scientific reports

OPEN

GAIT-GM integrative cross-omics analyses reveal cholinergic defects in a *C. elegans* model of Parkinson's disease

Lauren M. McIntyre[1,2] ✉, Francisco Huertas[1,3], Alison M. Morse[1,2], Rachel Kaletsky[4], Coleen T. Murphy[4], Vrinda Kalia[5], Gary W. Miller[5], Olexander Moskalenko[6], Ana Conesa[3,8] ✉ & Danielle E. Mor[7] ✉

Parkinson's disease (PD) is a disabling neurodegenerative disorder in which multiple cell types, including dopaminergic and cholinergic neurons, are affected. The mechanisms of neurodegeneration in PD are not fully understood, limiting the development of therapies directed at disease-relevant molecular targets. *C. elegans* is a genetically tractable model system that can be used to disentangle disease mechanisms in complex diseases such as PD. Such mechanisms can be studied combining high-throughput molecular profiling technologies such as transcriptomics and metabolomics. However, the integrative analysis of multi-omics data in order to unravel disease mechanisms is a challenging task without advanced bioinformatics training. Galaxy, a widely-used resource for enabling bioinformatics analysis by the broad scientific community, has poor representation of multi-omics integration pipelines. We present the integrative analysis of gene expression and metabolite levels of a *C. elegans* PD model using GAIT-GM, a new Galaxy tool for multi-omics data analysis. Using GAIT-GM, we discovered an association between branched-chain amino acid metabolism and cholinergic neurons in the *C. elegans* PD model. An independent follow-up experiment uncovered cholinergic neurodegeneration in the *C. elegans* model that is consistent with cholinergic cell loss observed in PD. GAIT-GM is an easy to use Galaxy-based tool for generating novel testable hypotheses of disease mechanisms involving gene-metabolite relationships.

Parkinson's disease (PD) is a progressive neurodegenerative disorder that is characterized by motor symptoms including resting tremor, muscle rigidity, and slowness of movement[1], as well as non-motor symptoms including sleep disturbances and autonomic dysfunction[2]. Several populations of neurons degenerate in PD, including dopaminergic neurons in the substantia nigra, and cholinergic neurons in the dorsal motor nucleus of the vagus and the nucleus basalis of Meynert[3]. Current treatment options mitigate symptoms of the disease but do not offer a cure, emphasizing the urgent need for greater understanding of disease mechanisms to guide the design of new potential therapeutics.

C. elegans is a small nematode worm with an exceptionally high degree of genetic tractability that can be used to uncover biological mechanisms in a tissue and cell type-specific manner. *C. elegans* possess orthologs for 60–80% of human genes, and the worm nervous system utilizes many of the same neurotransmitters as in mammals, including dopamine, acetylcholine, γ-amino butyric acid (GABA), and glutamate[4]. In our recently developed *C. elegans* model of PD, RNAi-mediated reduction of the branched-chain amino acid (BCAA) transferase *bcat-1* recapitulated several features of PD, including age-dependent and progressive motor dysfunction

[1]University of Florida Genetics Institute, Gainesville, FL, USA. [2]Department of Molecular Genetics and Microbiology, University of Florida, Gainesville, FL, USA. [3]Department of Microbiology and Cell Science, University of Florida, Gainesville, FL 32610, USA. [4]Department of Molecular Biology and Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA. [5]Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY 10032, USA. [6]University of Florida Research Computing, University of Florida, Gainesville, FL 32610, USA. [7]Department of Neuroscience and Regenerative Medicine, Medical College of Georgia at Augusta University, Augusta, GA 30912, USA. [8]Present address: Institute for Integrative Systems Biology, Spanish National Research Council, 46980 Paterna, Spain. ✉email: mcintyre@ufl.edu; ana.conesa@csic.es; dmor@augusta.edu
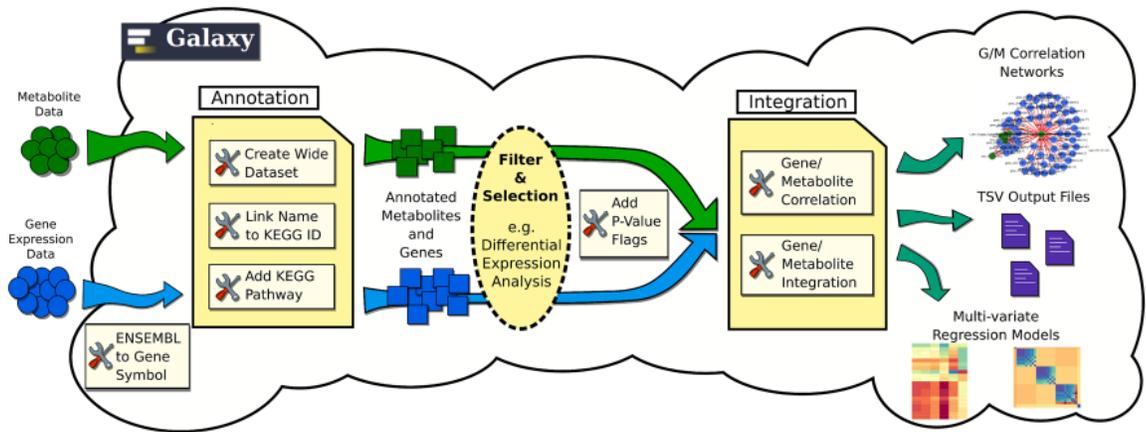
Figure 1. Galaxy pipeline construction enabled by GAIT-GM. Metabolomics and gene expression data are mapped to KEGG pathways. Relevant features are identified using existing statistical analysis approaches. Changes in metabolite levels are modeled as a result of changes in gene expression. Pipelines for data-driven and biology-informed integration are easily created.

characterized by spasm-like 'curling' behavior. *bcat-1* reduction also promoted the loss of dopaminergic cell bodies and degeneration of dopaminergic neurites, with a concomitant decline in dopaminergic neuron function[5,6]. Transcriptomics and metabolomics profiling were used to study underlying molecular mechanisms, and analysis of each of these datasets separately pointed to an important role of mitochondria in driving disease[6,7].

