere [19, 21]. (E) DL-based prediction in patient and disease classification and selection of metabolic predictors in biomarker discovery. Confusion matrix and receiver operating characteristic curve (ROC)-based model performance and diagnostics are shown. (F) Applications of DL in MS data processing that leads to identification of key metabolites and pathway regulators. 'DAENS', 'SOMs', 'GAN', 'SVR', stands for 'a deep autoencoder', 'self-organizing maps', 'a generative adversarial network',' support vector regression', respectively. 'TP', 'TN', 'FP', 'FN', 'TPR', 'FPR' corresponds to 'true positive', 'true negative', 'false positive', 'false negative', 'true positive rate', 'false positive rate', respectively.

biological fluids can be applied to infer the health status of an individual, as well as to aid in monitoring disease-specific changes [4].

Classical multivariate statistics and machine learning (ML) methods such as principal component analysis (PCA), partial least squares (PLSs) and its variants [partial least squares discriminant analysis (PLS-DA), orthogonal-(O)-PLSDA, sparse-(s)-PLSDA)], support vector machines (SVMs), random forests (RFs), kernel machines, Bayesian networks or fuzzy logic have been used to identify metabolic signatures and biomarkers in metabolomics studies [5–8] (Figure 1A). These and other ML techniques have been used to model, infer and predict the inherent (linear or non-linear) relationships found within biological data in the context of health and disease, e.g. in cancer, obesity and diabetes [9–11]. Among these techniques, dimensionality reduction techniques such as PCA and PLS have been widely used to explain variation in the metabolomics data [5].

Recent advances in mass spectrometry (MS) have allowed researchers to generate large volumes of metabolomics data at a high level of sensitivity and resolution. Acquisition and interpretation of complex biological information poses several challenges for conventional data analysis strategies (i.e. functional characterization, annotation and integration) and thus rises the need for the development of novel computational tools [12, 13].

Artificial neural networks (ANNs) are one of the commonly applied ML tools that are used for finding particularly complex patterns within the data [6, 14, 15]. A neural network is based on interconnecting layers of perceptrons, i.e. computational structures inspired by neurons in the brain. An ANN can have any number of layers. These perceptrons, sometimes called 'nodes', pass on a given value ('signal') when a chosen activation function meets specific criteria, much as a biological neuron fires on receiving sufficient stimulation [6, 14, 15]. Deep learning (DL) is an emerging field in ML and artificial intelligence (AI), which employs 'deep' neural networks (DNNs) to accomplish supervised, semi-supervised and unsupervised ML tasks. DL can explicitly learn and predict relationships from unstructured, diverse data sets [14, 16–22]. DNNs take input data and converts them into abstract feature representations in its multiple hidden layers (the characteristic that makes these neural networks 'deep' neural networks) by successively combining outputs from the preceding layer (Figure 1B–D). The intricate structure of DL models allows for modelling of the complex patterns exhibited by large and/or high-dimensional data sets [17]. Various DL approaches have been reviewed extensively [14, 16–21]. DL has improved the performance and interpretability of data analysis over conventional ML models [14, 16–20]. Such approaches thereby have great potential to improve disease prediction, diagnosis and drug discovery. Recently, DL has been applied to MS data acquisition and processing, such as noise

filtering, peak detection, integration and alignment, deconvolution of overlapping peaks, substructure prediction and metabolite identification [23–27]. A selection of open source tools and workflows for ML, particularly based on DL architectures, are listed in Table 1.

DL applied to metabolomics data acquisition and processing

The processing of raw data obtained directly from MS-based metabolomics (e.g. liquid chromatography coupled to MS, LC–MS or gas chromatography coupled to MS, GC–MS) includes several overlapping steps such as feature detection, deconvolution and metabolite identification. Open source platforms such as MZmine 2 [28], XCMS [29], MS-DIAL [30], as well as some vendor-specific software, are widely used for MS data processing. When comparing the methods employed for LC–MS and GC–MS, care must be taken to appreciate the differences between these two techniques. LC–MS is typically coupled to soft ionization methods such as electrospray ionization (ESI) and atmospheric-pressure chemical ionization (APCI) and putative metabolites are identified based on their mass-to-charge ratio and retention time, which is then complemented by tandem MS (fragment spectra). On the other hand, GC–MS is traditionally coupled to hard ionization techniques such as electron impact (EI) that cause in-source fragmentation of metabolites, thus generating a complex spectrum for subsequent spectral deconvolution. Softer ionization techniques can be applied in GC-based methods, e.g. chemical ionization (CI) and APCI. However, these softer ionization techniques often suffer from a loss of sensitivity. Recently, a multi-algorithm, ML-based workflow (WiPP) was introduced for peak detection in GC–MS data [31]. WiPP can evaluate the quality of detected peaks by using a classification scheme combined with seven peak classifiers, providing feature detection for metabolite identification and downstream analysis.

Recently, peak detection was treated as a pattern recognition problem [32] and an ANN was used for the purpose. The authors showed that, in contrast to classical, regression-based approaches, the ANN approach estimated the weighted probability of detection for a particular peak and decided upon its inclusion or exclusion in the final data set. However, this approach was computationally expensive and limited by the ability to calculate the peak area or deconvolute overlapping peaks.

In order to identify missing peaks in LC–MS metabolomics data, Peakonly, a convolutional neural network (CNN) approach was developed [26]. Peakonly can simultaneously recommend the exclusion of low-intensity (noisy) peaks and the inclusion of true positive peaks. However, this CNN-based framework is yet to be compared against other multivariate and ML methods such as PLS-DA, ANN and/or a weighted regression model. With the same objective in mind, Kantz *et al.* [24] used a combined DNN and multiple logistic regression (MLR) approach to classify spectral features from non-targeted LC–MS metabolomics data. The method removed 90% of false positive peaks (noise), without a reduction in the true positive rate. However, as the authors pointed out, this approach will probably require extensive training and validation on each new LC–MS method established in laboratories. This is because the approach itself is based on turning peak shapes into a graphical object and, therefore, each true peak will vary depending on specific analytical conditions. Using MLR was shown to help to overcome some of these issues, as it allows parameters obtained directly from the peak shape to be assessed. However, these parameters will still need to be tuned based on the specific analytical conditions.

As discussed, a chromatographic profile may change in both shape and location along the elution time-axis and therefore poses a challenge for a linear classifier (e.g. PLS-DA). Recently, a CNN classifier has been developed to automatically assess, evaluate and model elution profiles in GC–MS experiments [33]. This technique helped to adjust shifts and variation of peak shapes in the chromatogram. This classifier outperformed PLS-DA, ANN and a local weighted regression model for peak classification [33].

In addition to peak detection and classification, DL has been applied to peak alignment of GC–MS data. ChromAlignNet [34] was designed to improve alignment of GC spectra acquired using high-resolution MS instrument [time-of-flight (TOF) MS]. This technique uses a recurrent neural network (RNN) to train and learn from the data at the individual peak (specific m/z and retention time values) level. In general, the algorithm could detect true positives with a high degree of confidence (84–96%). When applied to breath samples, it was able to detect up to ∼30% of false positive peaks [34]. This shows that clearly more work is needed to reduce the false positive peaks detected form the GC–MS data.

