OPEN

FlySilico: Flux balance modeling of *Drosophila* larval growth and resource allocation

Jürgen Wilhelm Schönborn[1,2,4], Lisa Jehrke[1,2,4], Tabea Mettler-Altmann[3] & Mathias Beller [1,2*]

Organisms depend on a highly connected and regulated network of biochemical reactions fueling life sustaining and growth promoting functions. While details of this metabolic network are well established, knowledge of the superordinate regulatory design principles is limited. Here, we investigated by iterative wet lab and modeling experiments the resource allocation process during the larval development of *Drosophila melanogaster*. We chose this system, as survival of the animals depends on the successful allocation of their available resources to the conflicting processes of growth and storage metabolite deposition. First, we generated "FlySilico", a curated metabolic network of *Drosophila*, and performed time-resolved growth and metabolite measurements with larvae raised on a holidic diet. Subsequently, we performed flux balance analysis simulations and tested the predictive power of our model by simulating the impact of diet alterations on growth and metabolism. Our predictions correctly identified the essential amino acids as growth limiting factor, and metabolic flux differences in agreement with our experimental data. Thus, we present a framework to study important questions of resource allocation in a multicellular organism including process priorization and optimality principles.

Balancing limited resources to concurrent processes is an essential task in all areas of life. Every organism, for example, needs to allocate its available resources – mostly in the form of diet-derived nutrients – to concurrent processes such as live sustaining functions, reproduction, or storage metabolite synthesis. While the metabolic pathways involved in these processes, as well as some regulatory signaling pathways, are well established, the overarching design principles governing resource allocation and prioritization are elusive[1,2]. This is especially true for multicellular heterotrophic organisms, which often have an almost unlimited amount of destinations for channeling their available resources. On top of a plethora of energy consuming processes available to the whole organism (physical movement, growth and reproduction, energy storage metabolite deposition), higher order multicellular organisms are composed of a multitude of organs. This results in an even higher complexity, given that many organs have distinct and different metabolic preferences, which is important during health and disease states[3–5]. The mammalian brain, for example, depends on sugars for energy production, whereas most other body cells can additionally utilize other energy liberating pathways such as fatty acid beta-oxidation. To this end, a bottom-up understanding of the resource allocation regulation therefore appears close-to impossible based on this complexity and lack of detailed information. Thus, an abstraction in the form of a top-down modeling paradigm has the potential to reveal design principles, which serve as a starting point to investigate resource allocation principles. Yet, the various degrees of freedom available to complex organisms complicate the modeling procedures, as model solving usually targets the optimization (maximizing or minimizing) of a distinct objective function such as e.g. biomass production or growth[6]. Given that multicellular organisms can have multiple and often conflicting objective functions (e.g. reproduction, longevity), or objective functions without the aim of maximization or minimization (e.g. metabolic processes to sustain survival, growth in terms of sustaining healthy cell turnover, deposition of energy depots), the identification of a single and clear-cut objective function is difficult.

Development, however, appears to represent an exception. The development of most organisms follows a stereotyped program, which involves hard constraints e.g. in terms of the timing or metabolic thresholds. Holometabolous insects, for instance, need to deposit sufficient energy storage amounts to allow metamorphosis

[1]Institute for Mathematical Modeling of Biological Systems, Heinrich Heine University, Duesseldorf, Germany. [2]Systems Biology of Lipid Metabolism, Heinrich Heine University, Duesseldorf, Germany. [3]Institute of Plant Biochemistry & Cluster of Excellence on Plant Sciences, Heinrich Heine University, Duesseldorf, Germany. [4]These authors contributed equally: Jürgen Wilhelm Schönborn and Lisa Jehrke. *email: mathias.beller@hhu.de
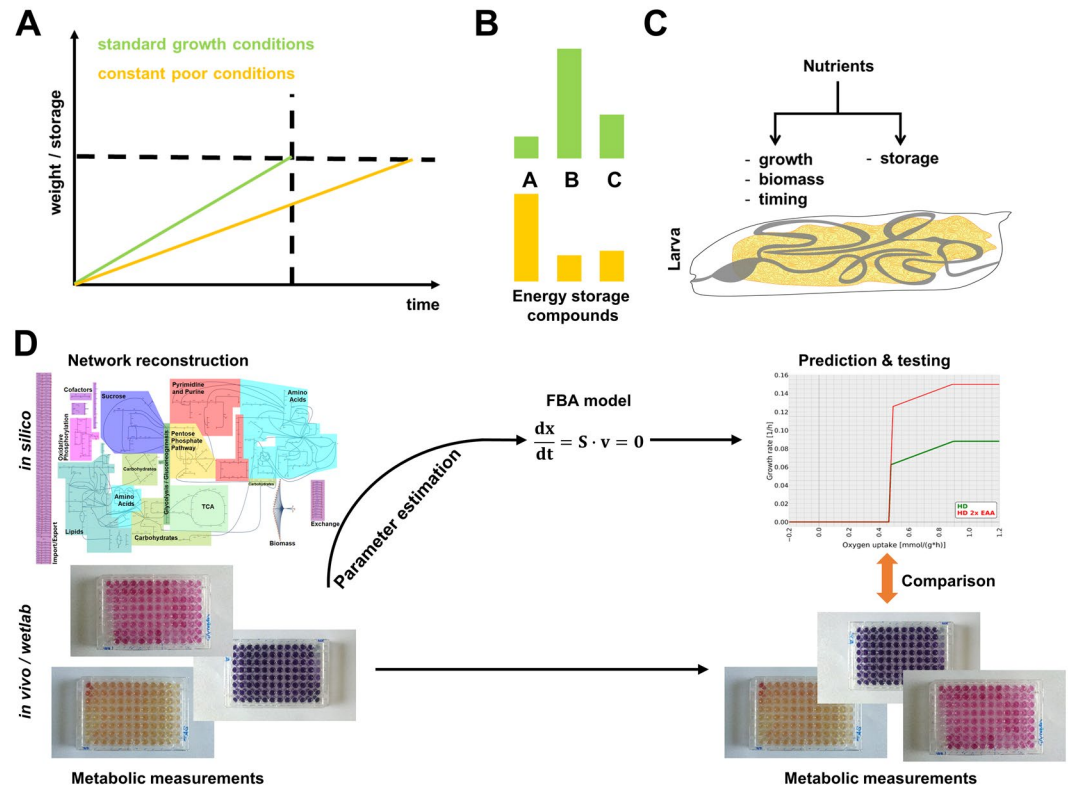
Figure 1. Plasticity of resource allocation and experimental design of the study. (**A**) Development involves growth and weight gain (in part due to the deposition of storage metabolites) over time. Altered environmental conditions, such as a rich (green) or poor (orange) nutrition results in an altered timing and/or (**B**) altered energy storage compound levels (compounds **A**–**C** show a different relative abundance under rich (green) and poor (orange) nutritional conditions). (**C**) Nutrients fuel conflicting processes. *Drosophila* larvae need to channel available nutrients to sustain either growth or the formation of storage metabolites. Following nutrient uptake in the gut (shown in grey) copious amounts of triglyceride and glycogen stores are built in the fat body (yellow), which largely fills the body cavity of developing larvae. (**D**) Our experimental design involved *in silico* and *in vivo* (wet lab) experiments. For the *in silico* studies, we first reconstructed main parts of the *Drosophila* metabolic network. For the parameter estimation, we performed metabolic profiling experiments during the *Drosophila* larval development. Subsequently, the metabolic network served the constraint-based flux balance modeling (FBA). With the FBA models, we predicted the consequences of alterations of the dietary composition and validated the modeling results using targeted experiments. High resolution versions of the metabolic network shown in (**D**) are provided in Fig. 2 as well as an interactive version of the metabolic network as supplemental data.

into an adult organism, which is often associated with a minimal weight, which is also termed critical weight in insects such as *Drosophila melanogaster*[7]. Given the variations many organisms face during development – based on e.g. fluctuation in temperature or food supply and quality – developmental programs are at the same time also flexible (Fig. 1A,B). The time needed to reach a certain weight or storage threshold necessary to complete development thus can be extended under poor nutritional conditions, for example (Fig. 1A). Besides adjusting the developmental timing, organisms can also alter the relative amounts of their energy storage compounds (Fig. 1B). Thus, development of multicellular organisms should be accessible to modeling campaigns, as there is a clear-cut objective function (growth) and sufficient plasticity (timing and metabolic fine-tuning) to test modeling predictions. A better knowledge of the resource allocation principles potentially answers the question whether the developmental growth and resource allocation are in multicellular organisms also (pareto-)optimal. A pareto-optimality was previously reported for bacterial growth and metabolism[8]. Pareto-optimality was also proposed for certain phenotypic traits of multicellular organisms[9,10], yet this is still under discussion[11,12].