Integration of gene expression and metabolomics data requires advanced bioinformatics skills that are not always within the reach of the biologists who develop and test disease models. However, the experimental scientists are in an advantageous position to interpret the molecular data due to their profound knowledge of the biology of the system and of the significance of analysis results. Therefore, user-friendly bioinformatics tools that provide experimentalists with access to advanced omics data integration analysis methods are needed. Galaxy is a widely used open-source community development platform with an easy to understand GUI interface and a 'mix and match' pipeline building philosophy that makes it attractive to users with limited bioinformatic background. Galaxy has become one of the most successful resources for reproducible omics data analysis and already contains a plethora of tools to analyze gene expression, and metabolomics[8]. However, tools for the integrative analysis of transcriptomics and metabolomics data are not readily available in Galaxy.

GAIT-GM is a new Galaxy tool developed to address these limitations. GAIT-GM leverages the plethora of tools available in Galaxy for the separate analysis of transcriptomics and metabolomics data to create an integration module that combines significant gene expression and metabolite annotations. This allows for studies of both data-driven and biologically oriented relationships among genes and metabolites, while seamlessly managing the data formatting issues that are endemic in trying to integrate these data types. GAIT-GM integration is based on the joint mapping of genes and metabolites to the KEGG database and in the application of dimension reduction[9] and clustering[10] techniques to combine gene expression and metabolite level signals. In addition to this pathway-based integration, GAIT-GM includes unbiased approaches that directly explore links between metabolites and gene expression by modeling metabolite changes as a function of transcriptional regulation (Fig. 1). We have applied these tools to the *bcat-1(RNAi) C. elegans* transcriptomics and metabolomics datasets[6] in order to gain deeper insights into *bcat-1*-associated PD phenotypes. Using GAIT-GM, a previously underappreciated relationship between BCAA metabolism and cholinergic signaling were identified. This relationship was independently tested with new experimental data. These independent follow-up experiments with *C. elegans* confirmed that *bcat-1* reduction causes cholinergic neurodegeneration, paralleling the loss of cholinergic neurons in PD.

Results

GAIT-GM Annotation Tool effectively maps genes and metabolites to pathways. Annotation using GAIT-GM text mining (see Supplementary Methods) resulted in 110/110 metabolites (100%) mapping to KEGG compound IDs and 3096/3146 genes (98%) mapping to KEGG IDs. While some gene and metabolite KEGG IDs did not have associated pathways, 79 metabolites were mapped to a total of 75 unique pathways, and 1187 genes mapped to a total of 115 unique pathways (Fig. S1). Pathways where both metabolites and genes were found included valine, leucine and isoleucine degradation; alanine, aspartate and glutamate metabolism; glycolysis / gluconeogenesis; and citrate cycle (TCA cycle), which recapitulates the pathway identification obtained at the separate analysis of these data[6].

