



**HAL**  
open science

# The mitoXplorer 2.0 update: integrating and interpreting mitochondrial expression dynamics within a cellular context

Fabio Marchiano, Margaux Haering, Bianca Hermine Habermann

## ► To cite this version:

Fabio Marchiano, Margaux Haering, Bianca Hermine Habermann. The mitoXplorer 2.0 update: integrating and interpreting mitochondrial expression dynamics within a cellular context. *Nucleic Acids Research*, 2022, 50 (W1), pp.W490-W499. 10.1093/nar/gkac306 . hal-03769539

**HAL Id: hal-03769539**

**<https://hal.science/hal-03769539>**

Submitted on 25 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The mitoXplorer 2.0 update: integrating and interpreting mitochondrial expression dynamics within a cellular context

Fabio Marchiano, Margaux Haering and Bianca Hermine Habermann \*

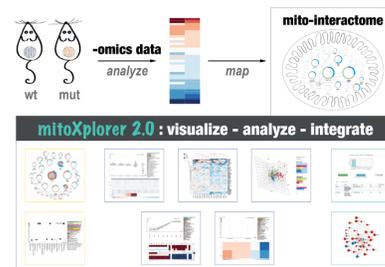
Aix-Marseille University, CNRS, IBDM UMR 7288, 13009 Marseille, France

Received January 25, 2022; Revised April 09, 2022; Editorial Decision April 12, 2022; Accepted April 15, 2022

## ABSTRACT

Mitochondria are subcellular organelles present in almost all eukaryotic cells, which play a central role in cellular metabolism. Different tissues, health and age conditions are characterized by a difference in mitochondrial structure and composition. The visual data mining platform mitoXplorer 1.0 was developed to explore the expression dynamics of genes associated with mitochondrial functions that could help explain these differences. It, however, lacked functions aimed at integrating mitochondria in the cellular context and thus identifying regulators that help mitochondria adapt to cellular needs. To fill this gap, we upgraded the mitoXplorer platform to version 2.0 (mitoXplorer 2.0). In this upgrade, we implemented two novel integrative functions, network analysis and transcription factor enrichment, to specifically help identify signalling or transcriptional regulators of mitochondrial processes. In addition, we implemented several other novel functions to allow the platform to go beyond simple data visualization, such as an enrichment function for mitochondrial processes, a function to explore time-series data, the possibility to compare datasets across species and an IDconverter to help facilitate data upload. We demonstrate the usefulness of these functions in three specific use cases. mitoXplorer 2.0 is freely available without login at <http://mitoxplorer2.ibdm.univ-mrs.fr>.

## GRAPHICAL ABSTRACT



## INTRODUCTION

Mitochondria are essential organelles in most eukaryotic cells and are involved in a multitude of cellular processes, including cellular energy production, metabolism, signalling or apoptosis. The mitochondrial protein content is estimated to be around 1000 proteins, varying slightly between species (1–3). Most of these proteins (mito-proteins) are encoded in the nucleus and imported into mitochondria in a controlled manner. Only few mito-proteins are encoded in the mitochondrial genome and transcribed and translated within mitochondria, encoding proteins of the respiratory chain, ribosomal RNAs required for the mitochondrial ribosome and transfer RNAs. Mitochondria are essential for the functioning of the cell; however, the requirements on mitochondria vary with the cell type. Consequently, mitochondrial structure and protein content differ between cell types (4) and we want to understand how mitochondria adapt to their cellular environment.

mitoXplorer version 1.0 (mitoXplorer 1.0) was developed as a visual data mining platform that allows us to analyse the dynamics of gene expression and mutations of all genes with a mitochondrial function (mito-genes) (1). It contains the most complete and up-to-date mitochondrial interactomes for four different model species (human, mouse, *Drosophila melanogaster* and budding yeast) as well as four different visualization interfaces that enable users to mine expression and mutation data in a comparative manner, as well as from a single dataset. mitoXplorer 1.0 was the first platform available to mine expression dynamics of

\*To whom correspondence should be addressed. Tel: +33 649 443263; Email: bianca.habermann@univ-amu.fr

mito-genes in different conditions. However, it fails to integrate the mitochondrial interactome into the cellular context and thus to help identify potential signalling and transcriptional regulators that adjust mitochondria and their functions to their cellular environment.

We here introduce mitoXplorer version 2.0 (mitoXplorer 2.0), in which we have added integrative data analysis functions to help address these questions. MitoXplorer 2.0 contains several new functions for data analysis, such as an enrichment function for mito-processes, a function to explore time-series data, the possibility to compare datasets across species and two novel data integration functions that aim at identifying mitochondrial regulators: first, a function to identify potential transcriptional regulators of mito-genes with a similar expression profile (co-regulated mito-genes) by making use of our recently published AnnoMiner web tool (5); and second, a function to identify potential signalling pathways from and to mitochondria, by embedding the mito-interactome in the cellular interactome and subsequently exploring the network neighbourhood of a selected mito-gene based on integrating the network with differential expression data (6). We demonstrate the usefulness of these new functions in three use cases, where we identify potential transcriptional regulators driving the mitochondrial metabolic switch during *Drosophila* flight muscle development (use case 1, Figure 2), identify potential active signalling pathways regulating  $\text{Ca}^{2+}$  signalling in ataxia (use case 2, Figure 3) and explore conserved mito-gene deregulation in fibroblasts from human and a mouse model of trisomy 21 (use case 3, Supplementary Figure S1).

## MATERIALS AND METHODS

Novel functions of mitoXplorer version 2.0, together with those already available in mitoXplorer version 1.0, are shown in Figure 1 as well as Supplementary Figure S2.