Identification of ‘unknown’ metabolites (i.e. those for which reference spectra and/or structure information are not available) using tandem mass-spectrometry (MS/MS) poses several challenges, as a limited number of MS/MS reference spectra are available. Moreover, MS/MS spectra of a compound may vary, depending on instrument type and collision energy applied. DeepMASS [23], an open-source DNN-based framework, was designed to identify ‘unknown’ metabolites. It employs three key steps: (i) training a DL model with MS/MS spectral data sets, (ii) scoring the structural similarity of ‘unknown’ compounds against an MS/MS spectral database and (iii) ranking the candidate metabolites. DeepMASS showed improved performance over *in silico*, combinatorial fragmentation methods such as MetFrag [35], CFM-ID [27] and MAGMa [36] for metabolite identification.

Neural electron–ionization mass spectrometry (NEIMS), a neural network (NN) framework was designed to predict the electron–ionization mass spectra of given metabolites [37]. The NN framework was able to generate mass spectra for thousands of possible candidates. The performance of NEIMS was tested by predicting the spectra for small molecules from the NIST mass spectral library. The authors showed that NEIMS approach was comparatively faster (5 ms per molecule, with an accuracy of 91.8%). However, this approach does not take into account the intensities of isotopic peaks [37].

In addition, Fan *et al.* [38] introduced an ANN model for metabolite identification using tandem MS data, which predicted the complete MS fingerprint of a metabolite. However, the ANN framework has to be evaluated using additional MS/MS data obtained from different types of instruments and mass spectral libraries.

Recently, Dührkop *et al.* [39] developed CANOPUS, a DNN-based framework, to predict compound classes from fragmentation spectra. CANOPUS was shown to assign class information to every metabolite (including ‘unknowns’) in an LC–MS/MS run. Here, the authors trained a DL model using compound structures from a simulated structure databases and predicted their classes using molecular fingerprints. Moreover, the authors annotated metabolites of ‘unknown’ classes based on their fragmentation spectra. This DNN framework, together with a kernel SVM method, enabled the authors to investigate the chemical

Table 1. A list of publically available resources or modules for DL

Resources	Programming language/libraries	Description	Neural network type	Sources (https://www/)
MXNet	Python, also has supports for Scala, Julia, Clojure, Java, C++, R and Perl	An open source, state-of-the-art and flexible DL framework.	CNN, RNN DNN	mxnet.apache.org
Theano	Python	Implemented for large-scale DL algorithms.	CNN, RNN	deeplearning.net/software/theano/
Torch	C, python	An old open source ML library. It has a library in python called Pytorch.	CNN, RNN	torch.ch, pytorch.org
Tensorflow	R, python, C++, R, CUDA	An open source end-to-end ML platform.	CNN, RNN DNN	tensorflow.org, cran.r-project.org
scikit-multilearn	Python	Implementation of a variety of multi-label classification algorithms including DL libraries.	DNN (multi-label)	scikit.ml
Keras	R, python	Keras is a high-level neural networks API (wrapper), capable of running on top of TensorFlow and Theano.	CNN, RNN	keras.io, cran.r-project.org
Caffe	Python, matlab, C++	Fastest available GPU based CNN implementation.	CNN, RNN	caffe.berkeleyvision.org
darch	R	It implements deep architectures and restricted Boltzmann machines.	DNN	cran.r-project.org
deepnet	R	It implements feed-forward neural network, restricted Boltzmann machines, deep belief network (DBN), stacked autoencoders.	DNN	cran.r-project.org
h2o	R	An open source ML platform that offers feed-forward neural network and deep autoencoders.	DNN	cran.r-project.org
RcppDL	R	Implements denoising autoencoder, stacked denoising autoencoder, restricted Boltzmann machine, DBNs.	DNN, DBN	cran.r-project.org
rnn	R	It implements RNN architectures in base R.	RNN	cran.r-project.org
Barista	Python	A graphical tool for designing and training DNN.	DNN	uni-muenster.de/PRIA/Barista/
superml	R	It builds ML models similar to Scikit-multilearn framework.	–	cran.r-project.org
DeeBNet	Matlab, Octave	A matlab toolbox for DBN.	DBN	ceit.aut.ac.ir/~keyvanrad/DeeBNet%20Toolbox.html
Toupee	Python	A library for running experiments with DL and ensembles on GPUs.	DNN	github.com/nitbix/toupee
deepdetect	C++	A DL API and server.	DNN	deepdetect.com
Crino	Python	A neural-network library based on Theano.	ANN, DNN	github.com/jlerouge/crino
Java deep neural networks	Java	GPU-accelerated java DNN.	DNN	github.com/ivan-vasilev/neuralnetworks

Table 2. DL tools applied to metabolomics (LC–MS and/or GC–MS) data acquisition, processing and downstream analysis. 'DNN', 'CNN' and 'GCN' denotes 'deep', 'convolutional', 'graph convolutional' neural networks, respectively and 'RF' denotes RF

Key steps in metabolic profiling	Neural network	Analytical technique	Description	References
Peak detection	*Peakonly* (CNN)	LC–MS	High-quality peak detection with precision.	[26]
Peak alignment	*ChromAlignNet* (DNN)	GC–MS	Designed to better align the GC spectra.	[34]
Normalization	*NormAE* (deep adversarial learning model)	LC–MS	To normalize the non-linear batch effects in the non-targeted LC–MS data.	[41]
Metabolite identification and structure prediction	*DeepMASS* (DNN)	LC–MS/MS	An open-source DNN based framework was designed to identify 'unknown' metabolites.	[23]
	MetDNA (DNN)	LC–MS/MS	A recursive algorithm applied to the metabolic network to annotate 'unknown' metabolites.	[40]
	CANOPUS (DNN)	LC–MS/MS	A framework to predict the compound classes from a fragmentation spectra.	[39]
Metabolic pathway prediction	GCN and RF	—	A hybrid GCN-based approach to predict metabolic pathways with 95% of accuracy.	[64]

diversity and microbial colonization of the digestive systems of these mice, inferring biological responses at level of compound classes [39].

Another approach is the metabolite annotation and dysregulated network analysis (MetDNA) algorithm [40] (a recursive DL algorithm). MetDNA putatively annotate 'unknown' metabolites by determining the structural similarity of the unknown compounds within a 'reaction pair' of a metabolic network that is assumed to share similar MS/MS spectra.

A novel DL technique called Normalization Autoencoder (*NormAE*) was developed to remove non-linear batch effects, commonly occurring in large-scale metabolomics studies, in non-targeted LC–MS data [41]. *NormAE* is based on non-linear autoencoders (AEs) and adversarial learning techniques. In this case, a regularized AE model was evaluated on two different metabolomics data sets. *NormAE* showed the best calibration results as compared to other conventional, normalization methods. A list of DL tools as applied to metabolomics data acquisition, processing and downstream analysis is shown in Table 2.