GAIT-GM Integrative Tools characterize the gene-metabolite co-variation network. We first hypothesized that metabolic changes in *bcat-1(RNAi)* worms may relate to global transcriptional patterns. To explore this idea, we used MMC to cluster metabolites into modules and summarized pathway transcriptional activity by obtaining metagene profiles using PANA[9] (see "Methods"). MMC identified nine metabolite modules (Fig. 2A). Several of these modules consisted of metabolites related to BCAA metabolism, including L-valine
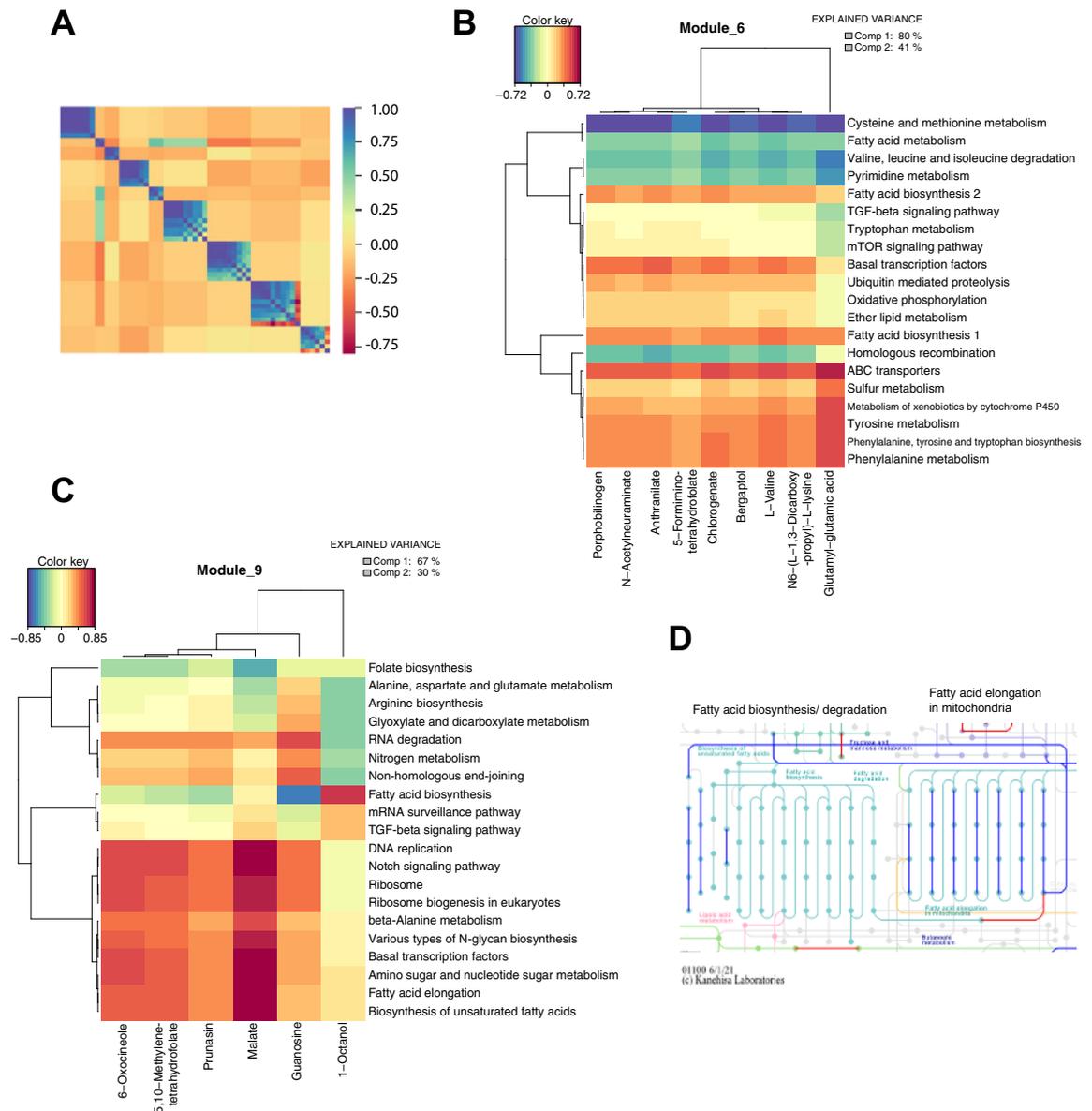
Figure 2. Integration analysis highlights gene-metabolite pathway relationships. (**A**) Modulated Modularity Clustering (MMC) analysis in which a smooth correlation ordering is performed in order to differentiate blocks of metabolites with the same behavior. Metabolites were grouped into 9 different modules. Blue is positively correlated and red is negatively correlated. (**B,C**) Heatmaps[28] of Modules 6 (**B**) and 9 (**C**) after a sparse PLS (sPLS) analysis was performed with the Gene/Metabolite Integration Tool[28]. Metagenes were estimated from the PANA approach[9]. (**D**) Representations of *C. elegans* fatty acid metabolic pathways in KEGG[31]. Genes significantly downregulated in *bcat-1(RNAi)* worms are blue, and genes significantly upregulated in *bcat-1(RNAi)* worms are red.

(Module 6), succinyl-CoA (Module 3), and thiamin diphosphate (Module 1). sPLS analysis of Module 6 revealed a negative correlation between L-valine and the valine, leucine and isoleucine degradation pathway (Fig. 2B), which is consistent with the known increase of BCAAs in the *bcat-1* background[6,11]. Also consistent with previous findings was the positive correlation between L-valine and the oxidative phosphorylation pathway, since *bcat-1* knockdown was reported to increase mitochondrial respiration[6]. In addition, L-valine was positively correlated with the phenylalanine, tyrosine and tryptophan biosynthesis pathway and negatively correlated with homologous recombination, entirely consistent with previously identified upregulation and downregulation of these pathways, respectively, in *bcat-1(RNAi)* worms[6]. In summary, integrative results involving L-valine recapitulated previous observations and provided confidence on the power of the GAIT-GM approach to recover meaningful biological insights.

Module 9 included the TCA cycle metabolite, malate (Fig. 2C), which is thought to be decreased upon *bcat-1* knockdown along with other TCA cycle substrates[6]. In Module 9, malate was highly positively correlated with DNA replication and Notch signaling (Fig. 2C), two pathways known to be downregulated in *bcat-1(RNAi)* worms[6], thus maintaining the expected directional relationship. Interestingly, both Modules 6 and 9 involved
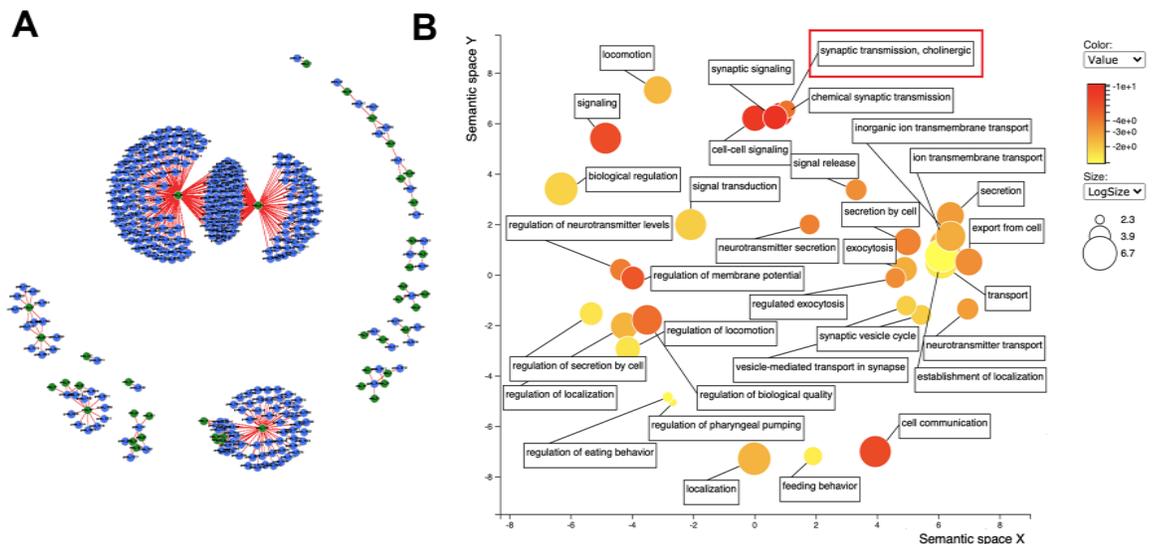
Figure 3. Correlation analysis between individual genes and metabolites points to potential cholinergic defects. (**A**) Output from the Gene/Metabolite Correlation Tool showing the 500 strongest correlations between individual genes and metabolites. Metabolites are represented in green, while genes are represented in blue. Red lines represent positive correlations and light blue lines signify negative correlations. (**B**) Gene Ontology analysis visualized by REVIGO for the largest cluster from (**A**) shows functional enrichment for neurotransmission, and in particular cholinergic synaptic transmission (red box). Gene lists were analyzed using gProfiler[29] and visualized using REVIGO[30].