### Novel functions in mitoXplorer 2.0

**Mito-process enrichment analysis.** In order to identify enrichment of our 38 manually curated mito-processes in a dataset, we implemented a gene set enrichment analysis (GSEA) function. This function helps guide users to focus on important mito-processes of a dataset for subsequent deeper analysis with other functions available in mitoXplorer 2.0.

Mito-process enrichment analysis was coded using the GSEAPy Python library (7,8). When  $P$ -value and  $\log_2\text{FC}$  are both present, genes are ranked by the combined score as defined by Xiao *et al.* (9):

$$\text{combined score} = \text{abs}(\log_2\text{FC}) \times -\log_{10}(P\text{-value}).$$

If no  $P$ -value is provided, genes are ranked by the absolute value of  $\log_2\text{FC}$ . The gene set enrichment score is tested against a null distribution of enrichment scores generated from 100 permuted gene sets composed of randomly selected genes from the input dataset. The results of the GSEA analysis are shown in the form of an interactive dot plot, as well as a bar plot (see Figure 2A for an example). In the dot plot, mito-processes are represented by bubbles, whereby the size reflects the normalized enrichment score

and the colour represents the combined score. The  $Y$ -axis shows the significance of the test ( $-\log_{10}\text{FDR}$ ), while the  $X$ -axis shows the different mito-processes. By clicking on one of the mito-process bubbles, all its values associated with the GSEAPy analysis are displayed in the information panel. In the bar plot, mito-processes are ranked by combined score, whose value is also given in parentheses next to the mito-process.

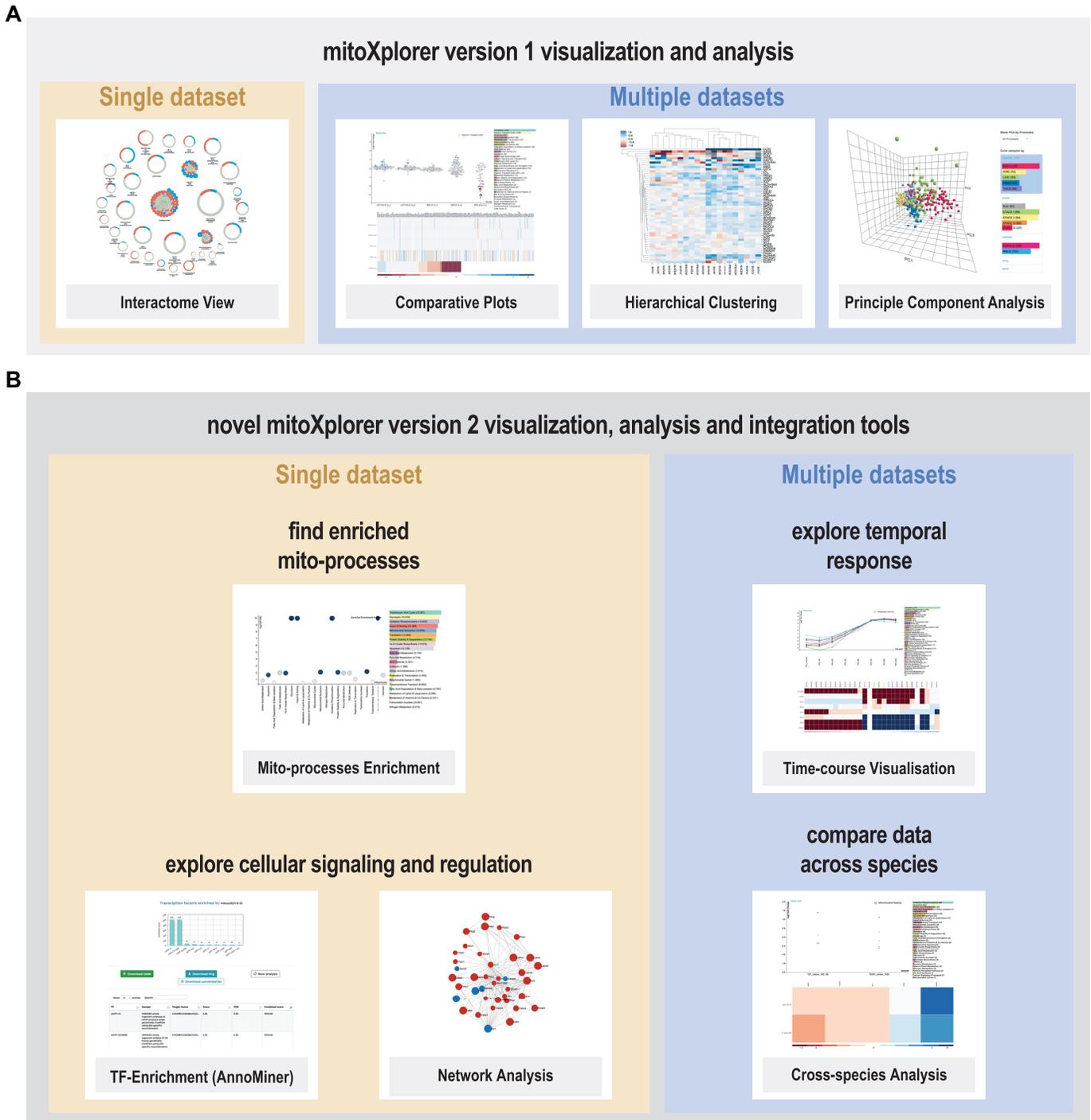
**Time-course visualization.** In mitoXplorer 1.0, only six datasets could be compared to each other with the comparative plots. To overcome this limitation and to effectively visualize an increasing number of available time-course data, we implemented a time-course visualization interface. With this function, we allow users to visualize up to 20 time points of a temporal dataset. The structure of this visualization interface is similar to the comparative plots, with all its interactivity, except that genes can be individually selected from the sortable heatmap for display in the line plot by clicking on the gene name. Time points on the  $X$ -axis are connected via a line to visualize expression profiles over time. The user can thus visualize the temporal gene expression profile of a single or multiple mito-genes. Colours are assigned automatically to help visualize individual genes (see Figure 2B). This visualization was developed using the D3.js library (10) in order to produce a fully interactive visual interface.