Robust metabolomics data processing requires a sufficient number of quality control (QC) samples. A commonly applied initial step is to ensure that any given metabolite peak is detected in a sizeable fraction of the samples analyzed, e.g. at least in ~70% of the samples. The next step is to scrutinize the QC/QA of the data processing as well as instrument performance. The subject of data quality in MS by way of inclusion of QC samples is a subject that has been reviewed extensively [42–44]. In addition to this, community-based reference standards are the gold standard for transparency in validating data quality. For human plasma metabolomics, the SRM1950 NIST standard reference plasma sample [42] is recommended to be included in every batch in order to ensure data consistency with other studies. If the NIST sample cannot be run this frequently, then it is possible to calibrate with an in-house QC sample by running it periodically over time. These methods are vital to create reproducible, accurate, robust metabolomics data. Furthermore, the residual standard error (% RSD, relative standard deviation) of each metabolite should be calculated and compared to pooled QC samples. A suitable RSD cut-off (by convention, an RSD of 30% is used) should be chosen and stated clearly. It is important to check if any novel metabolites are detected in a group as compared to other groups. If a metabolite appears to be exclusive but low concentration in one group, it might be missed in the pooled QC sample. The final aspect of data quality that needs to be considered is whether the biological variation is similar to the instrumentation error then changes in these metabolites cannot be considered reliable and will introduce noise into any further analysis. One way of assessing this is by using the dispersion ratio (D-ratio) [44]. The D-ratio for a given metabolite is estimated by comparing its standard deviation in the QC samples versus the standard deviation in the sample itself. With this in mind, DNNs can be used for imputation and gap-filling in sparse MS data generated by a large-scale, non-targeted experiment. However, the accuracy of DNNs tasked with gap-filling is yet to be evaluated. Taken together, there clearly are opportunities to apply DL to improve and automate QC/QA steps of metabolomics analysis.

DL for stratification of metabolic phenotypes

A typical metabolomics experiment results in the measurement of hundreds to thousands of metabolites, usually with more features than there are samples. Despite the non-linear nature of metabolomics data, the non-linear ML approaches for 'metabotyping' are not yet extensively used [15, 45]. As suggested by

Mendez *et al.* [15] this may be due to the lack of community acceptance because linear method such as PLS-DA are easier to interpret than the non-linear ML methods. Moreover, ANN or DNN models tend to overfit the models due to low sample size of metabolomics data sets, while SVMs and RFs are considered to be more robust when applied to small data sets but remains non-interpretable [15]. However, at present, there is a growing interest in use of DL for analysing metabolic profiles and to extract biologically meaningful information [46–48], especially when classical PLS methods fail to extract such information [49, 50].

DL applied to metabolomics data can capture metabolic features of complex traits. Alakwaa *et al.* [46] showed that DL can accurately predict estrogen receptor status in breast cancer samples. DL using feed-forward networks was applied to a data set with 162 metabolites (variables/features) and predictions were compared to conventional ML approaches such as RF, SVM, prediction analysis for microarrays (PAMs), generalized boosted models, recursive partitioning and regression trees (RPART) and linear discriminant analysis (LDA). The authors demonstrated that DL achieved the highest accuracy (area under the curve, AUC ~ 0.93) as compared to other ML methods. Additionally, this DL approach unveiled eight novel metabolic pathways appearing to promote breast cancer. Overall, the study showed that DL, when applied to a metabolomics data set, may provide network inference about the structure of affected biochemical pathways.

Recently, a novel variant of DNNs, i.e. DNN-MDA [49], an integrated supervised classification and regression technique, was introduced for sample stratification and variable (metabolite) selection. The classical DNN approach was extended to incorporate mean decrease in accuracy (MDA) estimates, which determined the relative 'importance' of each metabolite towards the outcome. The authors showed that DNN-MDA approach had highest classification accuracy (97.8%) as compared to other conventional ML methods examined in the study. However, the classification performance of DNN-MDA varied linearly with sample size. Moreover, the DNN-MDA approach showed reduced predictive ability when applied to highly biased (class-imbalanced) data sets. In a similar study, Asakura *et al.* [20] developed an ensemble DNN (EDNN) algorithm for the improvement of classification and regression performance. EDNN integrated several DNN classifiers and statistical methods such as bootstrap resampling, random sampling of variables and the results obtained from different classifiers were combined. EDNN showed better prediction of the outcome as compared to PCA, RF, SVM and classical DNN [20].

In another study, two recently developed ML algorithms including DL and extreme gradient boosting (XGBoost) were compared to RF in supervised classification. The implemented framework captured metabolic complexity in Alzheimer's disease (AD) [51]. Although the best model was XGBoost (AUC = 0.88), DL showed comparable performance (AUC = 0.85) among the other ML algorithms.

DL algorithms have been considered a huge breakthrough for unstructured data sets such as imaging data [52]. Intriguingly, DL approaches have also shown outstanding performance in imaging-based metabolomics studies. In an MS-based imaging study using DL, Inglese *et al.* [53] revealed the metabolic heterogeneity of cancer. The authors used unsupervised DNN-based technique in combination with parametric T-distributed stochastic neighbor embedding to map the MS-based imaging data set from human colorectal adenocarcinoma biopsies. Based on their findings, the authors provided a novel computational workflow that successfully identified the presence of tumor subgroups. Moreover, they captured novel chemical and biological interactions occurring in the tumor tissue [53]. Overall, these findings demonstrate the scope of DL that can be applied to multiple data types generated from metabolomics experiments. A list of DL (DNN) architectures as applied to metabolomics studies is given in Table 3.

DL-aided integration of multi-omics data

'Multi-omics' techniques interrogate different strata of information flow in biological samples. From genome to transcriptome to proteome and metabolome, cellular life displays a vast scope of information flow, feedback and control. Adding more layers of 'omics' provides an increasingly detailed and, ideally, more useful snapshot of biological processes [54], from subcellular to organismal and even at population levels [55].

Such analyses inevitably bring with them the curse of dimensionality, many redundant features and the concomitant, often widely different, signal-to-noise (S/N) ratios of the different 'omics' technologies [56]. To effectively identify and model underlying patterns of interest, conventional ML approaches cease to be adequate or practical [14]. These require manual feature engineering/selection and elimination and data normalization for each 'omics' data type to create a set of meaningful features. The results of the overarching analysis and their subsequent interpretation are therefore sensitive to the aforementioned choices.

In a multi-omics analysis, DL approaches develop their own internal abstractions of the features in an individual omics data set, eliminating the need for manual feature engineering/selection and increasing the S/N ratio [14, 18, 19, 21]. DL has already been successfully applied to bring together different omics-based technologies in the pursuit of biological research questions [14, 18, 19, 21]. DL, particularly the use of AEs and various forms of neural networks has enabled the successful integration of transcriptomic, genomic and even chemical structure data (SMILES) for the prediction of anti-cancer compound sensitivity [57]. DL was able to predict survival groups of patients with hepatocellular carcinoma (HCC) from six different cohorts. Here, RNA sequencing (RNA-Seq), miRNA sequencing (miRNA-Seq) and methylation data were used to train a DL model. DL explicitly identified multi-omics features associated with differential survival subgroups of patients with HCC [58]. The authors suggested that the DL workflow can be extended to even predict prognosis of HCC.

ML, particularly DL, has been used for the prediction of perinatal outcome in asymptomatic pregnant women with short cervical length. The ML framework integrated proteomics and metabolomics data obtained from amniotic fluid of the mothers in their second trimester of pregnancy, together with sonographic, clinical and demographic factors. The DL models outperformed (with the highest AUC) six other conventional ML techniques such as SVM, GLM, PAM, RF and LDA for the prediction of several outcomes related to child delivery and/or admission to neonatal intensive care unit [59].