Here, we use the larval development of *Drosophila melanogaster* as a model system for the analysis of the resource allocation of a multicellular organism using *in silico* and wet lab experiments (Fig. 1C,D). This system is particularly well suited for this endeavor based on the following key-points: (i) Larval development of *Drosophila* involves a massive increase in size and weight coupled to the deposition of copious energy stores necessary to fuel metamorphosis (Fig. 1C)[13]. In order to allow this massive size increase, larvae are constantly eating[14], which facilitates the estimation of energy expenditure and metabolite intake. (ii) *Drosophila* larval development and resource allocation shows an inherent plasticity. Poor nutritional conditions, for example, result in a prolonged developmental timing based on a lowered rate of development (Fig. 1A)[15]. This change of developmental timing is
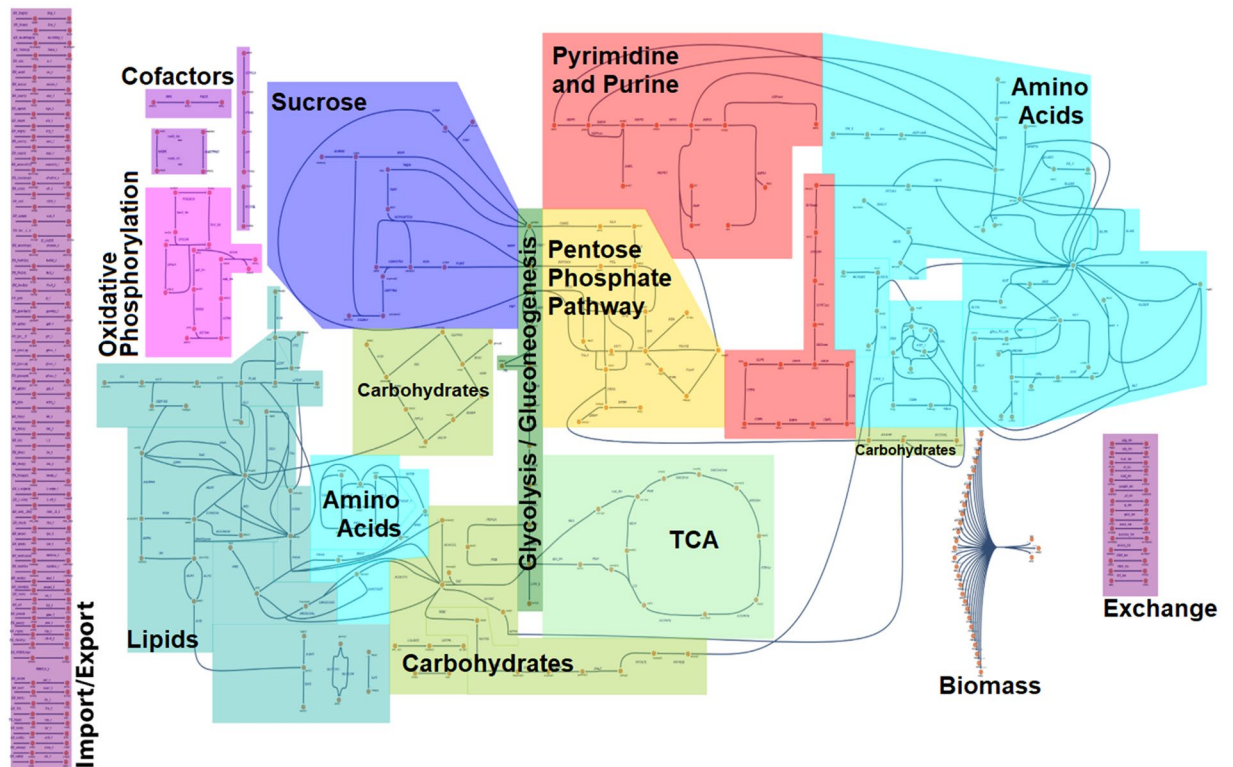
Figure 2. Map of the FlySilico V1.0 metabolic model of *Drosophila melanogaster*. On top of the central carbon metabolism, we included reaction complexes such as the lipid or amino acid metabolic pathways. A main goal was to include the pathways necessary for metabolizing the main constituents of the holidic diet and the pathways covering our experimentally quantified metabolites. For details, please see main text, Table S1 and Interactive Supplementary Fig. 1.

often paralleled by an altered metabolite composition of the organism[16–19]. Intriguingly, also e.g. mammals adapt their metabolism to *in utero* nutritional alterations[20] and the impact of the nutritional status during development and early life on the later stages is well described in terms of the life history theory[21,22] and the process of metabolic programming[23–25]. (iii) For the modeling procedures, we benefitted from a previously published chemically defined, fully synthetic minimal *Drosophila* food (*Holidic diet*; HD)[26], which allowed us to clearly define the input of our model. On top of the general advantages of working with small animals, such as the large number of progeny and accessibility of the developmental stages, these characteristics clearly facilitated our investigations.

To target resource allocation in *Drosophila* larval development, we generated the – to the best of our knowledge – to date largest curated metabolic network for fruit flies and subjected it to flux balance analyses. To validate and optimize our metabolic network, we used targeted metabolite quantifications and GC-coupled mass spectrometry metabolomics measurements of larvae grown on the HD. We built the model capturing previously known requirements of *Drosophila* metabolism, such as sterol auxotrophy[27], and tailored it to incorporate all prominent ingredients of the HD. Our model predictions allowed us to correctly identify the amount of essential amino acids as growth limiting factor. Further, the model predictions resulted in flux differences, which are in line with the measured metabolite alterations associated with growth of the larvae in food with elevated amounts of sucrose or essential amino acids. These proof-of-principle experiments provide a starting point to investigate the optimality principles of multicellular resource allocation.

Results

Drosophila metabolic network reconstruction.
In order to model *Drosophila* larval growth and resource allocation, we first constructed a flux balance capable metabolic network covering the biochemical pathways necessary to metabolize the major constituents of the minimal, synthetic medium (*Holidic diet*; HD)[26], which we used to grow the fruit flies during the wet lab procedures. On top of the central carbon metabolism, we therefore included e.g. amino acid, lipid, and carbohydrate metabolism (Fig. 2, Table S1, and Interactive Supplementary Fig. 1). In total, our model – termed FlySilico – covers 363 reactions and 293 metabolites. To date, there are surprisingly only two other *Drosophila* metabolic networks available. The first one focuses on the effects of hypoxia on ATP production[28–30]. The other one is a whole-genome computer generated model, which lacks curation (BMID000000141998; https://www.ebi.ac.uk/biomodels-main/BMID000000141998).