**Cross-species analysis.** mitoXplorer 2.0 offers the possibility to compare datasets across species. Cross-species analysis is possible with the comparative plots or the heatmap function. This will help unravel similarities and understand differences in mitochondrial dynamics between species.

Orthology information has been included in the backend MySQL database of mitoXplorer by adding orthology ID tables for each model organism. These tables contain dictionaries, having as key the ID of the gene of the model organism of interest and as values the orthologous gene IDs in the other species available in mitoXplorer 2.0. Orthologs between the four model species offered by mitoXplorer were initially collected using the last release of BioMart (11). We manually curated these datasets and only 1:1 relations were kept.

The user has to first choose ‘cross-species’ from the ‘Data Mining’ drop-down menu and subsequently select the species to analyse. Only the comparative plot and the heatmap are offered for cross-species analysis. In the resulting plots of both visualization interfaces, the IDs of the first organism selected are used to label genes. However, the user can retrieve the ID of the ortholog of the second organism together with underlying information, by clicking on the respective cell in the heatmap or the bubble in the scatterplot.

**Integrative analysis functions.** In mitoXplorer 2.0, we also wanted to provide the possibility to identify mitochondrial regulators. More specifically, we wanted users to be able to find potential transcriptional regulators of co-regulated mito-genes, as well as identify potentially active signalling pathways from or to mitochondria. The interactivity of mitoXplorer allows one to directly select the gene(s) for both new integrative analysis functions starting from the com-



**Figure 1.** Collection of functions available in mitoXplorer 2.0. **(A)** Visualization and analysis tools available already in mitoXplorer 1.0 include the interactome view, comparative plots, hierarchical clustering and principal component analysis. **(B)** Novel functions available in mitoXplorer 2.0 are enrichment of mito-processes, time-course visualization for up to 20 time points, cross-species analysis, as well as the data integration functions of transcription factor (TF)-enrichment via AnnoMiner and network analysis using the viPER algorithm.

parative plot, heatmap or time-course visualization functions.

**TF-enrichment.** Identifying potential transcriptional regulators is done by performing a TF-enrichment analysis via connecting mitoXplorer to the AnnoMiner web server (5) and through direct upload of data from mitoXplorer to AnnoMiner. In brief, AnnoMiner seeks enriched peaks of transcriptional regulators from the ENCODE (12), modENCODE (13) and modERN (14) databases in the promoter regions of a user-provided gene set. This set of genes for an AnnoMiner TF enrichment analysis can directly be selected in mitoXplorer 2.0 and uploaded to the AnnoMiner web server via the TF-enrichment function (see Supplementary Figure S2E for the TF-enrichment menu). The selection can be done from three different interfaces: comparative plots, heatmap and time-course visualization. After invoking the ‘Integrative Analysis’ TF-enrichment, the user needs to select the genes by clicking on the boxes of the genes in the (sortable) heatmap (see Supplementary Figure S3 and our mitoXplorer video tutorials for detailed instructions). The user must then perform TF-enrichment using the AnnoMiner web server (a detailed user tutorial of AnnoMiner can be found at <http://chimbrazo.ibdm.univ-mrs.fr/AnnoMiner/tutorial.html> and Supplementary Figure S3). As a result, the user receives a list of TFs that potentially regulate the selected mito-gene set within the AnnoMiner web server environment.

**Network analysis.** The integrative network analysis function aims at identifying active subnetworks (representing potential signalling paths) from and to mitochondria, starting from one selected mito-gene. To search for active subnetworks, mitoXplorer 2.0 first embeds the mito-interactome into the entire cellular interactome and then allows to integrate differential expression data with this cellular interactome. The user can then explore the expression dynamics of the network neighbourhood of a single selected mito-gene and, based on differential expression data, identify the active subnetwork surrounding this gene. This function can be invoked from the interfaces: comparative plots, heatmap and time-course visualization. To perform a network analysis, the user must first choose from the ‘Integrative Analysis’ panel the network analysis function (see Supplementary Figure S2F). Then, they can select a mito-gene to start network analysis by clicking on the box of the gene in the (sortable) heatmap, the maximum number of steps from the starting point (mito-gene), the  $\log_2FC$  threshold (to define a gene as deregulated or not) and whether unregulated nodes in the results are allowed (see Supplementary Figures S2 and S4 for the network analysis menu; see also our mitoXplorer tutorials for further instructions). Results will be displayed directly in a new window in the mitoXplorer platform.

The network analysis function has been developed using the NetworkX Python library (15). The cellular interactomes were created from the STRING database (v11) (16) using high-confidence physical protein–protein interactions and have been added to the MySQL database. The network neighbourhood of a selected gene is explored by using the ‘environment search’ algorithm described by the viPER Cy-

toscape app (6). Once the analysis has been performed, the active subnetwork is drawn using the D3.js library (10).