In other work, four separate 'omic' data types such as phenotypic, proteomic, metabolomic and genomic data were brought together by Kim *et al.* [60] to model genome-wide effects on downstream metabolic processes and consequent growth in *Escherichia coli*. This work has now resulted in the development of the multi-omics model and analytics platform.

Table 3. DNNs applied to metabolomics studies. 'MDA' denotes mean decrease in accuracy

Neural network	Metabolomics Study	Analytical technique	Outcome and performance	References
DNN	Predication of estrogen receptor status in breast cancer.	GC–TOFMS	Isoleucine, putrescine, glycerol, 5′-deoxy-5′-methylthioadenosine, ornithine, beta tocopherol, phenylalanine and arachidonic acid metabolic marker. AUC (area under the curve) of 0.93.	[46]
DNN – MDA (hybrid model)	Determining the geographical origins of specimens (fish) and identifying their biomarkers.	NMR	Trimethylamine N-oxide, inosinic acid and glycine as metabolic marker. Classification accuracy (97.8%).	[49]
EDNN (ensemble DNN)	Extended study of DNN-MDA.	NMR	EDNN outperformed classical DNN, RFs and SVM.	[20]
DNN	Diagnose Alzheimer's type dementia in blood.	UPLC–MS	20 circulatory metabolites can separate AD from controls; DL produced the AUC of 0.85, while XGBoost produced (AUC of 0.88).	[106]
DNN	Identification of metabolic heterogeneity of colorectal adenocarcinoma.	DESI–MS	Identification of presence of novel tumour subgroups.	[53]
DNN	Comparison of tissue samples from different organs of germ-free (GF) and specific-pathogen-free (SPF) mice.	LC–MS	Identification of novel conjugated bile acids, glucuronic acid derivatives and triterpenoids differed between GF and SPF mice.	[39]

DL in metabolic pathway prediction

ML techniques have been employed to determine metabolic pathway(s) associated with various chemical compounds [61, 62]. Li *et al.* [62] developed the mummichog framework that can successfully predict the functional activity of a metabolic pathway directly from a spectral feature table, without a priori identification. AlAkwaa *et al.* [46] developed Lilikoi, an R package that specializes in personalized pathway-based classification modelling using metabolomics data. In contrast to conventional pathway tools such as MetPA, IMPaLA and MPEA that use metabolite names for pathway mapping, integration and over-representation or enrichment analysis; Lilikoi can transform and incorporate metabolomic profiles from samples to generate personalized pathway profiles, using pathway deregulation scores. Further, it points out any pathway(s) found to be significantly associated with a given disease phenotype. A list of conventional ML methods used for metabolic pathway prediction, analysis and modelling is reviewed in [11, 63].

For this same kind of task, Baranwal *et al.* [64] proposed a hybrid DL architecture and ensemble learning approach that incorporated a graph convolutional network (GCN) and RF classifier to extract molecular structures (SMILES format) for metabolic pathway prediction. The GCN-based approach was able to predict metabolic pathways with 95% accuracy. Their method was able to estimate the relative contribution of metabolites engaged in multiple pathways (multi-class problem).

DL and genome-scale metabolic modelling

A cellular phenotype is mainly regulated by physiological, environmental factors and complex biochemical interactions. A multi-omics experiment may determine the molecular predictors that explain changes in biological signals as outcomes. Genome-scale metabolic modelling (GSMM) is a stoichiometric, constraint-based modelling approach that combines biochemical and genetic information within a computational framework [65–68]. Genome-scale metabolic models (GEMs) provide a comprehensive platform for modelling metabolic fluxes within biological systems. The method can also be used to interrogate, postulate and test hypotheses regarding links between genotype to phenotype [65, 69]. Furthermore, multi-omics data sets can be used to contextualize GEMs (Figure 2A) and thereby develop condition-specific models for cells, tissues and organs [70, 71]. GSMMs have already been used extensively to model human metabolism [72–74]. The GSMM method has also provided a catalogue of human gut microbes [75, 76] based on their metabolic activities. The detailed inner workings of GSMMs and their applications are reviewed elsewhere [69, 77].