several pathways related to fatty acid metabolism, including fatty acid biosynthesis, fatty acid elongation, and biosynthesis of unsaturated fatty acids (Fig. 2B,C). Closer examination of the fatty acid metabolic pathways revealed that the associated genes are largely downregulated, particularly those functioning in fatty acid elongation in mitochondria (Fig. 2D). Given that mitochondrial respiration is increased in *bcat-1(RNAi)* worms[6], it possible that fatty acid elongation is being downregulated in favor of fatty acid breakdown and utilization as a substrate for TCA cycle activity. Fatty acid metabolism was not previously highlighted in single omics analyses of the *bcat-1(RNAi)* worms, indicating that GAIT-GM integrated analysis may be used to reveal new potential relationships within networks of co-regulated genes and metabolites.

GAIT-GM data-driven analysis points to potential cholinergic neuron defects. To explore possible gene-metabolite relationships not captured by the pathway analysis, we performed an unbiased discovery analysis using the Pearson correlation between individual differentially expressed metabolites and genes (Fig. 3). Two prominent metabolite hubs with 226 and 60 highly correlated genes, respectively, were identified (Fig. 3A). In the larger cluster of genes, which are positively correlated with the metabolites retinoate and amylopectin, several functional terms related to neurotransmission were significantly overrepresented. These include chemical synaptic transmission, neurotransmitter secretion, and synaptic vesicle cycle (Fig. 3B). The known role of retinoic acid in neuronal development and maintenance[12] offers a possible explanation for the high number of neuronal genes correlated with this metabolite.

The second largest cluster was positively correlated with arachidonyltrifluoromethane, a derivative of arachidonic acid, which also plays a role in neuronal function[13]. However, in contrast to the first cluster, genes in the second largest cluster were functionally enriched for processes related to reproduction, including germ cell development, germ-line stem cell division, and cellular process involved in reproduction in multicellular organism (Fig. S2). These findings suggest the interesting possibility that *bcat-1* disruption may affect reproduction in addition to neuronal health. Thus, GAIT-GM analyses offer several avenues by which to further investigate *bcat-1* regulation of biological and disease processes.

Here were focus on the impact of *bcat-1* on the regulation of neuronal function and its role in neurodegenerative disease. We identified a significant functional enrichment related to cholinergic synaptic transmission (Fig. 3B) in GO analysis of the larger cluster. Further examination of cholinergic synapse components revealed that the associated genes are largely upregulated in *bcat-1(RNAi)* worms, including the vesicular acetylcholine transporter VACHT/*unc-17*, and the choline acetyltransferase CHAT/*cha-1* (Fig. 4). These findings suggested to us that *bcat-1(RNAi)* worms may have abnormal cholinergic signaling and/or damage to cholinergic neurons which may contribute to PD-like disease.