For our model reconstruction, we started from scratch and emphasized on avoiding biologically unfeasible reactions (dead-end, blocked, and unbalanced reactions) as well as on minimizing the number of exchange reactions (see methods section). Figure 3 shows a comparison between different aspects of our FlySilico and

| Organism	Model ID	# Reactions	# Metabolites	% dead-end reactions	% blocked reactions	% unbalanced reactions	% exchange reactions
E. coli	e_coli_core	95	72	0	0	1.1	21.1
D. melanogaster	FlySilico_v1	363	293	2.5	3.6	0.6	17.6
S. cerevisiae	iMM904	1577	1226	13.5	35.6	3.2	10.4
E. coli	iWFL_1372	2782	1973	9.5	13.8	1.4	14.1
Mouse	iMM1415	3726	2775	13.9	21.3*	0	12.2
Human	RECON1	3741	2766	11.6	34.1*	1.5	11.5
D. melanogaster	BMID000000141998	6198	2873	0.2	0.2*	6.6	46.2

Figure 3. Comparison of FlySilico V1.0 to other publically available metabolic networks. Dead-end reactions result in the production of metabolites, which are not further utilized in the network. Blocked-reactions are not accessible during the solving of the network. Unbalanced reactions are violating the conservation of mass and/or charge. Exchange reactions represent either biologically necessary transporters (such as for the import of nutrients into the system, or naturally occurring transporters for the export of end products), or transport reactions necessary for the modeling, to e.g. eliminate metabolites which are not further processed as the necessary biochemical reactions were omitted from the model. The values marked by an asterisk were calculated using non-loopless conditions. All other values are determined using loopless computations.

a selection of previously published other FBA-models of different organisms including the whole genome *Drosophila* model. While of course still limited in size, our model has a low amount of biologically unfeasible reactions. The importance of this became evident when we performed simulations with the whole genome, *in silico* generated *Drosophila* metabolic network. Here, simulations resulted in a positive Biomass production (Fig. S1) even without any inputs entering the model; i.e. that this model allows perpetual motion. For FlySilico, we did not detect such an artificial and erroneous behavior (data not shown).

Development of a Drosophila biomass function based on experimental data. Given our aim to investigate growth and resource allocation, we established the parameters of our model by incorporating experimental data. For this, we grew wildtype Oregon-R *Drosophila* animals on the HD. The complete larval development (until prepupae emerge) appears quasi linear and takes on the HD about 170 hours. In order to follow the development and metabolite profile over time, we collected larvae at three equally spaced time points during development (96, 132 and 168 hours after egg laying; AEL). To determine growth progression over time, we measured the wet and dry weight (Fig. 4A) as well as different size parameters (Fig. S2) of the larvae at the different time points. Larval weight increased almost linear over time (Fig. 4A). The water content was stable with values between 85 and 89% (Fig. 5A) and unaffected by alterations of the food composition (data not shown). For all time points, we performed absolute quantifications of free protein, glycogen, glucose, triacylglycerol (TAG), lactate, and glycerol (Fig. 4B–G) levels according to established protocols[31] (see Table S2). Further, we quantified various metabolites of the central carbon metabolism as well as free amino acids by GC-MS metabolomics measurements and external standard curves (Figs 4H,I, S3 and Table S2). Most per animal normalized measurements increased over time, as expected (Fig. 4). Only lactate levels reached a plateau after 132 hours of development. All in all, our measurements explained on average 79% (for 96 h AEL: ~81%, for 132 h AEL: ~96% and for 168 h AEL: ~60%) of the total dry weight with proteins and TAGs being the major contributors (Fig. 5B). Of course, we can not rule out that the larval midgut contained metabolites from the HD, which we also included in our measurements. Yet, while none of our targeted metabolite measurements covered compounds present in the HD, those represented the most abundant constituents of the larval biomass. The metabolites quantified by the GC-MS metabolomics strategy – which covered also metabolites present in the HD as e.g. the singular amino acids – were only present in minute amounts. Thus, our measurements covered a large part of the body composition and should provide a sufficient approximation to model larval growth.

In order to perform growth rate simulations, we formulated a biomass function based on the previously reported yeast biomass function of the model iMM904[32]. Yet, from the original yeast biomass function, we only utilized the value for the growth associated maintenance (GAM) costs, as this measure is difficult if not impossible to obtain for a multicellular organism. The other coefficients of the biomass function are based on our own measurements (Fig. 5C and methods section). Next, we used the HD food ingredients as constraints for the model solution procedure. Although we already knew the exact composition of the food, we still needed to approximate the amount of food consumed by a single larva over time. Given that the measurement of the amount of internalized solid food is difficult, we decided to follow a theoretical approach (for details see methods section). For this purpose, we used data available for the average number of mouth hook movements ("bites") per minute of larvae. The bites per minute only show low variability across different food compositions[33,34], suggesting that this assumption is reasonable. Further, we approximated the volume of the mouth of the larva based on previous[35] and own measurements (Fig. S2; approximated volume of the mouth = 0.011 mm³). Calculations based on both parameters resulted in an amount of 0.064 g/h food consumed per hour. We used this approximated food intake amount to calculate for each food ingredient the maximum amount consumed per hour. Of course, the calculated amount of consumed food is likely a prominent overestimation. For example, the larvae will most likely not fill their mouth completely with every bite. Further, not all food ingredients passage the gut barrier with 100% efficiency and at this point, we do not know the resorption rate for the different nutrients. Thus, we sought to identify a correction factor to limit the nutrient influxes to a reasonable level. For this purpose, we solved the
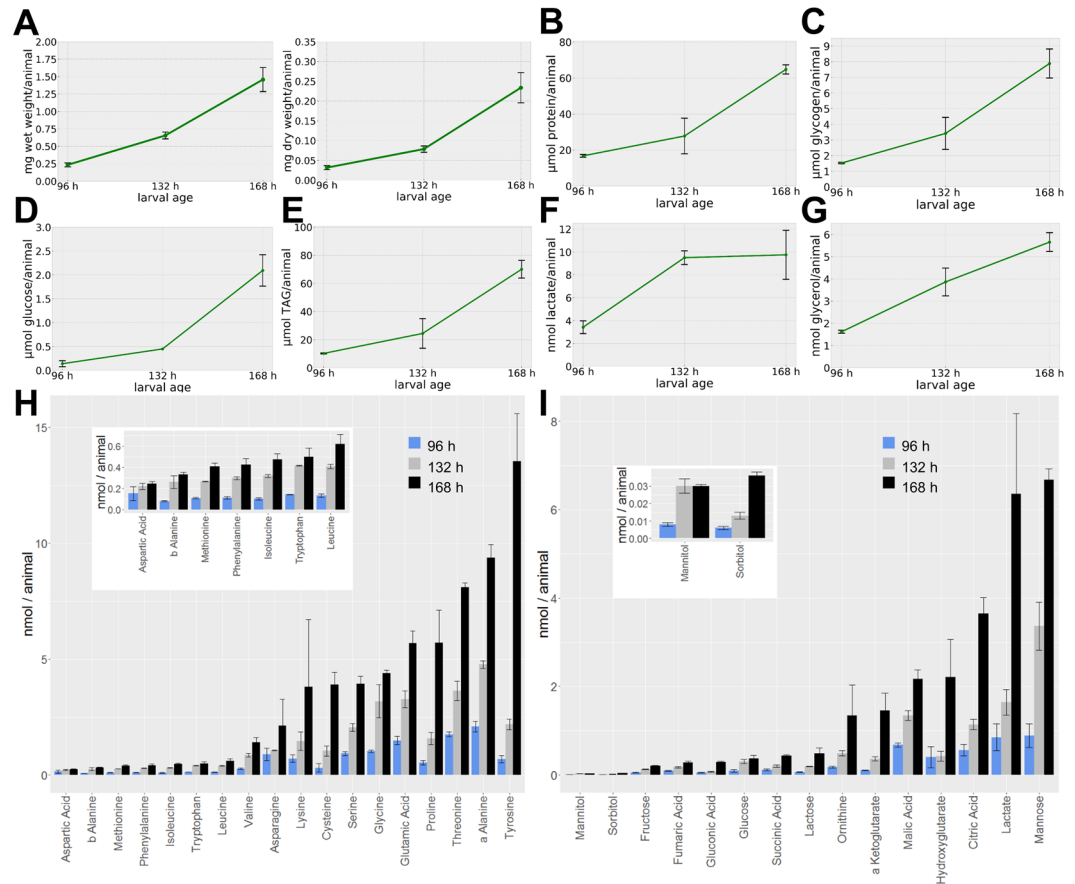
Figure 4. Growth and metabolic profiling of the larval developmental of *Drosophila melanogaster* 96, 132, and 168 hours after egg laying. (**A**) Wet weight (left plot) and dry weight (right plot) measurements of larvae at the indicated developmental time points. (**B**–**G**) Absolute quantification of protein, glycogen, glucose, triglyceride (TAG), lactate, and glycerol levels. The data represent mean values ± standard deviation normalized to the amount of animals per sample of at least triplicate measurements. (**H**,**I**) GC-MS metabolomics measurements of proteinogenic amino acids (**H**) and different metabolites (**I**) of *Drosophila* larval extracts from the indicated time points. Insets in (**H**,**I**) provide a zoom-in view on the low-abundant metabolite data. Metabolites were quantified using five point calibration curves (see methods, Fig. S3, and Table S2) and sorted for increasing abundance during the 168 hours time point.