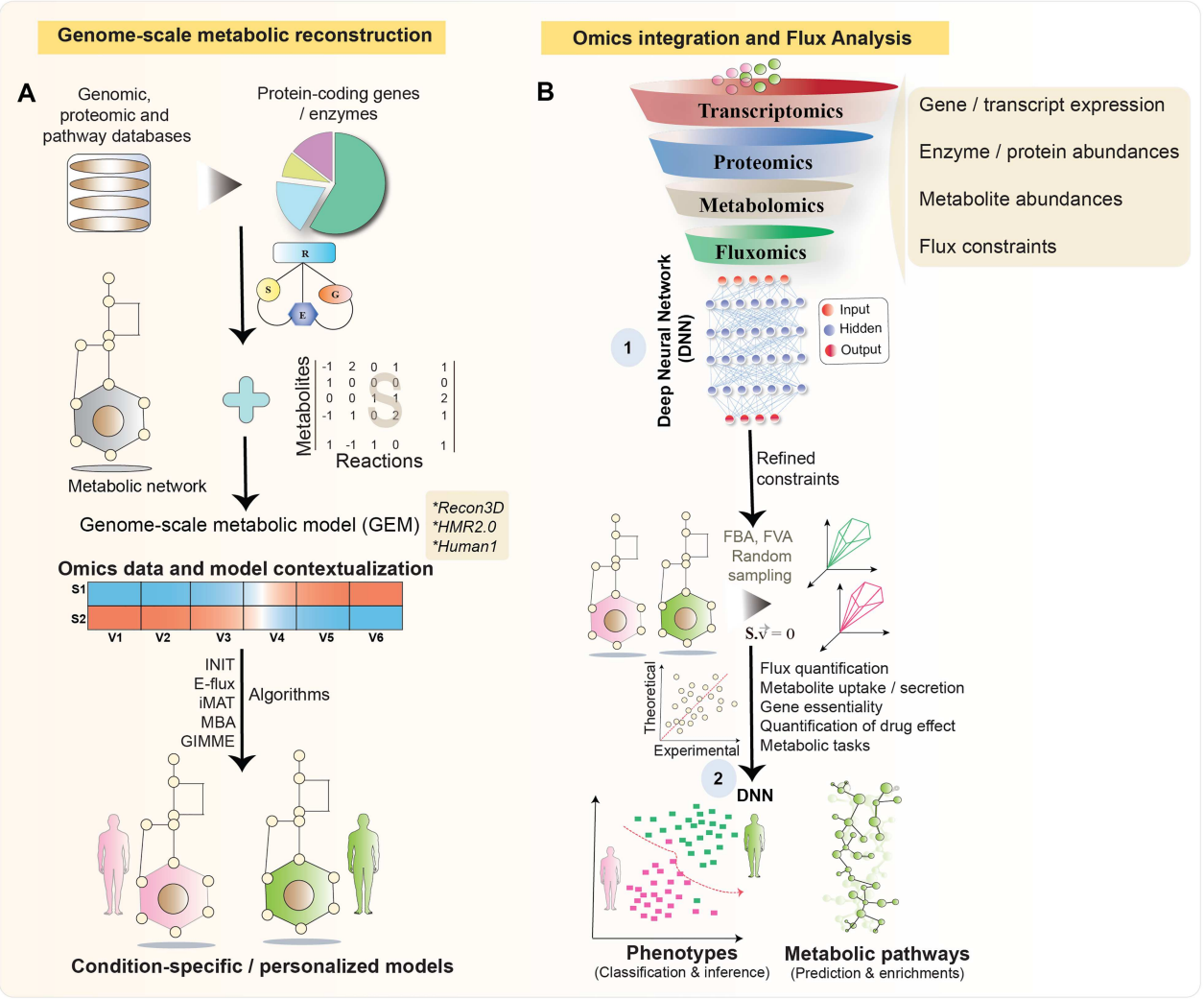

Figure 2. (A) Reconstruction and contextualization of GEMs. Graphical overview of GEM reconstructed from multi-omics data, pathway databases and generic human metabolic reconstructions for the selection of organism-specific metabolic genes, enzymes and pathways/reactions. Reaction components (R) of GEMs include: S = substrates/products (metabolites), E = enzymes, G = genes. Stoichiometric matrix (S) of M_n metabolites and R_n reactions that represents directionality of each metabolites consumed (−1) or produced (+1) or not involved in the reaction (0). Multi-omics data used for selection of key reaction components for building condition-specific/personalized GEMs. S_n and V_n in the heatmap denotes samples and variables (gene expression, metabolite abundances, reaction fluxes, etc.), respectively. (B) DNN used for integration of heterogeneous and complex multi-omics data and derivation of input constraints for GEMs. Various constraints such as reaction rates/flux bounds, protein abundances, enzyme turnover rates and missing thermodynamic constraints can be estimated by harmonizing organism-related published/bibliographic data using DNN. In addition, DL can be used for classification of phenotypes based on flux/biomass predictions and pathway simulations. 'FBA' and 'FVA' denotes 'flux balance analysis' and 'flux variability analysis', respectively.

Recently, ML has been used to study metabolism at the genome-scale [63, 78–82]. This approach brings together structural, biochemical and metabolic network properties [78, 82–84]. However, such modelling efforts do not provide a mechanistic framework for prediction of biological processes [79, 82]. There is a growing interest in combining ML with GSMM to improve biomarker identification (genes, proteins/enzymes and metabolites), e.g. for potential application in personalized medicine [11, 63, 81]. ML can explicitly fine-tune/refine input constraints for GEMs [82] (Figure 2B). Conversely, GEM simulations can be analyzed by ML and compared to experimental data [79] (Figure 2B). Previously, ML methods, including LR, decision trees and naive Bayes were used for gap-filling, i.e. identifying and adding missing reactions and related components in draft GEMs [78, 84].

A prospective GEM was coupled with ML to expedite and investigate causal mechanisms underlying antibiotic efficacy. The framework was used to counter-screen diverse metabolites for bactericidal antibiotics and thereby aided in the mechanistic understanding of antibiotic sensitivity and lethality [85]. Moreover, ML methods were used to successfully predict catalytic (enzymatic) turnover rates in *E. coli*. These values were used to parameterize two mechanistic GEMs, the results improved accuracy in the prediction of the quantitative proteome data [83]. In addition, GSMM and ML methods have been used to evaluate metabolite secretion, quantify metabolite flux, estimate

protein turnover rates, determine the essentiality of genes, predict metabolic genes and assess drug effects, as discussed in [11, 63, 79].

Guo *et al.* [86] have developed DeepMetabolism, a robust, knowledge-based DL system to predict cellular phenotypes from transcriptomics data. The method combines unsupervised (pre-training) and supervised (training) approaches for the determination of phenotypes from genotypes in *E. coli*. The method incorporates a five-layered autoencoder, the connections between the layers determined by biological information, with flux balance analysis (FBA) used to determine and evaluate the connectivity between the proteomic and the phenomic layer. Together, a DNN and a differential search algorithm were applied to design a gene deletion experiment in *E. coli* concerning production of xylitol.

DL *en route* to human metabolism

Metabolomics, when integrated with other omics, can provide highly detailed information about a biological system [14]. The past decade has witnessed the rapid growth of human omics and pathway databases such as the human metabolome database [87, 88], serum metabolome database [89], METLIN [90], REACTOME [91], Kyoto Encyclopedia of Genes and Genomes [92], small molecule pathway database [93] and HumanCyc [94] in parallel with the advancements in high-throughput molecular and analytical technologies.

Along the way, several human metabolic reconstructions were developed and extended to study human metabolism. The first *in silico* global reconstruction of human metabolic networks, Recon 1 and the Edinburgh Human Metabolic Network [95] were reconstructed to combine and analyze biological data sets [66]. However, these models were parsimonious and provided limited knowledge about human metabolism. Subsequently, Recon 2 [96], Recon 2.2 [97] and Recon 3D [72], a community-driven, consensus human metabolic reconstruction and Human Metabolic Reaction [67, 98] were developed with the goal of capturing metabolism under health and disease conditions. Recently, Human1, a rigorously curated human metabolic model was evaluated for its ability to generate cell- and tissue-specific models using transcriptomic, proteomic and kinetic data [73].

The increasing volume of data generated by human metabolic studies necessitates the development of novel computational tools to study underlying pathway(s) and their regulation. Here, DL thrives on massive amounts of biological data and so has started to play a crucial role in the advancement of human metabolomics [99] (Figure 1). Therefore, the utility of DL may not only be limited to effectively profiling metabolites and pathway dynamics in an automated fashion but may also be of great value with contributions to precision medicine [14]. DL has already shown massive potential for disease modelling, drug discovery and the identification of functional metabolomics biomarkers for human diseases [100, 101]. In particular, an ML (elastic net-based) model that utilized serum metabolomics data for chronic kidney disease (CKD) modelling was able to correctly classify (93% accuracy) G1-G4 CKD patients [101]. Likewise, DL architectures involving deep GCNs have been proposed for the prediction of metabolic pathways, molecular shapes and features with accuracies of over 90% [64]. Recent implementations of DL for non-targeted metabolite profiling have focused on the use of high-throughput coverage data-independent acquisition in MS. This has led to improvements in precision, accuracy and robustness of MS-based metabolite identification techniques, suggesting a promising role of ML in human metabolite studies [102].

Concluding remarks and future perspectives

DL has been shown to be of great use in disease phenotyping and patient stratification, including in facilitating the discovery of new therapies [17–19, 58, 59]. However, the application of DL in metabolomics studies is still in its infancy. The reasons for this may include low sample sizes, lack of interpretability and general lack of sufficient reference, training and validation data available.