Follow-up confirmation of cholinergic degeneration in the *C. elegans* PD model. To investigate the possibility of cholinergic defects, we crossed transgenic worms expressing GFP in cholinergic neurons with a neuronal RNAi-sensitive strain[14], and fed these animals either *bcat-1* or control (empty vector) RNAi starting at the onset of adulthood. On day 8 of adulthood, when *C. elegans* are considered aged, we imaged cholinergic neurons expressing GFP and quantified neurodegeneration by assessing neurite morphology. Remarkably,
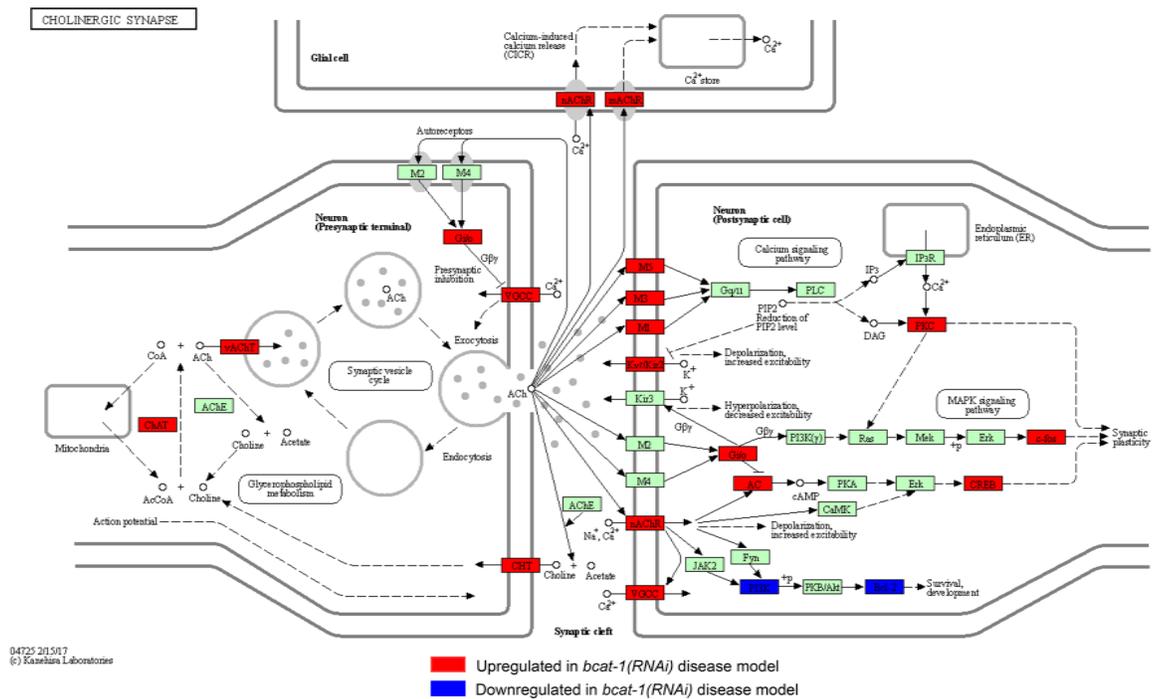
Figure 4. KEGG pathway for the human cholinergic synapse[31] with highlighted genes representing those that are affected in the *C. elegans bcat-1(RNAi)* PD model. Gene orthologs significantly upregulated in *bcat-1(RNAi)* worms are red, gene orthologs significantly downregulated in *bcat-1(RNAi)* worms are blue, and all other genes with KEGG annotations are green.
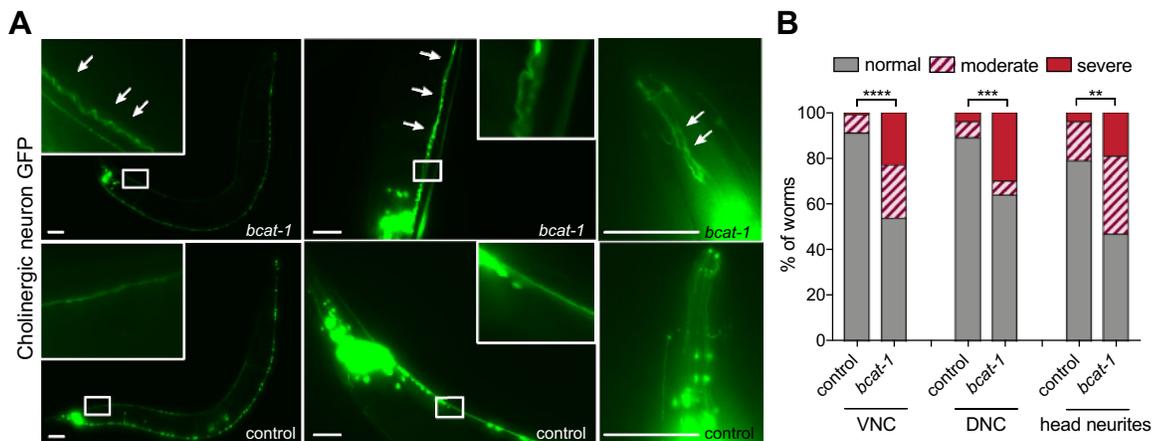
Figure 5. *bcat-1* RNAi-treated worms exhibit cholinergic neurodegeneration. (**A,B**) Cholinergic neuron projections in the head, ventral nerve cord (VNC), and dorsal nerve cord (DNC) of RNAi-sensitive worms (*unc119p::sid-1;unc-17p::GFP*) were scored for the degree of degenerated appearance on day 8 following *bcat-1* or control RNAi treatment during adulthood. In (**A**), representative images of DNC (left), VNC (middle), and head neurites (right) are shown, with arrows indicating abnormal neurite regions. Scale bars, 50 μm. Quantification in (**B**). Chi-square test in, n ≥ 60 worms per group. ** $p < 0.01$; *** $p < 0.001$, **** $p < 0.0001$.

RNAi-mediated reduction of *bcat-1* resulted in significant degeneration of cholinergic neurites in the head, ventral nerve cord, and dorsal nerve cord (Fig. 5). Specifically, cholinergic neurites in *bcat-1(RNAi)* worms had more severe wavy appearances and defasciculation compared with control RNAi-treated animals (Fig. 5). These findings indicate that BCAA metabolism plays a crucial role in the health of cholinergic neurons and illustrate the power of GAIT-GM to integrate metabolite and gene expression data to reveal novel mechanisms of disease.

Discussion

Parkinson's disease (PD) is increasingly recognized as a multi-system disorder involving multiple brain regions and neurotransmitter systems, yet the mechanisms causing neurodegeneration of various neuronal subtypes is largely unknown. The separate analysis of gene expression and metabolomics data in a *C. elegans* model of PD

resulted in valuable new discoveries, in particular that *bcat-1* knockdown in worms alters levels of TCA cycle components and increases mitochondrial respiration. Reducing mitochondrial activity with low-dose treatment of sodium azide mitigates PD-like motor dysfunction and dopaminergic neurodegeneration in *bcat-1(RNAi)* worms[6], underscoring the power of these technologies.