**IDconverter.** In order to facilitate data upload, we added an IDconverter to allow usage of any gene identifier. This function converts input gene IDs (e.g. ENSEMBL, Entrez, RefSeq, Genecode, Flybase) on the fly to gene symbols, without any intervention from the user. The conversion is performed using the latest release of BioMart (11). BioMart data were downloaded and stored in text format. We retrieved them locally for reasons of speed.

**Upgrade to Python.** In order to enable local deployments and future development of the platform, we upgraded the backend of the web tool from Python 2.7 to Python 3.6.

The new mitoXplorer 2.0 menu is detailed in Supplementary Figure S2A–F; details on how to use new mitoXplorer 2.0 functions are shown in Supplementary Figures S3 and S4 and online at the mitoXplorer 2.0 tutorial pages (<http://mitoxplorer2.ibdm.univ-mrs.fr/tutorials>).

## Data processing for use cases 1 and 2

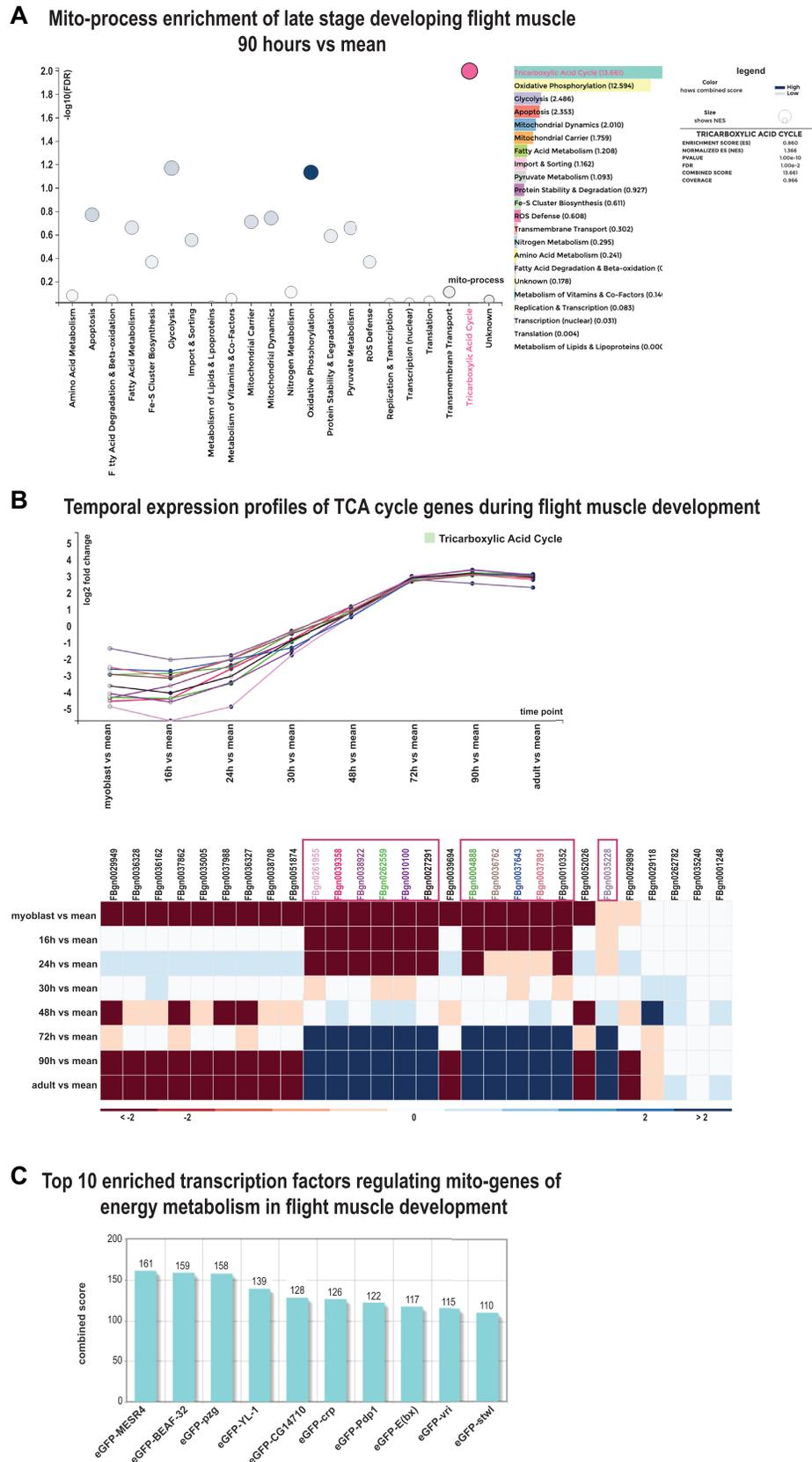
**Flight muscle data (use case 1).** Temporal expression data of developing *D. melanogaster* IFM were taken from GSE107247 (17) in the form of raw read counts. In order to obtain differential expression values for the different time points versus the mean over all time points, we created two pseudo-replicates of the mean by first calculating the mean over all time points using one replicate per time point each. Normalization of the raw read counts as well as pairwise differential expression analysis was performed using DESeq2 (18).

**Ataxia data (use case 2).** Data from the ataxia mouse model ATXN1.82Q.Tg from cerebellum were taken from GSE122099 (19). We downloaded sequencing reads and mapped them to mouse genome version mm10 using STAR (20) with default parameters. FeatureCounts (21) was used to calculate read counts. Normalization and differential expression analysis were done using DESeq2 via RN-fuzzyApp (22). We compared all conditions to each other: 5-week wild type (WT) versus ATXN1.82Q.Tg, 12-week WT versus ATXN1.82Q.Tg, 5- versus 12-week WT, and finally 5- versus 12-week ATXN1.82Q.Tg. We uploaded the resulting data to mitoXplorer 2.0. For enrichment analysis of genes found in the active subnetworks extracted by mitoXplorer 2.0, we first downloaded and parsed the respective JSON files to obtain gene lists. These were then submitted to EnrichR (7); for further downstream analysis, we used the KEGG resource (23).