DL models require large numbers of well-characterized samples/data sets to train. Therefore, a small sample size is likely to decrease DL models' ability to predict and stratify metabolic phenotypes. It also restricts the implementation of DL algorithms developed in other fields such as imaging, which is iteratively trained on very large data sets. To address this issue, hybrid and/or consensus neural networks are needed for the effective interrogation of the data. Additionally, quality of the training data is crucial, an imbalanced (case-control) or skewed data may embed biases in the model prediction. Although DL provides relief from the need for manual selection or elimination of features, DL-based prediction still needs manual curation and therefore, validation of key results from a DL model is pertinent to biomarker discovery. Further, in order to be used by more metabolomics researchers, who may only have basic computational expertise, DL tools and workflows should be user-friendly. Moreover, the formulation and predictions of the DL models themselves should be simplified and easily interpretable by non-experts.

It is also necessary to choose suitable DL methods for the various input data types and to achieve a particular research goal. For instance, DNNs are adept at establishing inherent correlations in high-dimensional data, CNNs excel in delineating spatial information, while RNNs are particularly effective for the analysis of sequential information [18, 19, 21]. Therefore, the choice of a suitable DL tool is of primary importance, and such methodological choices remain open for discussion.

Extending DL architectures to metabolome-wide association studies (MWASs) is a promising future area. DL has been successfully used in genome-wide association studies (GWAS) [103, 104]. GWAS and MWAS share similar data structures [105]; thus, DL methods applied to GWAS can be adapted for MWAS.

In addition, DL models can fine-tune input constraints/parameters for GEMs, when trained with appropriate data sets. As discussed before, these parameters are required to constrain GEMs that otherwise remained poorly defined for lack of experimental data. A fully parameterized GEM may reduce uncertainties across various degrees of freedom and significantly improve the prediction of phenotype (e.g. flux pattern, growth rate, gene essentiality, etc.).

In conclusion, we foresee DL having great potential for resolving several bottlenecks in metabolomics data acquisition, processing, analysis and interpretation. By integrating metabolomics data and DL into bioinformatics pipelines, DL stands to greatly contribute to, among other areas, medical diagnostics, biomarker identification and drug discovery. However, to improve this field, the community seeks more reference and evaluation data to assess and benchmark various DL architectures that might lead towards a responsible innovation.

Key Points

- Mass spectrometry-based metabolomics has generated large volumes of data.

- Acquisition and interpretation of these complex data sets poses several challenges for conventional data analysis strategies, and novel computational tools are needed.
- Deep learning (DL), an emerging field in the machine learning and artificial intelligence, has markedly advanced over the past years.
- Recently, DL has been used in metabolic phenotyping and biomarker discovery. Besides, DL played an essential role in integration of multi-omics data sets, metabolic pathway predictions and metabolic modelling.
- Our review provides insights regarding the practical usage of DL-based frameworks for integration, prediction and drawing of statistical inference about biological outcomes using metabolomics data. In addition, it emphasizes how DL can be used to sort out subsequent bottlenecks in metabolomics data acquisition and processing.

Funding

This study was supported by the Novo Nordisk Foundation (NNF19OC0057418 to M.O.), Juvenile Diabetes Research Foundation (2-SRA-2016-289-S-B to M.O.) and FPU scholarship from the Spanish Ministry of Education, Culture and Sport (FPU15/02373 to M.A.L.-B.).

S.K. is supported by Thailand Research Fund (MRG6280206). The Center of Excellence for Innovation in Chemistry (PERCH-CIC), Ministry of Higher Education, Science, Research and Innovation Thailand, is gratefully acknowledged.

S.L is supported by Academy of Finland postdoctoral grant (no. 323171).

Conflict of Interest

The authors declare no conflict of interest.

References

1. Fiehn O. Metabolomics—the link between genotypes and phenotypes. *Plant Mol Biol* 2002;**48**:155–71.
2. Nicholson JK, Lindon JC, Holmes E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* 1999;**29**:1181–9.
3. Goodacre R. Metabolomics of a superorganism. *J Nutr* 2007;**137**:259s–66.
4. Beckonert O, Keun HC, Ebbels TM, *et al*. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat Protoc* 2007;**2**:2692–703.
5. Yamamoto H, Yamaji H, Abe Y, *et al*. Dimensionality reduction for metabolome data using PCA, PLS, OPLS, and RFDA with differential penalties to latent variables. *Chemom Intel Lab Syst* 2009;**98**:136–42.
6. Mendez KM, Broadhurst DI, Reinke SN. Migrating from partial least squares discriminant analysis to artificial neural networks: a comparison of functionally equivalent visualisation and feature contribution tools using jupyter notebooks. *Metabolomics* 2020;**16**:17.
7. Gromski PS, Muhamadali H, Ellis DI, *et al*. A tutorial review: metabolomics and partial least squares-discriminant analysis—a marriage of convenience or a shotgun wedding. *Anal Chim Acta* 2015;**879**:10–23.
8. O'Shea K, Misra BB. Software tools, databases and resources in metabolomics: updates from 2018 to 2019. *Metabolomics* 2020;**16**:36.
9. Sen P, Dickens AM, Lopez-Bascon MA, *et al*. Metabolic alterations in immune cells associate with progression to type 1 diabetes. *Diabetologia* 2020;**63**:1017–31.
10. Pedersen HK, Gudmundsdottir V, Nielsen HB, *et al*. Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* 2016;**535**:376–81.
11. Cuperlovic-Culf M. Machine learning methods for analysis of metabolic data and metabolic pathway modeling. *Metabolites* 2018;**8**.
12. Cambiaghi A, Ferrario M, Masseroli M. Analysis of metabolomic data: tools, current strategies and future challenges for omics data integration. *Brief Bioinform* 2017;**18**:498–510.
13. Berger B, Peng J, Singh M. Computational solutions for omics data. *Nat Rev Genet* 2013;**14**:333–46.
14. Grapov D, Fahrmann J, Wanichthanarak K, *et al*. Rise of deep learning for genomic, proteomic, and metabolomic data integration in precision medicine. *OMICS* 2018;**22**:630–6.
15. Mendez KM, Broadhurst DI, Reinke SN. The application of artificial neural networks in metabolomics: a historical perspective. *Metabolomics* 2019;**15**:142.
16. Eraslan G, Avsec Z, Gagneur J, *et al*. Deep learning: new computational modelling techniques for genomics. *Nat Rev Genet* 2019;**20**:389–403.
17. Zou J, Huss M, Abid A, *et al*. A primer on deep learning in genomics. *Nat Genet* 2019;**51**:12–8.
18. Wainberg M, Merico D, Delong A, *et al*. Deep learning in biomedicine. *Nat Biotechnol* 2018;**36**:829–38.
19. Angermueller C, Parnamaa T, Parts L, *et al*. Deep learning for computational biology. *Mol Syst Biol* 2016;**12**:878.
20. Asakura T, Date Y, Kikuchi J. Application of ensemble deep neural network to metabolomics studies. *Anal Chim Acta* 2018;**1037**:230–6.
21. Min S, Lee B, Yoon S. Deep learning in bioinformatics. *Brief Bioinform* 2017;**18**:851–69.
22. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;**521**:436–44.
23. Ji H, Xu Y, Lu H, *et al*. Deep MS/MS-aided structural-similarity scoring for unknown metabolite identification. *Anal Chem* 2019;**91**:5629–37.
24. Kantz ED, Tiwari S, Watrous JD, *et al*. Deep neural networks for classification of LC-MS spectral peaks. *Anal Chem* 2019;**91**:12407–13.
25. Colby SM, Nunez JR, Hodas NO, *et al*. Deep learning to generate in silico chemical property libraries and candidate molecules for small molecule identification in complex samples. *Anal Chem* 2020;**92**:1720–9.
26. Melnikov AD, Tsentalovich YP, Yanshole VV. Deep learning for the precise peak detection in high-resolution LC-MS data. *Anal Chem* 2020;**92**:588–92.
27. Allen F, Pon A, Wilson M, *et al*. CFM-ID: a web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra. *Nucleic Acids Res* 2014;**42**:W94–9.