An integrative analysis of the transcriptomics and metabolomics data has the potential to reveal deeper insights into the interplay between gene expression and metabolic changes associated with PD. Existing methods for the integration of metabolomics and gene expression data are often based on multivariate dimension reduction techniques that inform on the covariation structures between transcript and metabolite levels[15,16]. These methods rarely incorporate existing knowledge, which hampers the interpretation of analysis results. Pathway databases such as KEGG, that include both metabolites and genes, can be used to map features to a common biological process and to guide integrative analysis and interpretation. However, mapping to pathways such as in KEGG is limited by the sparsity of metabolites captured in particular metabolic pathways. Conversely, many detected metabolites that cannot be mapped to the pathway data are eventually left aside in the integrative analysis. Moreover, despite the clear evidence of the importance of using metabolite identifiers[17], names are still prevalent, impeding the use of databases that rely on identifiers. Finally, existing pathway tools have limited or no choices for the preprocessing of gene expression and metabolomics data, thus fragmenting the analysis pipeline.

To meet these challenges, we developed GAIT-GM, a Galaxy-based tool for the integrative analysis of multi-omics data. Galaxy is a platform widely used by experimentalists who analyze their own omics data because it allows flexible configuration of analysis pipelines to accommodate a wide range of experimental designs and analysis needs. However, few tools for integrative analysis of transcriptomics and metabolomics datasets exist in Galaxy. The GAIT-GM tool fills a gap in multi-omics data analysis and enables both hypothesis-driven and unbiased hypothesis-generating integrated data analysis in a well-supported open source environment. The tool provides a user-friendly, modular analytical framework that enables the construction of reproducible pipelines and transparent sharing of analytical approaches and results. Additionally, there are two important yet unmet needs in the analysis of metabolomics data specifically developed here: the efficient mapping of compound names to KEGG, and the analysis of the heterogeneity and co-regulation with gene expression to improve functional characterization.

Here, unbiased analysis of the association between metabolites and gene expression in *bcat-1(RNAi)* versus control *C. elegans* using GAIT-GM led to the discovery of cholinergic neurons as a vulnerable cell type to *bcat-1* disruption. We observed severe degeneration and defasciculation of cholinergic axons/neurites in aged *bcat-1(RNAi)* worms. Consistent with this, several populations of cholinergic cells degenerate in PD, including in the dorsal motor nucleus of the vagus and the nucleus basalis of Meynert[3]. In particular, the loss of cholinergic neurons in the pedunculopontine nucleus (PPN) is implicated in gait abnormalities and akinesia in PD[18], and lesions of the PPN in monkeys produce parkinsonian symptoms[19,20]. To our knowledge the present study is the first to link *bcat-1* with cholinergic neurodegeneration. *C. elegans unc-17* and *cha-1* cholinergic synaptic mutants show highly similar motor dysfunction phenotypes to *bcat-1(RNAi)* worms[21], further supporting a connection between BCAA metabolism and cholinergic neuron health. Follow-up studies may wish to dissect which BCAAs are particularly important for cholinergic signaling and/or viability using BCAA and metabolite supplementation, as well as RNAi knockdown of corresponding enzymes.

Our results herein suggest that BCAA metabolic defects may contribute to the vulnerability of multiple cell types in PD, providing a potential unifying mechanism of PD neurodegeneration. We previously found that *bcat-1* reduction promotes dopaminergic neurodegeneration[5,6], and we now show cholinergic degeneration in this PD model, as well. Genome-wide association studies have pointed to other BCAA pathway genes, methylcrotonoylCoA carboxylase 1 (*MCCC1*) and branched-chain ketoacid dehydrogenase kinase (*BCKDK*), as being associated with PD[22], further implicating BCAA metabolism in PD pathogenesis. Moreover, metabolomics studies of PD patient urine and cerebrospinal fluid have shown elevated levels of BCAAs[23] and altered TCA cycle metabolites[24], respectively, consistent with our findings in the *bcat-1(RNAi) C. elegans* model[6].

As demonstrated here, the use of GAIT-GM to uncover previously unknown relationships between genes, metabolites, and functional processes can generate novel testable hypotheses for continued investigation and discovery. Interestingly, we also observed an association between *bcat-1* and reproduction pathways; since *bcat-1* is known to regulate longevity in *C. elegans*[11] and *C. elegans* lifespan regulators are functionally related to human longevity GWAS genes[5], these findings may suggest another avenue of study that could lend insights into the common trade-off in evolution between reproduction and lifespan. Taken together, our present analyses provide greater understanding of the complex relationship between BCAA metabolism and neuronal health, and our new GAIT-GM tool can be widely applied to biological questions requiring transcriptomic and metabolomic data integration.

Methods

All method details for the generation of transcriptomics and metabolomics datasets were previously published[6].

Transcriptomics.   TU3311 uIs60 *(unc-119p::sid-1, unc119p::yfp)* worms were fed either *bcat-1* or control RNAi starting at the L4 stage. On day 5 of adulthood, neurons were isolated and RNA-sequenced[25] for five independent collections from each feeding group. Using Galaxy, reads were mapped to the *C. elegans* genome (WS245) using RNA STAR, mapped reads that overlap with gene features were counted using htseq-count (mode = union), and differential gene expression analysis was performed using DESeq2. Raw sequencing reads are available at National Center for Biotechnology Information BioProject PRJNA599166. A total of 3,146 differentially-expressed genes (false discovery rate < 0.05) were selected for the integrated analysis.

Metabolomics. CF512 *fem-1(hc17); fer-15(b26)* worms were fed either *bcat-1* or control RNAi starting at the L4 stage. On day 5 of adulthood, 6 sets of independent replicates from each feeding group were collected for metabolomics[6]. Raw data are available at Dryad (https://doi.org/10.5061/dryad.5mkkwh72q). Metabolites were annotated through mummichog (level 3 confidence[26]), and those that were significantly different ($p < 0.05$) between the two groups were further examined. If a metabolic feature had multiple putative annotations, the annotation with the least difference between the theoretical and observed mass was retained, and if multiple features were assigned the same putative annotation, the putative annotation with the least difference between the theoretical and observed mass was retained. Adducts and isomers were manually curated for a final list of 110 significantly increased/decreased metabolites selected for the integrated analysis.