FBA-model with a wide range of diminishing correction factors (Fig. S4) and compared the resulting calculated growth rates with our experimentally determined growth rate of approximately 8.8% dry weight increase per hour (see methods section). In this way, we identified a correction factor of the calculated food and thus nutritional uptake of about 12.2%. So far, little data concerning the food conversion and uptake rates of insects are available. Yet, Waldbauer *et al.* measured the efficiency of conversion of ingested food to body substance for various *Lepidoptera* species and found conversion rates ranging from 2 to 38%[36]. Our theory-derived value of 12.2% thus is in the range of the experimentally determined values of other insects.

Model verification. The first step in our model verification procedure was to test whether it operates in a reasonable manner and whether it recapitulates known behaviors of the fly system, in contrast to e.g. the computer-generated model mentioned above. *Drosophila*, for example, is sterol auxotroph[27]. We built our model to recapitulate this behavior and indeed a steep increase in the growth rate for positive, non-zero sterol uptake rates is visible (Fig. 6A). In contrast, the amount of the non-essential amino acid aspartic acid had no effect on the growth rate, as expected (Fig. 6B).

As a next step, we performed more complex simulations. First, we investigated growth in response to varying oxygen levels. Here, a certain minimal oxygen influx was needed to support suboptimal growth before increasing oxygen levels resulted in a plateau of the growth rate (Fig. 6C). As a test for the predictive power of our system, we decided to test next whether we could predict the impact of diet alterations on the growth and metabolism. While a reduction of certain nutrients would have been possible, we decided to rather test for a possible limitation of certain nutrients given that the growth of the larvae was on the HD much slower as compared to a complex and rich diet. Based on the growth properties of the animals, the HD in fact is classified as a minimal medium, which was designed to mirror dietary restriction characteristics[26]. Thus, we increased either the amount of dietary sugar or essential amino acids that was possible to enter the model. A doubling of the sucrose input limit had no effect on the calculated growth rate (Fig. 6C). However, when we doubled the amount of essential amino acids (EAA)
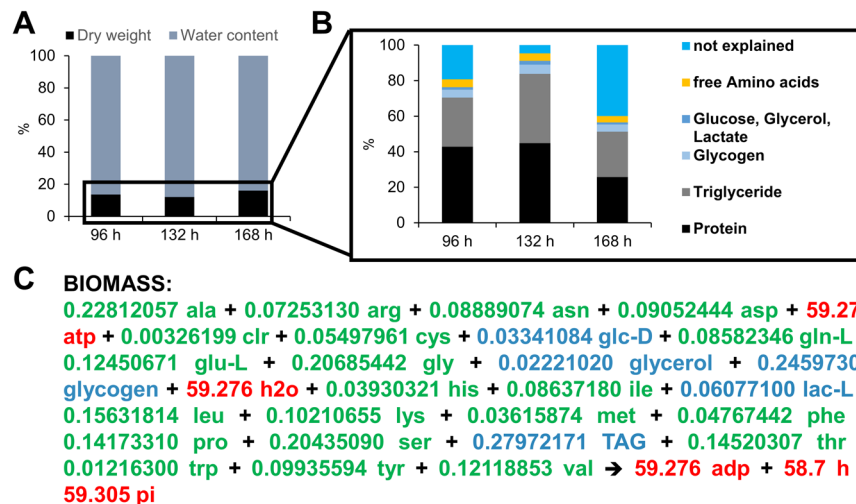
Figure 5. *Drosophila* body composition and biomass function. (**A**) Water content of *Drosophila* larvae raised on the HD during the three investigated time points. The black part of the stacked bar plots shows the dry mass of the larvae. (**B**) Our combined targeted and GC-MS metabolite measurements explain on average about 79% of the dry mass. Triglycerides and protein are the main contributors to dry mass. (**C**) *Drosophila melanogaster* biomass function based on our experimental data and literature. Green values indicate indices based on the GC-MS measurements and on the absolute biochemical metabolite quantifications, blue indices are based on the absolute biochemical metabolite quantifications. Red indices are based on information from the yeast whole-genome FBA model iMM904[32].

potentially entering the system, the predicted growth rate prominently increased (Fig. 6D). We subsequently tested our modeling predictions by performing corresponding experiments. Larvae reared on a HD containing the double amount of sucrose ("2x sucrose food") indeed did not show an increased growth-associated weight gain as compared to larvae raised on the standard HD (Fig. 6E). Protein levels were even lower than the ones of the control animals at the middle time point (Fig. 6F). Doubling the amount of EAA ("2x EAA food"), however, indeed resulted in a higher growth rate (Fig. 6E) as well as higher amounts of protein (Fig. 6F). Thus, our modeling data are consistent with the experimental results, which suggests a high predictive power of the model.

Modeling resource allocation. As a next step, we investigated whether the model could recapitulate resource allocation differences driven by the increase of sucrose or EAA in the HD. Therefore, we performed a flux variability analysis for our model with the given elevated maximum input limits (Table S3). Subsequently, we percent normalized the maximum and minimum flux values obtained for the 2x sucrose or 2x EAA food, respectively, based on the flux variability values of the standard HD (Table S4). Further, we split the metabolic reactions into functional groups and plotted selected reactions with altered fluxes (Fig. 7). On the simulated 2x sucrose diet, most reactions of the central carbon metabolism (Fig. 7A) showed a prominently increased maximum flux rate (e.g. FBA = fructose-bisphosphate aldolase, GAPD = glyceraldehye-3-P dehydrogenase, PFK = phosphofructokinase, PGI = glucose-6-P isomerase). On the simulated 2x EAA diet, however, most maximum flux rates did not change, and only the lower flux rate bounds were increased, thus resulting in a more narrow range of possible flux variations (Fig. 7A). For few reactions, the diet alterations resulted in opposite flux changes (HEX1 = hexokinase 1, LDH_L = l-lactate dehydrogenase, PPPH = diphosphate phosphohydrolase, PRPPS = ribose-phosphate diphosphokinase, PhnN = ribose 1,5-bisphosphate phosphokinase, R1Pk = Ribose 1-phosphokinase). The diphosphate phosphohydrolase (PPPH) flux showed a largely increased minimal flux on the simulated 2x EAA diet (Fig. 7A). Diphosphate phosphohydrolase activity takes place very early in the lipid degradation, as it acts as the force to activate fatty acids for the beta-oxidation[37]. The higher minimal flux following the elevated EAA input, suggested an enhanced rate of lipid activation, which potentially fuels the elevated growth. Further, the increased flux of the lactate dehydrogenase suggested increased lactate levels of animals reared on the diet with 2x EAA.

As expected, fluxes of reactions involved in sucrose metabolism (AF6P = ATP:D-fructose 6-phosphotransferase, AGMH1 = 1,4-alpha-D-Glucan maltohydrolase AMYTRA = 1,4-alpha-D-Glucan:1,4-alpha-D-glucan 6-alpha-D-(1,4-alpha-D-glucano)-transferase, MGH = maltose glucohydrolase SGH = sucrose glucohydrolase, UDPGTRA = UDP-glucose:glycogen 4-alpha-D-glucosyltransferase) increased prominently if the sucrose input was increased (Fig. 7B). The simulation of elevated EAA levels resulted in increased minimal fluxes of enzymes involved in lipid metabolism (ACS3 = acyl coenzyme A synthetase, DGA = 1,2-diacylglycerol acyltransferase, GK = glycerol kinase, GPPA = alpha-glycerophosphate acyltransferase), as well as situations where both the lower and upper flux limits were elevated (FASN = fatty-acid synthase), which suggest elevated lipid storage levels as a result (Fig. 7C).