28. Pluskal T, Castillo S, Villar-Briones A, *et al*. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinform* 2010;**11**:395.
29. Smith CA, Want EJ, O'Maille G, *et al*. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching. and identification. *Anal Chem* 2006;**78**:779–87.
30. Tsugawa H, Cajka T, Kind T, *et al*. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat Methods* 2015;**12**:523–6.
31. Borgsmuller N, Gloaguen Y, Opialla T, *et al*. WiPP: workflow for improved peak picking for gas chromatography-mass spectrometry (GC-MS) data. *Metabolites* 2019;**9**:171.
32. Woldegebriel M, Derks E. Artificial neural network for probabilistic feature recognition in liquid chromatography coupled to high-resolution mass spectrometry. *Anal Chem* 2017;**89**:1212–21.
33. Risum AB, Bro R. Using deep learning to evaluate peaks in chromatographic data. *Talanta* 2019;**204**:255–60.
34. Li M, Wang XR. Peak alignment of gas chromatography-mass spectrometry data with deep learning. *J Chromatogr A* 2019;**1604**:460476.
35. Wolf S, Schmidt S, Muller-Hannemann M, *et al*. In silico fragmentation for computer assisted identification of metabolite mass spectra. *BMC Bioinformatics* 2010;**11**:148.
36. Ridder L, van der Hooft JJ, Verhoeven S. Automatic compound annotation from mass spectrometry data using MAGMa. *Mass Spectrom* 2014;**3**:–S0033.
37. Wei JN, Belanger D, Adams RP, *et al*. Rapid prediction of electron-ionization mass spectrometry using neural networks. *ACS Cent Sci* 2019;**5**:700–8.
38. Fan Z, Ghaffari K, Alley A *et al*. Metabolite identification using artificial neural network. In: *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 2019, p. 244–8. IEEE.
39. Dührkop K, Nothias LF, Fleischauer M, *et al*. Classes for the masses: systematic classification of unknowns using fragmentation spectra. *bioRxiv* 2020; 2020.2004.2017.046672.
40. Shen X, Wang R, Xiong X, *et al*. Metabolic reaction network-based recursive metabolite annotation for untargeted metabolomics. *Nat Commun* 2019;**10**:1516.
41. Rong Z, Tan Q, Cao L, *et al*. NormAE: deep adversarial learning model to remove batch effects in liquid chromatography mass spectrometry-based metabolomics data. *Anal Chem* 2020;**92**:5082–90.
42. Hyotylainen T, Oresic M. Optimizing the lipidomics workflow for clinical studies—practical considerations. *Anal Bioanal Chem* 2015;**407**:4973–93.
43. Hyotylainen T, Oresic M. Bioanalytical techniques in non-targeted clinical lipidomics. *Bioanalysis* 2016;**8**:351–64.
44. Broadhurst D, Goodacre R, Reinke SN, *et al*. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics* 2018;**14**:72.
45. Ala-Korpela M, Changani KK, Hiltunen Y, *et al*. Assessment of quantitative artificial neural network analysis in a metabolically dynamic ex vivo 31P NMR pig liver study. *Magn Reson Med* 1997;**38**:840–4.
46. Alakwaa FM, Chaudhary K, Garmire LX. Deep learning accurately predicts estrogen receptor status in breast cancer metabolomics data. *J Proteome Res* 2018;**17**:337–47.
47. Dias-Audibert FL, Navarro LC, de Oliveira DN, *et al*. Combining machine learning and metabolomics to identify weight gain biomarkers. *Front Bioeng Biotechnol* 2020;**8**:6.
48. Ghaffari MH, Jahanbekam A, Sadri H, *et al*. Metabolomics meets machine learning: longitudinal metabolite profiling in serum of normal versus overconditioned cows and pathway analysis. *J Dairy Sci* 2019;**102**:11561–85.
49. Date Y, Kikuchi J. Application of a deep neural network to metabolomics studies and its performance in determining important variables. *Anal Chem* 2018;**90**:1805–10.
50. Hochrein J, Klein MS, Zacharias HU, *et al*. Performance evaluation of algorithms for the classification of metabolic 1H NMR fingerprints. *J Proteome Res* 2012;**11**:6242–51.
51. Stamate D, Kim M, Proitsi P, *et al*. A metabolite-based machine learning approach to diagnose Alzheimer-type dementia in blood: results from the European medical information framework for Alzheimer disease biomarker discovery cohort. *Alzheimers Dement (N Y)* 2019;**5**:933–8.
52. Strack R. Deep learning in imaging. *Nat Methods* 2019;**16**:17.
53. Inglese P, McKenzie JS, Mroz A, *et al*. Deep learning and 3D-DESI imaging reveal the hidden metabolic heterogeneity of cancer. *Chem Sci* 2017;**8**:3500–11.
54. Fabres PJ, Collins C, Cavagnaro TR, *et al*. A concise review on multi-omics data integration for terroir analysis in Vitis vinifera. *Front Plant Sci* 2017;**8**:1065.
55. Sen P, Kemppainen E, Orešič M. Perspectives on systems modeling of human peripheral blood mononuclear cells. *Front Mol Biosci* 2018;**4**:96.
56. Palsson B, Zengler K. The challenges of integrating multi-omic data sets. *Nat Chem Biol* 2010;**6**:787–9.
57. Manica M, Oskooei A, Born J, *et al*. Toward explainable anticancer compound sensitivity prediction via multimodal attention-based convolutional encoders. *Mol Pharm* 2019.
58. Chaudhary K, Poirion OB, Lu L, *et al*. Deep learning-based multi-omics integration robustly predicts survival in liver cancer. *Clin Cancer Res* 2018;**24**:1248–59.
59. Bahado-Singh RO, Sonek J, McKenna D, *et al*. Artificial intelligence and amniotic fluid multiomics: prediction of perinatal outcome in asymptomatic women with short cervix. *Ultrasound Obstet Gynecol* 2019;**54**:110–8.
60. Kim M, Rai N, Zorraquino V, *et al*. Multi-omics integration accurately predicts cellular state in unexplored conditions for Escherichia coli. *Nat Commun* 2016;**7**:13090.
61. Toubiana D, Puzis R, Wen L, *et al*. Combined network analysis and machine learning allows the prediction of metabolic pathways from tomato metabolomics data. *Commun Biol* 2019;**2**:214.
62. Li S, Park Y, Duraisingham S, *et al*. Predicting network activity from high throughput metabolomics. *PLoS Comput Biol* 2013;**9**:e1003123.
63. Angione C. Human systems biology and metabolic modelling: a review from disease metabolism to precision medicine. *Biomed Res Int* 2019;**2019**:8304260.
64. Baranwal M, Magner A, Elvati P, *et al*. A deep learning architecture for metabolic pathway prediction. *Bioinformatics* 2019;**36**:2547–53.
65. Orth JD, Thiele I, Palsson BØ. What is flux balance analysis? *Nat Biotechnol* 2010;**28**:245–8.
66. Duarte NC, Becker SA, Jamshidi N, *et al*. Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc Natl Acad Sci U S A* 2007;**104**:1777–82.