## RESULTS

### Use case 1: identifying transcriptional regulators of mitogenes during *D. melanogaster* indirect flight muscle development

We analysed temporal mito-gene expression dynamics during *D. melanogaster* IFM development using the new mitoXplorer 2.0 functions. In the developing indirect flight muscle, mitochondria undergo significant structural,



**Figure 2.** mitoXplorer 2.0 analysis of *D. melanogaster* developing indirect flight muscle (IFM). (A) Mito-process enrichment analysis of 90 h IFM after puparium formation (APF) revealed an enrichment of tricarboxylic acid (TCA) cycle, as well as OXPHOS. Mito-processes with a high combined score

but also bioenergetic changes (17,24). We wanted to explore these mitochondrial changes further using mitoXplorer 2.0. We used a transcriptomic resource generated by Spletter *et al.*, consisting of transcriptomic data from *Drosophila* isolated IFM at myoblast stage, 16, 24, 30, 48, 72 and 90 h after puparium formation (APF), and the adult stage (17). To observe temporal changes during IFM development, we compared each time point to the mean over all time points (see the ‘Materials and Methods’ section) and uploaded the data to mitoXplorer 2.0.

We first investigated which mito-processes are enriched during IFM development and used the new mito-process enrichment function. We used nearly mature IFM at 90 h APF for mito-process enrichment analysis. We found mito-processes related to OXPHOS-dependent energy metabolism enriched at this time point (Figure 2A and Supplementary Figure S3A).

We next investigated the temporal expression profiles of mito-genes from the mito-processes OXPHOS, OXPHOS mt and TCA cycle using the time-course visualization function of mitoXplorer 2.0. As exemplified in Figure 2B, we found several mito-genes within these three mito-processes that showed a specific temporal expression profile, being first downregulated (myoblast stage, 16, 24 and 30 h APF) and subsequently strongly induced from 72 h APF to the adult stage (see also Supplementary Figure S3B).

We were interested in identifying potential TFs responsible for regulating these co-regulated genes. We therefore made use of mitoXplorer 2.0’s new integrative analysis function TF-enrichment. We selected all genes from the three mito-processes that showed the above-described temporal expression profile and uploaded them via mitoXplorer’s TF-enrichment function to AnnoMiner (see Supplementary Table S1A for the list of genes selected and Supplementary Figure S3C and D). AnnoMiner returned several interesting TFs as enriched in the promoters of the co-regulated mito-genes (Supplementary Table S1B). Among the top 10 hits are *MESR4*, which has been shown to be involved in development and in the cellular response to hypoxia (25); *pzg*, which is known to regulate developmental processes (26,27); *Pdp1*, which is known to regulate muscle genes (28); *crp*, which controls cell growth and tracheal terminal branching (29); or *vri*, which has been shown to be involved in tracheal development (30,31).

To conclude, this use case demonstrates how the new functions of mitoXplorer 2.0, mito-process enrichment, time-course visualization and the integrative analysis function TF-enrichment could help identify the gene expression dynamics behind the mitochondrial bioenergetic switch in developing *Drosophila* indirect flight muscle and predict potential transcriptional regulators responsible for this switch.

## Use case 2: identification of signalling cascades regulating Ca<sup>2+</sup> signalling in a mouse model of spinocerebellar ataxia type 1

Spinocerebellar ataxias (SCAs) are a group of dominantly inherited neurodegenerative diseases, which are defined by a loss of coordination of body movements and cerebellar degeneration. They can be caused by mutations in nearly 40 genes and are currently untreatable. A disruption of Ca<sup>2+</sup> signalling in cerebellar cells, and more specifically in the Purkinje neurons, is considered to play a key role in disease onset and progression (32). We wanted to explore the contribution of mitochondria to Ca<sup>2+</sup> signalling deregulation in SCAs. To this end, we used a mouse model of spinocerebellar ataxia type 1 (SCA1). SCA1 affects the cerebellum as well as the inferior olive and is caused by polyglutamine expansion in the *ATAXIN1* gene, which encodes a transcriptional regulator (33). We used RNA expression data from a mouse model of SCA1 from (19), which looked at cerebellum and inferior olive of 5- and 12-week-old transgenic ATXN1.82Q (ATXN1.82Q\_Tg) mice. We focused our analysis on the cerebellum and were mostly interested in the temporal differences of pathways possibly affecting mitochondrial Ca<sup>2+</sup> signalling and transport.

We uploaded differential expression data of ATXN1.82Q\_Tg compared against WT at 5 and 12 weeks, as well as the temporal comparisons of WT (5 and 12 weeks) and ATXN1.82Q\_Tg (5 and 12 weeks), to mitoXplorer 2.0. We used comparative plots to visualize differences in gene expression in ‘Ca<sup>2+</sup> signalling and transport’ in those four datasets. Among the genes most affected in this mito-process was *Itp1* (Figure 3A and Supplementary Figure S4A), an intracellular receptor for inositol 1,4,5-trisphosphate that is located at the endoplasmic reticulum (ER). *Itp1* is one of the main regulators of mitochondrial Ca<sup>2+</sup> signalling (34) and is itself regulated by mitochondria (35). Furthermore, mutations in *ITPR1* cause SCA15 (36).