GAIT-GM tool. GAIT-GM (Galaxy Annotation and Integration Tools for Genes and Metabolites) consists of two major analysis modules (Fig. 1, see Supplementary Methods for details) that jointly identify, classify, group and link gene expression and metabolomics data. The *annotation tool* maps gene and metabolites to the KEGG database to enable subsequent pathway-based integration. This novel tool implements a text mining approach to parse metabolite names and match them to KEGG compound IDs. Ambiguous metabolites such as lipids of the same type but different formula's (i.e. sphingomyelins) are assigned to the same KEGG ID and considered a metabolite class. Gene names are maximally mapped to KEGG through a combination of databases cross-referencing as described[27]. The *integration tool* provides different options for the clustering and summarization of each omic data modality, as well as methods for integrative analysis (see integration workflow examples in Fig. S3 and Fig. S4). Metabolites might be grouped by KEGG pathway ID, by metabolite class (i.e. all sphingo-myelins), or by abundance patterns (e.g. using the Modulated Modularity Clustering (MMC) tool[10]. Similarly, genes can be grouped by pathway and pathway metagenes representing the expression trend in the pathway are computed using an implementation of the PANA algorithm[9]. Using *correlation tool*, bipartite networks are created to explore the relationship between metabolite and gene expression levels in an unbiased-way and without prior knowledge (Fig. S4). Alternatively, metabolite-gene associations can be studied by multivariate regression methods (Sparse Partial Least Squares or sPLS[28]) where any grouping of explanatory (typically pathway-specific genes or metagenes) and response (i.e. metabolite classes of MMC clusters) variables might be combined. Further information on the GAIT-GM use is provided in the Supplementary User Guide, Supplementary Table 1, and Supplementary Methods. The tool has been implemented as conda and pypi packages and available at the galaxy toolshed repository (see "Data availability" section).

Integrative analysis of *C. elegans* transcriptomics and metabolomics data. Five biological samples of each experimental condition and omics type were used for integrative analysis. Significant genes and metabolites were uploaded to GAIT-GM and annotated to KEGG pathways using the annotation tool. We used GAIT-GM to analyze the pair-wise correlation between individual metabolites and genes and obtained the bipartite network containing the top 500 metabolite-gene pairs with the largest absolute correlation values. Additionally, we analyzed the relationship between metabolite changes and pathway gene expression changes using the multivariate regression method available in the GAIT-GM. For this, we used KEGG pathway metagenes as explanatory variables and MMC-derived clusters for metabolites as response variables.

Gene ontology. Gene lists were analyzed using gProfiler[29] with the following settings: *C. elegans* organism, only annotated genes, g:SCS threshold, user threshold 0.05, ENTREZGENE_ACC. REVIGO[30] was used to cluster and plot GO terms with a q value < 0.05.

Validation experiments. The following *C. elegans* strain was used for validation experiments: CQ491 vsIs48 [*unc-17p::gfp*]; vIs69 [pCFJ90 *(myo-2p::mCherry + unc-119p::sid-1)*]. This strain was generated by crossing LX929 vsIs48 [*unc-17p::gfp*] with LC108 uIs69 [pCFJ90 *(myo-2p::mCherry + unc-119p::sid-1)*][14], both of which were obtained from the *C. elegans* Genetics Center. Worms were maintained at 20 °C on standard high growth medium (HG) plates seeded with OP50 *E. coli*. Following synchronization by bleaching, worms were transferred at day 1 of adulthood to standard nematode growth medium (NGM) plates supplemented with car-benicillin and isopropyl β-D-1-thiogalactopyranoside (IPTG), seeded with HT115 RNAi *E. coli* (either *bcat-1* or L4440 empty vector control). RNAi plates were pre-induced with 0.1 M IPTG at 1 h before transfer, and worms were transferred onto fresh RNAi plates every 2–3 days. On day 8, animals were mounted on 2% agarose pads in M9 and sodium azide for cholinergic neuron imaging. Images were captured on a Nikon Eclipse Ti inverted microscope and processed using Nikon NIS elements software. Neurite morphology was scored as normal, mod-erately degenerated, or severely degenerated at the anterior portion of the ventral and dorsal nerve cords, as well as the neurites projecting from cholinergic neurons in the head ganglia. At least 15 worms were imaged per condition in each replicate.

Data availability

GAIT-GM scripts, test data, Galaxy xmls are available at https://github.com/secimTools/gait-gm, as a PyPi repository (https://pypi.org/project/gait-gm/), and as a bioconda package (https://anaconda.org/bioconda/gait-gm). A detailed Galaxy User Guide providing step-by-step instructions for running each tool in Galaxy is included as Supplementary Material. All tools are deposited in the Galaxy ToolShed for download and installation (https://toolshed.g2.bx.psu.edu/view/malex/gait_gm/ec9ee8edb84d). r-mixomincs version 6.3.2[28], python version 3.7 and R version 4.1.1 and SECOMTools version 21.6.3. Raw RNAseq reads are available at National Center for Biotechnology Information BioProject PRJNA599166 and raw metabolomics data are available at Dryad (https://doi.org/10.5061/dryad.5mkkwh72q).