67. Mardinoglu A, Agren R, Kampf C, *et al*. Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. *Nat Commun* 2014;**5**:1–11.
68. Oberhardt MA, Palsson BO, Papin JA. Applications of genome-scale metabolic reconstructions. *Mol Syst Biol* 2009;**5**:320.
69. O'Brien EJ, Monk JM, Palsson BO. Using genome-scale models to predict biological capabilities. *Cell* 2015;**161**:971–87.
70. Wang Y, Eddy JA, Price ND. Reconstruction of genome-scale metabolic models for 126 human tissues using mCADRE. *BMC Syst Biol* 2012;**6**:153.
71. Agren R, Bordel S, Mardinoglu A, *et al*. Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT. *PLoS Comput Biol* 2012;**8**:e1002518.
72. Brunk E, Sahoo S, Zielinski DC, *et al*. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nat Biotechnol* 2018;**36**:272.
73. Robinson JL, Kocabas P, Wang H, *et al*. An atlas of human metabolism. *Sci Signal* 2020;**13**:624.
74. Sen P, Mardinogulu A, Nielsen J. Selection of complementary foods based on optimal nutritional values. *Sci Rep* 2017;**7**:1–9.
75. Li J, Jia H, Cai X, *et al*. An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol* 2014;**32**:834–41.
76. Qin J, Li R, Raes J, *et al*. A human gut microbial gene catalog established by metagenomic sequencing. *Nature* 2010;**464**:59.
77. Sen P, Kemppainen E, Orešič M. Perspectives on systems modelling of human peripheral blood mononuclear cells. *Front Mol Biosci* 2017;**4**:96.
78. Kotera M, Tabei Y, Yamanishi Y, *et al*. Supervised de novo reconstruction of metabolic pathways from metabolome-scale compound sets. *Bioinformatics* 2013;**29**:i135–44.
79. Rana P, Berry C, Ghosh P, *et al*. Recent advances on constraint-based models by integrating machine learning. *Curr Opin Biotechnol* 2019;**64**:85–91.
80. Moore BM, Wang P, Fan P, *et al*. Robust predictions of specialized metabolism genes through machine learning. *Proc Natl Acad Sci U S A* 2019;**116**:2344–53.
81. Zampieri G, Vijayakumar S, Yaneske E, *et al*. Machine and deep learning meet genome-scale metabolic modeling. *PLoS Comput Biol* 2019;**15**:e1007084.
82. Medlock GL, Papin JA. Guiding the refinement of biochemical knowledgebases with ensembles of metabolic networks and machine learning. *Cell Syst* 2020;**10**:109–19.
83. Heckmann D, Lloyd CJ, Mih N, *et al*. Machine learning applied to enzyme turnover numbers reveals protein structural correlates and improves metabolic models. *Nat Commun* 2018;**9**:5252.
84. Liberal R, Pinney JW. Simple topological properties predict functional misannotations in a metabolic network. *Bioinformatics* 2013;**29**:i154–61.
85. Yang JH, Wright SN, Hamblin M, *et al*. A white-box machine learning approach for revealing antibiotic mechanisms of action. *Cell* 2019;**177**:1649–61.
86. Guo W, Xu Y, Feng X. DeepMetabolism: a deep learning system to predict phenotype from genome sequencing. arXiv preprint, arXiv1705.03094, 2017.
87. Wishart DS, Feunang YD, Marcu A, *et al*. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res* 2018;**46**:D608–17.
88. Wishart DS, Knox C, Guo AC, *et al*. HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res* 2009;**37**:D603–10.
89. Psychogios N, Hau DD, Peng J, *et al*. The human serum metabolome. *PLoS One* 2011;**6**:e16957.
90. Smith CA, O'Maille G, Want EJ, *et al*. METLIN: a metabolite mass spectral database. *Ther Drug Monit* 2005;**27**:747–51.
91. Croft D, O'Kelly G, Wu G, *et al*. Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res* 2011;**39**:D691–7.
92. Kanehisa M, Sato Y, Kawashima M, *et al*. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 2015;**44**:D457–62.
93. Jewison T, Su Y, Disfany FM, *et al*. SMPDB 2.0: big improvements to the small molecule pathway database. *Nucleic Acids Res* 2014;**42**:D478–84.
94. Romero P, Wagg J, Green ML, *et al*. Computational prediction of human metabolic pathways from the complete human genome. *Genome Biol* 2005;**6**:R2.
95. Ma H, Sorokin A, Mazein A, *et al*. The Edinburgh human metabolic network reconstruction and its functional analysis. *Mol Syst Biol* 2007;**3**:135.
96. Thiele I, Swainston N, Fleming RM, *et al*. A community-driven global reconstruction of human metabolism. *Nat Biotechnol* 2013;**31**:419–25.
97. Swainston N, Smallbone K, Hefzi H, *et al*. Recon 2.2: from reconstruction to model of human metabolism. *Metabolomics* 2016;**12**:1–7.
98. Mardinoglu A, Agren R, Kampf C, *et al*. Integration of clinical data with a genome-scale metabolic model of the human adipocyte. *Mol Syst Biol* 2013;**9**:649.
99. Costello Z, Martin HG. A machine learning approach to predict metabolic pathway dynamics from time-series multiomics data. *NPJ Syst Biol Appl* 2018;**4**:1–14.
100. Masania J, Faustmann G, Anwar A, *et al*. Urinary metabolomic markers of protein glycation, oxidation, and nitration in early-stage decline in metabolic, vascular, and renal health. *Oxid Med Cell Longev* 2019;**2019**.
101. Yang Y, Tong W-W, Liu Z. A machine learning model to characterize chronic kidney disease with metabolomics data. In: *2019 IEEE International Conference on Healthcare Informatics (ICHI)*. 2019, p. 1–6. IEEE.
102. Ji H, Lu H, Zhang Z. Deep learning enable untargeted metabolite extraction from high throughput coverage data-independent acquisition. *bioRxiv* 2020.
103. Zhou J, Troyanskaya OG. Predicting effects of noncoding variants with deep learning-based sequence model. *Nat Methods* 2015;**12**:931–4.
104. Fergus P, Montanez CC, Abdulaimma B, *et al*. Utilizing deep learning and genome wide association studies for epistatic-driven preterm birth classification in African-American women. *IEEE/ACM Trans Comput Biol Bioinform* 2020;**17**:668–78.
105. Wei J, Wang A, Li R, *et al*. Metabolome-wide association studies for agronomic traits of rice. *Heredity (Edinb)* 2018;**120**:342–55.
106. Kim M, Snowden S, Suvitaival T, *et al*. Primary fatty amides in plasma associated with brain amyloid burden, hippocampal volume, and memory in the European medical information framework for Alzheimer's disease biomarker discovery cohort. *Alzheimers Dement* 2019;**15**:817–27.