We next wanted to know how the network neighbourhood and thus potential signalling pathways change between the two disease stages. We therefore made use of the new integrative network analysis function in mitoXplorer 2.0 (see Supplementary Figure S4B). We extracted active subnetworks surrounding the *Itp1* gene using differential expression data of 5- and 12-week-old ATXN1 mice, respectively, and then compared them first by visual inspection (Figure 3B and C; no deregulated nodes allowed, log<sub>2</sub>FC ≥ 11.21, step size 3). The extracted active subnetwork surrounding *Itp1* is substantially larger at 12 weeks with more differentially expressed nodes within a three-step neighbourhood (Figure 3C) compared to the earlier stage (Figure 3B). We then downloaded the two extracted, active subnetworks (Supplementary Figure S4C) for enrichment analysis using EnrichR and KEGG. Many of the en-

are highlighted in dark blue; the process TCA cycle has been selected and appears in pink. (B) Time-course visualization of developing *Drosophila* IFM, showing TCA cycle as an example. Genes that were downregulated early in development (myoblast, 30 h APF) and induced in late stages (72 h APF, adult) were selected for display in the line plot. Each visualized gene is shown in a different colour. (C) Top 10 enriched TFs potentially regulating mito-genes with the described expression profile (down early, up late) of the mito-processes OXPHOS, OXPHOS mt (mitochondrial-encoded genes) and TCA cycle. The plot is directly generated in AnnoMiner upon automatic upload of selected mito-genes within mitoXplorer 2.0’s time-course visualization function. See Supplementary Figure S3 as well as mitoXplorer’s video tutorials for instructions on how to generate these plots.



riched pathways overlap between the two time points, such as calcium signalling, neuroactive ligand–receptor interaction or phosphatidylinositol signalling. At 12 weeks, some more metabolic pathways as well as SCA appear to be enriched (Supplementary Table 2B–D). Moreover, three genes responsible for SCAs are found in the 12-week network, including *Itpr1* itself, together with *Trpc3* and *Cacna1g* (Supplementary Table S2E).

In conclusion, we used the comparative plots and network analysis functions to investigate  $\text{Ca}^{2+}$  signalling deregulation in an SCA1 mouse model and found several signalling pathways consistently induced from 5 to 12 weeks in the cerebellum. Moreover, two more genes related to SCA were found in the 12-week active subnetwork.

## DISCUSSION

Here, we present an important upgrade of the mitoXplorer visual data mining platform. While mitoXplorer version 1.0 (1) allowed to inspect mitochondrial datasets in multiple ways, mitoXplorer 2.0's integrative analysis functions permit the user to go beyond simple data visualization. Our new integrative functions can be used to better understand the interaction between mitochondria and their cellular environment, as we have shown in two of our use cases. In the first use case, we showed how the new functions of mitoXplorer 2.0 (the mito-process enrichment, time-course visualization and TF-enrichment) can be used to identify key transcriptional regulators of important mitochondrial processes. In the second one, we employed network analysis, which, by integrating the mitochondrial interactome into the complete cellular interactome, allows the identification of important non-mitochondrial genes that are connected with and potentially regulate mitochondrial processes. These results demonstrate the power of mitoXplorer 2.0's downstream and integrative analysis functions.

For TF-enrichment, we decided to rely on AnnoMiner (5). By connecting mitoXplorer 2.0 to AnnoMiner, direct data transfer from one to the other allows us to execute AnnoMiner's *TF enrichment analysis*, to identify potential transcriptional regulators of mitochondrial functions. To date, several tools for TF-enrichment are available, but to our knowledge only few of these, namely modEnrichr (37), iCisTarget (38) and AnnoMiner, work on model organisms other than human and mouse and thus can be used for three out of the four model organisms present in mitoXplorer 2.0: human, mouse and *D. melanogaster*. AnnoMiner has also data for *Caenorhabditis elegans*, a model species we plan to offer soon on mitoXplorer 2.0. For network analysis, we implemented a function for active subnetwork extraction starting from a single gene based on the algorithm of the Cytoscape app viPER (6). The exploration of the network neighbourhood of a gene using viPER requires only minimal user input: a selected gene; the number of steps from this

gene for extracting an active subnetwork; a  $\log_2\text{FC}$  threshold for defining genes as deregulated; and finally, whether to accept connecting genes that are not deregulated. The algorithm thus avoids cumbersome tuning of parameters. The fact that we implemented this algorithm using NetworkX (15) and D3.js (10) permits a fully interactive exploration of the extracted subnetwork, facilitating its interpretation. In the future, we plan to include downstream analysis of the resulting network, such as pathway annotation through enrichment analysis. This should help to better identify potential signalling paths regulating mitochondrial functions. We find these two integrative analysis functions as an important novel upgrade of mitoXplorer, which should help to unravel how mito-gene expression and thus mitochondrial functions adjust to a specific cellular environment.

Additional future developments we are currently working on include mitochondrial metabolic modelling using flux balance analysis as well as logical modelling. Combining mitoXplorer directly with metabolic modelling would allow users to assess the metabolic states of biological samples based on mito-gene expression data directly within the mitoXplorer platform. Starting mitochondrial models are available for human and budding yeast (39–41), which we will adapt to our needs. In addition, we plan to extend mitoXplorer to allow analysis of single-cell RNA-seq data of tissues on cell-type level. All these further implementations are simplified by the modular fashion in which mitoXplorer 2.0 has been developed.