When we performed corresponding metabolite measurements with animals raised under the different growth conditions, TAG levels indeed showed a modest, but significant, increase following growth on the 2x EAA diet (Fig. 8A). Free glycerol levels showed a larger amount of variation (Fig. 8B). Yet, the trends clearly differed in

6
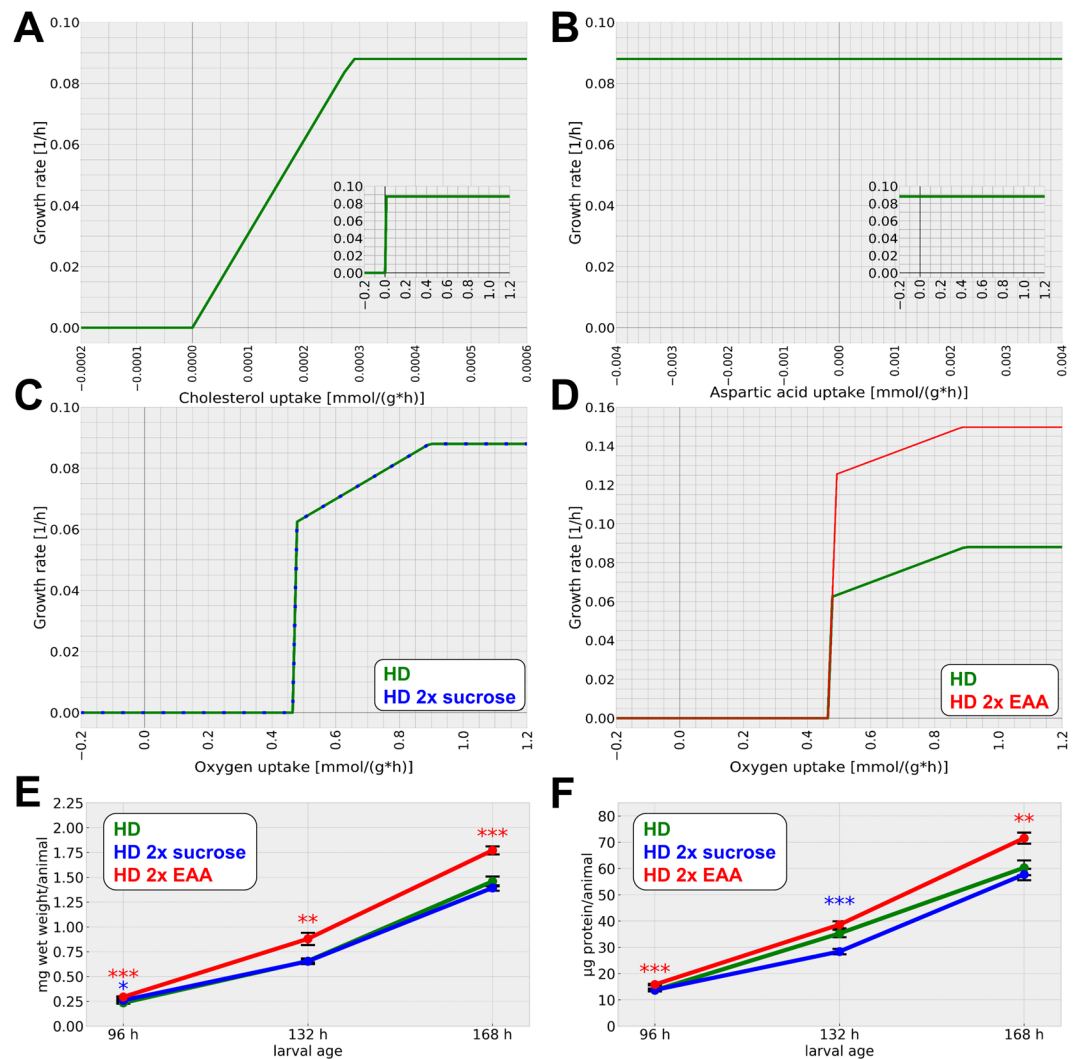
Figure 6. Model verification, predicted growth and comparison to real life. (**A–D**) Modeled growth rate in response to different input parameter variations. Negative uptake rates correspond to an excretion. (**A**) A certain level of cholesterol is needed for optimal growth while less result in a suboptimal growth phase. The zoom-in represents a larger cholesterol uptake flux range. (**B**) Levels of the non-essential aspartic amino acid do not affect biomass production. The zoom-in represents a larger aspartic acid uptake flux range. (**C**) Oxygen levels need to surpass a threshold to allow biomass production. In the following, the biomass production increases until it reaches a plateau. Increased sucrose levels (blue color) do not alter biomass production as compared to the standard HD (green color). (**D**) A doubling of the amount of essential amino acids (EAA) increases the biomass production (red color) as compared to the standard HD (green color). (**E,F**) Experimental testing of our model-based predictions. Animals were either reared on HD (green), HD containing the double amount of sucrose (blue) or the double amount of EAA (red). The wet weight (**E**) and protein (**F**) content of the larvae was measured 96, 132, and 168 hours after egg laying. While the altered sugar content did not affect the growth rate, the addition of more EAA resulted in a higher growth rate (**E**). The protein measurements show similar results (**F**). Measurements in (**E,F**) represent the mean values of three biologically independent experiments. Each experiment consisted of quadruplicate samples. Whiskers represent the standard error of the mean (SEM). Please note that the wet weight data for the HD is identical with the one shown in Fig. 4A. Statistical significance was tested by an unpaired two-sample T-Test for each time point. Significance levels are: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

response to the varying diet compositions. Increased dietary sugar levels resulted in lower levels of free glycerol, whereas increased amounts of EAA resulted in an on average increased free glycerol levels indicative of elevated lipolytic and/or lipogenic activity. Intuitively, we expected an altered sugar content of the animals raised on the 2x sucrose diet. Yet, we did not detect any prominent differences (Fig. 8C), which seems to be in line with the modeling results, which indicate a larger flux of the glycolysis reactions with a simultaneous activation of the complete TCA cycle or the reactions involved in oxidative phosphorylation (Fig. 7A). This increase of flux values suggest that the metabolism of *Drosophila* activates a metabolic program for an overflow metabolism perhaps associated with a larger burning or excretion of sugars. The 2x EAA diet, however, resulted in a significantly higher glucose
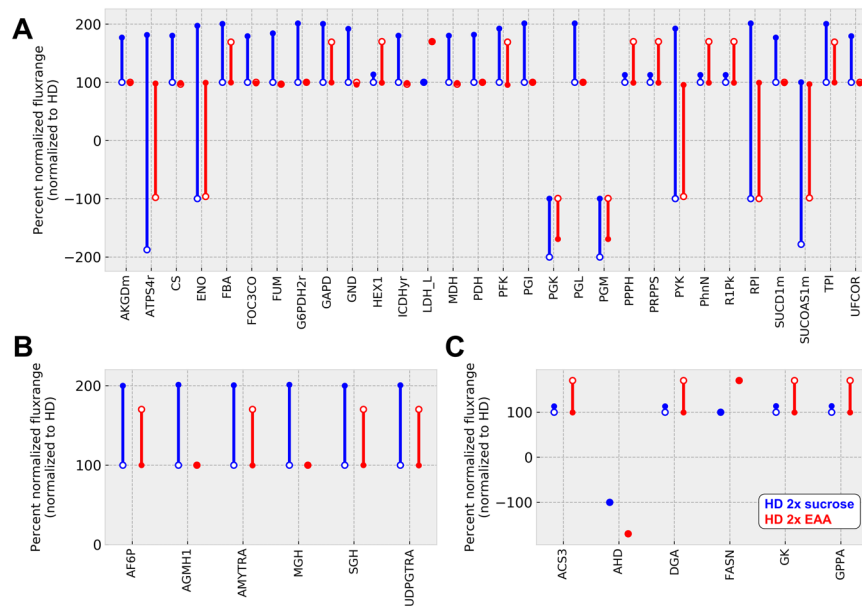
Figure 7. Modeling *Drosophila* larval resource allocation. The figure shows the results of the percent normalized flux variability analysis results for the simulations based on the HD with 2x sucrose (blue) and HD with 2x EAA (red), respectively. In brief, we percent normalized the values of the minimal (open circles) and maximal (filled circles) fluxes on the basis of the results for the standard HD. The sign of the flux percentage indicates the reaction direction, where a positive sign indicates the forward reaction, a negative sign indicates the backwards reaction and values spanning both signs represent reversible reactions. (**A**) Barplot of the central carbon metabolism flux change of the model solutions for the comparisons HD and HD with 2x sucrose (blue) or HD and HD with 2x EAA (red), respectively. (**B**) Flux change for the sucrose metabolism (description as in (**A**)). (**C**) Flux change for the lipid metabolism (description as in (**A**)). Each reaction description can be found in Table S1. Flux variability analysis results are given in Table S3 and the normalized flux data are presented in Table S4.