References

1. Hoehn, M. M. & Yahr, M. D. Parkinsonism: onset, progression and mortality. *Neurology* **17**, 427–442 (1967).
2. Lim, S. Y., Fox, S. H. & Lang, A. E. Overview of the extranigral aspects of Parkinson disease. *Arch. Neurol.* **66**, 167–172 (2009).
3. Forno, L. S. Neuropathology of Parkinson's disease. *J. Neuropathol. Exp. Neurol.* **55**, 259–272 (1996).
4. Corsi, A. K., Wightman, B. & Chalfie, M. A transparent window into biology: A primer on *Caenorhabditis elegans*. *WormBook* https://doi.org/10.1895/wormbook.1.177.1 (2015).
5. Yao, V. *et al.* An integrative tissue-network approach to identify and test human disease genes. *Nat. Biotechnol.* https://doi.org/10.1038/nbt.4246 (2018).
6. Mor, D. E. *et al.* Metformin rescues Parkinson's disease phenotypes caused by hyperactive mitochondria. *Proc. Natl. Acad. Sci. USA* **117**, 26438–26447 (2020).
7. Mor, D. E. & Murphy, C. T. Mitochondrial hyperactivity as a potential therapeutic target in Parkinson's disease. *Transl. Med. Aging* **4**, 117–120 (2020).
8. Kirpich, A. S. *et al.* SECIMTools: A suite of metabolomics data analysis tools. *BMC Bioinform.* **19**, 151 (2018).
9. Ponzoni, I. *et al.* Pathway network inference from gene expression data. *BMC Syst. Biol.* **8**(Suppl 2), S7 (2014).
10. Stone, E. A. & Ayroles, J. F. Modulated modularity clustering as an exploratory tool for functional genomic inference. *PLoS Genet.* **5**, e100079 (2009).
11. Mansfeld, J. *et al.* Branched-chain amino acid catabolism is a conserved regulator of physiological ageing. *Nat. Commun.* **6**, 10043 (2015).
12. Maden, M. Retinoic acid in the development, regeneration and maintenance of the nervous system. *Nat. Rev. Neurosci.* **8**, 755–765 (2007).
13. Bosetti, F. Arachidonic acid metabolism in brain physiology and pathology: Lessons from genetically altered mouse models. *J. Neurochem.* **102**, 577–586 (2007).
14. Calixto, A., Chelur, D., Topalidou, I., Chen, X. & Chalfie, M. Enhanced neuronal RNAi in *C. elegans* using SID-1. *Nat. Methods* **7**, 554–559 (2010).
15. Cavill, R., Jennen, D., Kleinjans, J. & Briedé, J. J. Transcriptomic and metabolomic data integration. *Brief. Bioinform.* **17**, 891–901 (2016).
16. Meng, C. *et al.* Dimension reduction techniques for the integrative analysis of multi-omics data. *Brief. Bioinform.* **17**, 628–641 (2016).
17. Kind, T., Scholz, M. & Fiehn, O. How large is the metabolome? A critical analysis of data exchange practices in chemistry. *PLoS ONE* **4**, e5440 (2009).
18. Pahapill, P. A. & Lozano, A. M. The pedunculopontine nucleus and Parkinson's disease. *Brain* **123**, 1767–1783 (2000).
19. Kojima, J. *et al.* Excitotoxic lesions of the pedunculopontine tegmental nucleus produce contralateral hemiparkinsonism in the monkey. *Neurosci. Lett.* **226**, 111–114 (1997).
20. Aziz, T. Z., Davies, L., Stein, J. & France, S. The role of descending basal ganglia connections to the brain stem in Parkinsonian Akinesia. *Br. J. Neurosurg.* **12**, 245–249 (1998).
21. Sohrabi, S., Mor, D. E., Kaletsky, R., Keyes, W. & Murphy, C. T. High-throughput behavioral screen in *C. elegans* reveals Parkinson's disease drug candidates. *Commun. Biol.* **4**, 203 (2021).
22. Nalls, M. A. *et al.* Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat. Genet.* **46**, 989–993 (2014).
23. Luan, H. *et al.* Comprehensive urinary metabolomic profiling and identification of potential noninvasive marker for idiopathic Parkinson's disease. *Sci. Rep.* **5**, 13888 (2015).
24. Willkommen, D. *et al.* Metabolomic investigations in cerebrospinal fluid of Parkinson's disease. *PLoS ONE* **13**, e0208752 (2018).
25. Kaletsky, R. *et al.* The *C. elegans* adult neuronal IIS/FOXO transcriptome reveals adult phenotype regulators. *Nature* **529**, 92–96 (2016).
26. Schymanski, E. L. *et al.* Identifying small molecules via high resolution mass spectrometry: Communicating confidence. *Environ. Sci. Technol.* **48**, 2097–2098 (2014).
27. Hernández-de-Diego, R. *et al.* PaintOmics 3: A web resource for the pathway analysis and visualization of multi-omics data. *Nucleic Acids Res.* **46**, W503–W509 (2018).
28. Rohart, F., Gautier, B., Singh, A. & Lê Cao, K. A. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput. Biol.* **13**, e1005752 (2017).
29. Raudvere, U. *et al.* g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).
30. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* **6**, e21800 (2011).
31. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

Author contributions

D.E.M., A.M.M., R.K., C.T.M., V.K., G.W.M., A.C., and L.M.M. designed the research; F.H., O.M. and A.M.M. coded, implemented and tested the GAIT_GM tool; A.M.M. wrote GAIT-GM documentation. L.M.M. and A.C. supervised GAIT-GM design and implementation. D.E.M., A.M.M., R.K., V.K. performed research and analyzed the data; D.E.M., A.C., and L.M.M. wrote the main manuscript. A.M.M. wrote the supplementary methods and User Guide.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-07238-9.

Correspondence and requests for materials should be addressed to L.M.M., A.C. or D.E.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.