## DATA AVAILABILITY

mitoXplorer 2.0 is available as a web server at <http://mitoxplorer2.ibdm.univ-mrs.fr>. The source code together with installation instructions is available in our GitLab repository ([https://gitlab.com/habermann\\_lab/MitoX2](https://gitlab.com/habermann_lab/MitoX2)).

Public gene expression data used for this study were obtained from the Gene Expression Omnibus repository at NCBI (42). *Drosophila* developmental IFM time-course data have been downloaded from GSE107247 (17). Mouse ATXN1 raw sequencing data have been downloaded from GSE122099 (19).

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank Alice Carrier, Benoit Giannesini, Frank Schnorner, Nuno Luis, Jerome Avellaneda, and the members of the IBDM Computational Biology team for helpful discussion on mitoXplorer 2.0. We thank Stephen Chapman for critical reading of the manuscript.

*Author contributions:* The upgrade of the mitoXplorer web server was developed by F.M. Data analysis was done by

(C) Network neighbourhood of *Itpr1* at 12 weeks. Genes found to cause a form of SCA are marked with an asterisk. Mito-genes are labelled in cyan, while non-mitochondrial genes are labelled in black. The user can interactively change gene positions and labels in order to show different information (gene name or ID, mutation,  $\log_2\text{FC}$ ). The visualization is information dense. Node size, colour and border represent magnitude of  $\log_2\text{FC}$ , direction of the deregulation (up or down) and mutation information, when available. See Supplementary Figure S4 as well as mitoXplorer's video tutorials for instructions on how to generate these plots.

B.H.H., F.M. and M.H. The project was conceived by B.H.H. The manuscript was written by F.M. and B.H.H with input from M.H. All authors read and approved the final version of the manuscript.

## FUNDING

Agence Nationale de la Recherche [ANR-18-CE45-0016-01 to B.H.H.]; Fondation pour la Recherche Médicale [MND202003011460 to B.H.H.]; Centre National de la Recherche Scientifique; Aix-Marseille University. Funding for open access charge: Agence Nationale de la Recherche [ANR-18-CE45-0016-01]; Fondation pour la Recherche Médicale [MND202003011460].

*Conflict of interest statement.* None declared.