content of the animals (Fig. 8C), which appears to be in line as the higher flux rate of diphosphate phosphohydrolase or fatty-acid synthase suggest a higher beta-oxidation besides an increased lipid storage predicted by the model. Glycogen levels were mostly unaffected by the altered nutritional conditions (Fig. 8D). Lactate showed under basal HD growth conditions a plateau 132 hours after egg laying (Figs 4F, 8E). If additional sucrose was present, lactate levels even decreased 132 hours after egg laying (Fig. 8E). The 2x EAA diet yet resulted in statistically increased lactate levels (Fig. 8E), which again is in line with our modeling results, which suggested an increased lactate dehydrogenase flux (Fig. 7A). Thus, we also could largely align our modeled flux values with the corresponding experimental data.

Discussion

Flux balance analyses (FBA) with metabolic networks of heterotrophic multicellular organisms usually appears complicated due to the difficulty of identifying a clear-cut objective function. Further, the various cell types of complex organisms, with their signature gene expression profiles and distinct metabolic and functional tasks, complicate modeling approaches even more. To tackle these difficulties, methods and strategies were designed[38], as for example, to model the metabolic flow in the context of gene regulation or other advanced constraints such as dynamic changes. Here, we decided to use a top-down modeling approach focusing on a generalized and averaged simple model of the growing larva, which we identified as a suitable system for FBA given its clear-cut objective function. We used this simplified model given that many details of the overall organismic physiology are still unclear. Therefore, we rationalized that adding uncertain information might rather act detrimental as compared to a simplified model, which could better catch the more general schemes. A benefit from starting our model mostly from scratch was that we could pay special attention on avoiding dead-end and blocked-reactions (Fig. 3), as well as to assure the biological feasibility of the subsequent FBA modeling results. Our resulting model is to the best of our knowledge the currently most advanced metabolic network of *Drosophila* metabolism and we provide it together with the code to perform FBA analyses in different ready-to-use formats to the community (see methods section). Our initial rationale that a simple model might be already useful appears valid, given that it is already able to recapitulate basic biology (Fig. 6) and successfully predicted the impact of dietary alterations on the larval growth and metabolism (Figs 6–8). In the future, the refinement of the model by e.g. incorporating gene regulation or a dynamic modeling[38], potentially further enhances the predictive power of the model. While these enhancements will be likely less important for the modeling of a general behavior of the system – and thus general resource allocation questions as there are still many uncertainties resulting in a high number of degrees of freedom –, they will facilitate the studying of new questions. For example, it will be
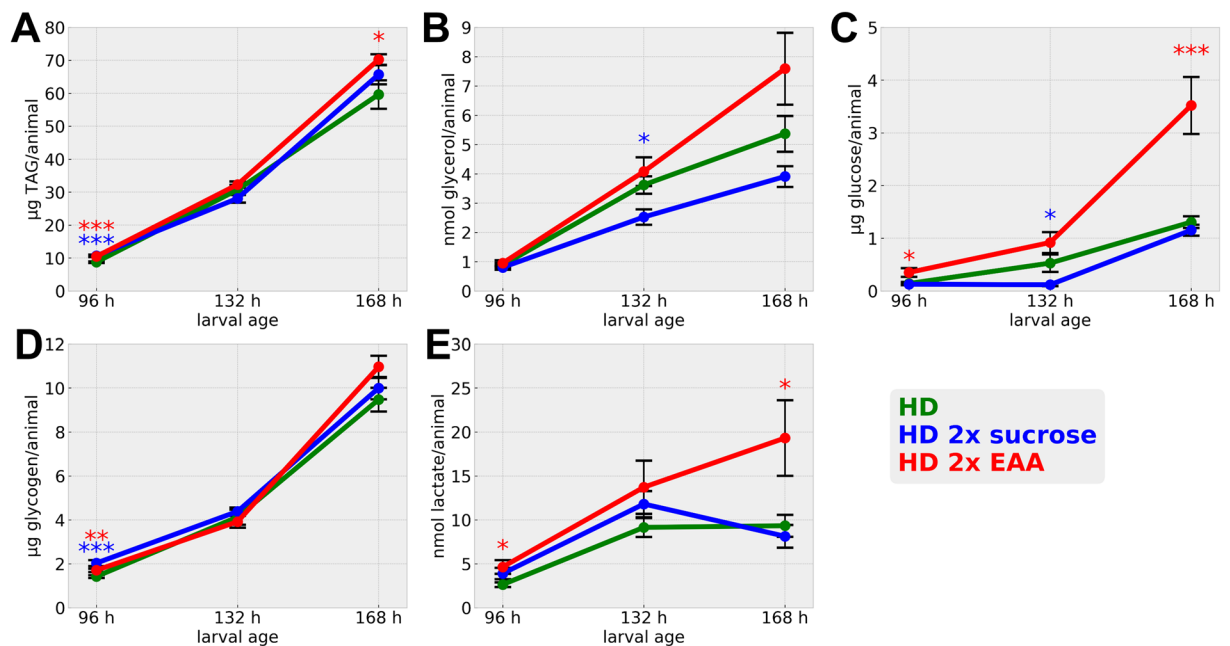
Figure 8. Metabolic profiling of the larval developmental of *Drosophila melanogaster* 96, 132, and 168 hours after egg laying. (**A–E**) Absolute quantification of triglyceride (TAG), glycerol, glucose, glycogen, and lactate levels. Measurements in (**A–E**) represent the mean values of three biologically independent experiments. Each experiment consisted of quadruplicate samples. Whiskers represent the standard error of the mean (SEM). Statistical significance was tested by an unpaired two-sample T-Test for each time point. Significance levels are: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

interesting to investigate the interplay between different organs, as well as to study inter-organismal metabolic connections such as between the host and its gut microbiome constituents or host-symbiont/host-parasite mutualism. *Drosophila* is an exquisite model for such kind of studies given its well characterized and relatively simple gut microbiome[39]. Further, the mutualism between insects and the endosymbiont *Wolbachia*, which can also act as pathogen, is well described[40,41]. In *Aedes aegypti*, an impact of an infection with the pathogenic *Wolbachia* strain *w*MelPop, for example, was recently shown to affect the TAG and cholesterol metabolism of the host[42]. Further, the time-resolved analysis of metabolite amounts as well as the consequences of gene dosage and protein activity alterations, will be interesting avenues to follow.

Our lack of knowledge concerning the energy costs associated with growth and non-growth associated processes (often referred to as GAM – growth associated maintenance, and NGAM – non-growth associated maintenance costs) might appear as a weak spot of our approach. So far, we determined the NGAM values in an iterative process based on the oxygen consumption rate of *Drosophila* S2R+ cells[43] (methods section and Fig. S5) and used the GAM values from a yeast model[32], as an experimental estimation of the values for *Drosophila* is difficult. Our approach appears legitimate, at least at the current point of time, as simulations testing a wide array of different GAM and NGAM values with our model demonstrated only a limited impact on the biomass production (Fig. S6 and Interactive Supplementary Fig. 2). The impact of GAM and NGAM uncertainties on the resource allocation problem might be bigger. The size of the linear programming solution space decreases with increasing GAM and/or NGAM values, as fluxes have to compensate for the increased energy requirements as for example the rate of oxygen consumption. Thus, the future estimation of the real GAM and NGAM values for larvae under different physiological conditions is a challenging, yet valuable goal.

We parameterized our model using absolute enzyme-based biochemical quantifications as well as via GC-MS based metabolomics measurements (Figs 4–6 and Table S2). Overall, our measurements explained a large portion of the dry weight of the animals (Fig. 5B). The prominent drop of explained dry weight at 168 hours after egg laying (60% explained dry mass versus 81 or 95% (96 or 132 hours after egg laying, respectively)) is intriguing. Given that despite of lactate all metabolites measured by us increased over time (Figs 4 and 8), this result suggests that the production of another metabolite not measured by us is increasing dramatically during this growth phase. Candidates are, for example, nucleic acids or components of the cuticle. In support of this notion, we noted an elevated slope of the larval weight gain from the second to the last time point (Fig. 4A), whereas size increase on the HD rather stalled during this phase (Fig. S2B–D).

The use of the chemically defined HD significantly facilitated the connection of the biological data to the modeling both under basal (Figs 4–8) as well as dietary altered (Figs 4, 6 and 8) conditions. The presence of the chemically defined food is a big advantage as compared to complex and ill-defined food compositions, as uncertainties in terms of the diet obfuscates defining the real inputs entering the system in the experiment as well as the model. An obvious point for optimization is our lack of knowledge concerning the exact amounts of food consumed and metabolite resorption rates. The mouth hook contraction values used by us, for example, could

in theory vary in response to different diets. Several reports, however, suggested only a limited impact of the diet on the mouth hook contraction frequencies[33,34]. Nevertheless, in the future, methods that are more sophisticated should be used to eliminate this shortcoming using e.g. radiolabeled tracer experiments, which will also allow the direct estimation of metabolite flux rates. These in turn will allow a much better comparison to the predicted flux rates as our endpoint measurements. Another point to consider is that the HD represents a minimal medium, which results in a slowed down growth. Here, we started to investigate which nutrients might act as growth limiting factors and our strategy appears to be successful in predicting nutrients allowing a faster growth. For the investigation of resource allocation and optimality principles, the slowed-down growth appears less problematic, given that optimality of resource allocation is situation dependent, and thus an inadequate diet can also be utilized in an optimal manner.

Investigations concerning the minimal nutritional requirements of an organism, as well as concerning the impact of nutritional alterations on the physiology of an organism, is an active field of research[44–46]. Our early results with the predictions of the impact of altering sucrose or EAA levels in the food are particularly promising. They suggest that simulations with FlySilico should facilitate the identification of suitable parameter ranges for experiments targeting e.g. the nutritional requirements of larvae and flies or the impact of diet alterations on the metabolic and/or growth phenotype. The future FlySilico-based investigations concerning the impact of varying amounts of essential and branched chain amino acids on growth processes, life history traits such as fecundity, ageing related diseases and cancer will be exciting as these aspects gained a lot of attention recently[47–50].

Drosophila larval growth is marked by an impressive increase in size and mass. This expansion is necessary for a successful completion of metamorphosis, which involves a drastic remodeling of body structures and a food intake cessation. Therefore, the larval development is subject to hard biological constraints. Still, the organism can react to fluctuations in the quality or quantity of food by adjusting the rate of development and resource allocation, e.g. by channeling less nutrients into storage forms[15]. Our experiments, where we raised animals on either standard HD or HD supplemented with additional sucrose or EAA (Figs 4, 6 and 8), demonstrate this plasticity. Further, our model predictions correctly identified the growth limiting nutritional parameters and revealed flux differences, which relate to the observed metabolic changes based on the diet alterations. The precision of the model predictions will increase with further improvements as outlined above. Intriguingly, the metabolic adaptation to altered nutritional conditions are not limited to insects, but have also been described e.g. for mammals[20]. Thus, future investigations targeting other species are possible to test for a possible generalization of our findings.

An important aspect is that solving the flux balance model is by definition following optimality principles. Given that our relatively simple model already was able to result in correct predictions suggests that *Drosophila* resource allocation is operating in a quasi-optimal state. Future studies with additional parameter variations (e.g. cost functions for protein and DNA synthesis and distinguishing different type of organs or tissues[38]) and incorporation of additional fly genotypes and perhaps single gene mutations will help to further elucidate this intriguing possibility.

Materials and Methods

Drosophila fly stocks and rearing.
For all our experiments, we used the wild type Oregon-R fly strain reared under standard culture conditions (25 °C, 12 h light-dark rhythm and 60–70% humidity).

Chemically defined fly medium.
The chemically defined (holidic) medium (*Holidic diet*; HD) was introduced in[26]. Animals developing on HD are viable, fertile, and have no aberrant phenotypes, although the development is slowed-down and the HD is thus classified a minimal medium. Food was prepared according to the instructions of the Piper *et al.* publication with use of the Yeast-like amino acid composition ("Yaa"). For our perturbation experiments we either added the double amount of sucrose (2x sucrose) or the double amount of essential amino acids (2x EAA).

Larvae collection procedure.
In order to minimize a possible impact of environmental effects, we kept the parental density constant with 15 male flies and 30 female flies per vial. These adult flies were kept on a standard complex cornmeal diet (per 100 mL: 0.5 g agar (Becton Dickinson, 214010), 7.1 g polenta (Verival, Pronurel Bio, 265250), 0.95 g soy flour (Bauck Hof, Amazon.de, B004RG3C0I), 1.68 g yeast (Bruggeman, lieferello. de, 14874413), 4 g treacle (Original Grafschafer Goldsaf, lieferello.de, 10231869), 4.5 g malt extract (Demeter, Amazon.de, B00GU029LW), 0.45 mL propionic acid (Acros Organics, 220130010, CAS 79094) and 1.5 mL nipagin (Sigma-Aldrich, H3647-100G) (1:10 stock solution in 70% Ethanol; Riedel-de Haën, 16202S-1L, CAS 64-17-5) before we transferred them to the chemically defined medium to allow oviposition. After six hours we discarded the parental flies to have a defined time period for the egg-laying. As the maximum time point for collecting larvae we used 168 h after egg laying as afterwards the larvae start to pupariate on the HD. Based on this terminal point, we added two equally spaced time points earlier in development (132 h and 96 h after egg laying, respectively) as growth is quasi linear.

The possibility that larval growth and metabolism is showing a sexual dimorphism appeared intriguing. Pilot experiments using animals reared on the holidic or a standard diet, however, showed that at the latest timepoint used by us female and male larvae do not significantly differ in terms of the size, the triglyceride, glucose or glycogen levels (data not shown). Thus, we did not consider the sex of the animals during our experimental timeframe a prominent factor and therefore collected unsexed larvae 96 h, 132 h and 168 h after egg laying and washed them in PBS with 0.1% Tween-20 (PBT) for the weight measurements and metabolic assays or in HPLC-graded water for the GC-MS analytics. We used quadruplicates for every condition and collected 25 larvae (96 h) or eight larvae (132 h, 168 h) for the GC-MS and the metabolic assays. For the dry and wet weight measurements, we collected 100 animals from the 96 h time point and 40 animals from the 132 h and 168 h time points, respectively.

Wet and dry weight measurements.   For the determination of the wet weight, we transferred the washed larvae into pre-weighed 1.5 mL tubes and weighed them on an analytical scale. The animals were then snap-frozen in liquid nitrogen and dried in an oven at 60 °C with the tube lids open. After 24 h we measured again the weight (= dry weight) and calculated the water content by subtraction of the dry weight from the wet weight.

Larval size measurements.   For the larval size measurements, we collected at the indicated time points the animals in ice cold PBS to minimize their movements and to ensure their elongation. Subsequently, we recorded images with a Zeiss SteREO Discovery.V8 dissection microscope, which were analyzed with the Zeiss Zen Software (Zen 2.3 lite – blue edition). For each larva, we measured the area, the length and its width. In total, we performed three biologically independent experiments and measured 20 to 30 animals per repetition.