## REFERENCES

1. Yim, A., Koti, P., Bonnard, A., Marchiano, F., Dürrbaum, M., Garcia-Perez, C., Villaveces, J., Gamal, S., Cardone, G., Perocchi, F. *et al.* (2020) mitoXplorer, a visual data mining platform to systematically analyze and visualize mitochondrial expression dynamics and mutations. *Nucleic Acids Res.*, **48**, 605–632.
2. Friedman, J.R. and Nunnari, J. (2014) Mitochondrial form and function. *Nature*, **505**, 335–343.
3. Rath, S., Sharma, R., Gupta, R., Ast, T., Chan, C., Durham, T.J., Goodman, R.P., Grabarek, Z., Haas, M.E., Hung, W.H.W. *et al.* (2021) MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and pathway annotations. *Nucleic Acids Res.*, **49**, D1541–D1547.
4. Glancy, B. (2020) Visualizing mitochondrial form and function within the cell. *Trends Mol. Med.*, **26**, 58–70.
5. Meiler, A., Marchiano, F., Haering, M., Weitkunat, M., Schnorrer, F. and Habermann, B.H. (2021) AnnoMiner is a new web-tool to integrate epigenetics, transcription factor occupancy and transcriptomics data to predict transcriptional regulators. *Sci. Rep.*, **11**, 15463.
6. Garmhausen, M., Hofmann, F., Senderov, V., Thomas, M., Kandel, B.A. and Habermann, B.H. (2015) Virtual pathway explorer (viPEr) and pathway enrichment analysis tool (PEANuT): creating and analyzing focus networks to identify cross-talk between molecules and pathways. *BMC Genomics*, **16**, 790.
7. Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A. *et al.* (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.*, **44**, W90–W97.
8. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. U.S.A.*, **102**, 15545–15550.
9. Xiao, Y., Hsiao, T.-H., Suresh, U., Chen, H.-I.H., Wu, X., Wolf, S.E. and Chen, Y. (2014) A novel significance score for gene selection and ranking. *Bioinformatics*, **30**, 801–807.
10. Bostock, M., Ogievetsky, V. and Heer, J. (2011) d<sup>3</sup>: data-driven documents. *IEEE Trans. Vis. Comput. Graph.*, **17**, 2301–2309.
11. Durinck, S., Spellman, P.T., Birney, E. and Huber, W. (2009) Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.*, **4**, 1184–1191.
12. ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57–74.
13. Muers, M. (2011) Functional genomics: the modENCODE guide to the genome. *Nat. Rev. Genet.*, **12**, 80.
14. Kudron, M.M., Victorsen, A., Gevirtzman, L., Hillier, L.W., Fisher, W.W., Vafeados, D., Kirkey, M., Hammonds, A.S., Gersch, J., Ammouri, H. *et al.* (2018) The modERN resource: genome-wide binding profiles for hundreds of *Drosophila* and *Caenorhabditis elegans* transcription factors. *Genetics*, **208**, 937–949.
15. Hagberg, A., Swart, P. and Schult, D. (2008) *Exploring network structure, dynamics, and function using NetworkX*. Los Alamos National Lab, Los Alamos, NM.
16. Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legeay, M., Fang, T., Bork, P. *et al.* (2021) The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.*, **49**, D605–D612.
17. Spletter, M.L., Barz, C., Yeroslaviz, A., Zhang, X., Lemke, S.B., Bonnard, A., Brunner, E., Cardone, G., Basler, K., Habermann, B.H. *et al.* (2018) A transcriptomics resource reveals a transcriptional transition during ordered sarcomere morphogenesis in flight muscle. *eLife*, **7**, e34058.
18. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
19. Driessen, T.M., Lee, P.J. and Lim, J. (2018) Molecular pathway analysis towards understanding tissue vulnerability in spinocerebellar ataxia type 1. *eLife*, **7**, e39981.
20. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
21. Liao, Y., Smyth, G.K. and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.
22. Haering, M. and Habermann, B.H. (2021) RNfuzzyApp: an R shiny RNA-seq data analysis app for visualisation, differential expression analysis, time-series clustering and enrichment analysis. *FI000Research*, **10**, 654.
23. Kanehisa, M., Sato, Y. and Kawashima, M. (2022) KEGG mapping tools for uncovering hidden features in biological data. *Protein Sci.*, **31**, 47–53.
24. Avellaneda, J., Rodier, C., Daian, F., Brouilly, N., Rival, T., Luis, N.M. and Schnorrer, F. (2021) Myofibril and mitochondria morphogenesis are coordinated by a mechanical feedback mechanism in muscle. *Nat. Commun.*, **12**, 2091.
25. Lee, S.-J., Feldman, R. and O'Farrell, P.H. (2008) An RNA interference screen identifies a novel regulator of target of rapamycin that mediates hypoxia suppression of translation in *Drosophila* S2 cells. *Mol. Biol. Cell*, **19**, 4051–4061.
26. Kugler, S.J. and Nagel, A.C. (2007) putzig is required for cell proliferation and regulates notch activity in *Drosophila*. *Mol. Biol. Cell*, **18**, 3733–3740.
27. Zimmermann, M., Kugler, S.J., Schulz, A. and Nagel, A.C. (2015) Loss of putzig activity results in apoptosis during wing imaginal development in *Drosophila*. *PLoS One*, **10**, e0124652.
28. Lin, S.C., Lin, M.H., Horváth, P., Reddy, K.L. and Storti, R.V. (1997) PDP1, a novel *Drosophila* PAR domain bZIP transcription factor expressed in developing mesoderm, endoderm and ectoderm, is a transcriptional regulator of somatic muscle genes. *Development*, **124**, 4685–4696.
29. Wong, M.M.-K., Liu, M.-F. and Chiu, S.K. (2015) Cropped, *Drosophila* transcription factor AP-4, controls tracheal terminal branching and cell growth. *BMC Dev. Biol.*, **15**, 20.
30. Gunawardhana, K.L., Rivas, G.B.S., Caster, C. and Hardin, P.E. (2021) Crosstalk between vrille transcripts, proteins, and regulatory elements controlling circadian rhythms and development in *Drosophila*. *iScience*, **24**, 101893.
31. Szuplewski, S., Fraise-Véron, I., George, H. and Terracol, R. (2010) vrille is required to ensure tracheal integrity in *Drosophila* embryo. *Dev. Growth Differ.*, **52**, 409–418.
32. Paulson, H.L. (2009) The spinocerebellar ataxias. *J. Neuroophthalmol.*, **29**, 227–237.
33. Orr, H.T., Chung, M.Y., Banfi, S., Kwiatkowski, T.J.J., Servadio, A., Beaudet, A.L., McCall, A.E., Duvick, L.A., Ranum, L.P. and Zoghbi, H.Y. (1993) Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat. Genet.*, **4**, 221–226.
34. Bezprozvanny, I. and Hayden, M.R. (2004) Deranged neuronal calcium signaling and Huntington disease. *Biochem. Biophys. Res. Commun.*, **322**, 1310–1317.
35. Parekh, A.B. (2003) Mitochondrial regulation of intracellular Ca<sup>2+</sup> signaling: more than just simple Ca<sup>2+</sup> buffers. *News Physiol. Sci.*, **18**, 252–256.

36. van de Leemput, J., Chandran, J., Knight, M.A., Holtzclaw, L.A., Scholz, S., Cookson, M.R., Houlden, H., Gwinn-Hardy, K., Fung, H.-C., Lin, X. *et al.* (2007) Deletion at ITPR1 underlies ataxia in mice and spinocerebellar ataxia 15 in humans. *PLoS Genet.*, **3**, e108.
37. Kuleshov, M.V., Diaz, J.E.L., Flamholz, Z.N., Keenan, A.B., Lachmann, A., Wojciechowicz, M.L., Cagan, R.L. and Ma'ayan, A. (2019) modEnrichr: a suite of gene set enrichment analysis tools for model organisms. *Nucleic Acids Res.*, **47**, W183–W190.
38. Herrmann, C., Van de Sande, B., Potier, D. and Aerts, S. (2012) i-cisTarget: an integrative genomics method for the prediction of regulatory features and *cis*-regulatory modules. *Nucleic Acids Res.*, **40**, e114.
39. Smith, A.C. and Robinson, A.J. (2011) A metabolic model of the mitochondrion and its use in modelling diseases of the tricarboxylic acid cycle. *BMC Syst. Biol.*, **5**, 102.
40. Smith, A.C., Eyassu, F., Mazat, J.-P. and Robinson, A.J. (2017) MitoCore: a curated constraint-based model for simulating human central metabolism. *BMC Syst. Biol.*, **11**, 114.
41. Malina, C., Di Bartolomeo, F., Kerkhoven, E.J. and Nielsen, J. (2021) Constraint-based modeling of yeast mitochondria reveals the dynamics of protein import and iron–sulfur cluster biogenesis. *iScience*, **24**, 103294.
42. Clough, E. and Barrett, T. (2016) The gene expression omnibus database. *Methods Mol. Biol.*, **1418**, 93–110.