Biochemical measurements.   All targeted biochemical measurements were essentially carried out as described in[31]. We collected the larvae from three different time points and snap-froze the animals in liquid nitrogen before storage at −80 °C. For the homogenization, we used 1 mL 0.05% Tween 20 in water in 2 mL screw-cap tubes and a Fast Prep FP120 machine (Bio101 Savant). After homogenization and heat-inactivation for 5 minutes at 70 °C, the supernatant was transferred to 1.5 mL tubes as a reservoir for the metabolic assays, which were performed in 96-well plates. We normalized each measurement to the amount of animals per sample.

Protein.   The free protein content was measured using the Pierce BCA assay kit (Life Technologies) according to the manufacturer's instructions. We used bovine serum albumin (BSA) as a standard to determine the protein content of the samples. The 0.05% Tween-20 used by us in the homogenization procedure most likely was not sufficient to solubilize also integral membrane proteins. Thus, the protein amounts per animal might represent an underestimation of the real total protein content.

Triglycerides (TAG).   For the determination of the triglyceride levels in the samples, we used the Triglycerides Reagent (Thermo Scientific). We transferred 50 μL of the samples and a serial dilution (1:2 in 0.05% Tween 20 in water) of the glycerol standard (Sigma Aldrich) to a 96-well plate and added 200 μL of the Triglycerides reagent. The samples and the standard were incubated 45 minutes at 37 °C and the absorbance was read at 510 nm.

Glycerol.   The glycerol content of the samples was determined using the Glycerol Assay Kit (Sigma-Aldrich). We followed the manufacturer's instruction for fluorometric measurements.

Glucose and Glycogen.   For the determination of glucose and glycogen, we used the GO Assay Reagent (Sigma-Aldrich) and a modified form of a protocol described in[51]. For both measurements, we transferred 30 μL of the undiluted samples and the standards to a 96-well plate. We added 100 μL GO reagent to measure free glucose and 100 μL GO reagent with amyloglucosidase (1 μL per 1 mL GO reagent) to measure the total glucose content (free glucose plus glucose liberated from the glycogen). After 60 minutes incubation by 37 °C we stopped the reactions by adding 100 μL 12 N H_2SO_4 and measured the absorbance at 540 nm. We calculated the glycogen content by subtraction of the free glucose from the total glucose content.

Lactate.   For the quantification of lactate, we used the Lactate Assay Kit (Biovision). We transferred 50 μL of the pre-diluted samples (1:50) to a 96 well plate and followed then the manufacturer's instruction for fluorometric measurements.

Cholesterol.   Measurements were performed as described in[52].

GC-MS measurements.   Metabolites were extracted using 105 μL chloroform and 245 μL methanol. After incubation for 1 h at −20 °C we added 560 μL HPLC grade water twice. The samples were centrifuged for two minutes at high speed in a table top centrifuge at 4 °C and the aqueous phases were collected for the GC-MS measurements (in total about 1.3 mL).

For the metabolite analysis a gas chromatography – mass spectrometry (GC-MS) system (7200 GC-QTOF from Agilent) was used as described in[53]. The data were analyzed with the Mass Hunter Software (Agilent). For absolute quantifications, we used five different dilutions of the standard mix (resulting in effective metabolite concentrations: 1 μM, 5 μM, 10 μM, 15 μM and 20 μM; Fig. S3) and calculated for each metabolite a standard curve which we used to determine the amount of the respective metabolite in our samples.

Network reconstruction.   For the *in silico* reconstruction of *Drosophila melanogaster* growth and metabolism we focused on core metabolic pathways required to metabolize the HD ingredients, and used the cameo package for the Python programming language[54]. The Yeast iMM904 model from the BiGG data base[32] and a previously published *Drosophila* model for hypoxia investigations[30] served as starting points for our network reconstruction. First, we incorporated major metabolic pathways of the carbohydrate metabolism (including glycolysis, gluconeogenesis, tricarboxylic acid (TCA) cycle and pyruvate metabolism), the lipid metabolism (with a focus on the glycerolipid metabolism), the energy liberating metabolic reactions (e.g. oxidative phosphorylation) and anaplerotic reactions. Subsequently, we successively integrated metabolic reactions necessary to metabolize the HD ingredients, such as amino acid metabolic pathways (including e.g. glycine, threonine, cysteine, and phenylalanine metabolism), pathways of vitamin synthesis (e.g. folate and riboflavin) and various pathways needed for the transport and synthesis of cofactors. For the initial version of FlySilico we only focused on three different compartments (extracellular, cellular and mitochondrial) and adapted the transport reactions accordingly. We manually curated all reactions by cross-validation with multiple resources (BioCyc - https://biocyc.org/ [55], BRENDA - https://www.brenda-enzymes.org/ [56], ChEBI - http://www.ebi.ac.uk/chebi/init.do [57],

KEGG - http://www.genome.jp/kegg/ [58], PubChem - https://pubchem.ncbi.nlm.nih.gov/ [59], BiGG - http://bigg.ucsd.edu/ [60], and FlyBase - http://flybase.org/ [61]) and paid special attention on the biochemical pathways and the genetics of *Drosophila melanogaster*. We attached to each reaction a confidence score based on evidence of sequence, physiological, genetic or biochemical data (as previously suggested by e.g. [62]). Reactions required for the modeling, but without any evidence of correctness, received the lowest confidence scores. FlySilico version 1.0 covers 293 metabolites and 363 reactions. Supplemental Table S1 summarizes all pathways, reactions and metabolites present in the model.

Constraint-based modeling. After the reconstruction process, the coefficients of the mass-balanced reactions form a mathematical representation as stoichiometric matrix S. The constraint-based modeling approach follows equation:

$$\frac{dx}{dt} = \mathbf{S} \cdot \boldsymbol{v} = 0 \tag{1}$$

with

$$\boldsymbol{\alpha_i} \leq 0 \leq \boldsymbol{\beta_i} \tag{2}$$

where x is a vector with all metabolites, S is the stoichiometric matrix and v is a vector with all fluxes under steady state conditions. The lower (α_i) and upper (β_i) bounds to each flux v_i impose additional constraints to the system. The null space of S includes any v that satisfies the solution under this steady state assumption with the given constraints. The model is solved by optimizing the system for a given objective function, i.e. the primary goal of an organism (such as biomass production for fast growing unicellular organisms as for example *E. coli*), using linear programming. A detailed explanation of constraint-based modeling, flux balance analysis and linear programming is provided e.g. in:[63] or[64]. For our model solutions we compared solutions allowing loops as well as loopless[65,66] variants (Fig. S7). The loopless solution reflects the biology better, and thus we used this method throughout the study unless otherwise noted.

Flux variability analysis. Flux variability analysis (FVA[67]) is a computational method to identify the maximum and minimum fluxes of reactions from a given network while it preserves a certain network state (e.g. maximum biomass production rate). FVA solves two optimization problems for each reaction v_i after solving a given objective function.

$$
\begin{aligned}
& max_v/min_v \ v_i \\
& \text{subject to} \quad \mathbf{S} \cdot \boldsymbol{v} = 0 \\
& \qquad\qquad \boldsymbol{c}^T \ast \boldsymbol{v} \geq \gamma \ast Z_0 \\
& \qquad\qquad \boldsymbol{\alpha_i} \leq 0 \leq \boldsymbol{\beta_i}
\end{aligned}
\tag{3}
$$

where $Z_0 = \boldsymbol{c}^T \boldsymbol{v_o}$ is an optimal solution for the objective function, γ is a parameter which controls the optimality of the solution (suboptimal: $0 \leq \gamma < 1$; optimal: $\gamma = 1$), $\mathbf{c}$ is the vector which represents the linear objective function.

Calculation of biomass and uptake rates. In order to identify an appropriate biomass function, we estimated the body composition of differently aged larvae by targeted absolute biochemical quantifications as well as GC-MS based metabolomics measurements. We reasoned that the main larval constituents are water, proteins, carbohydrates and storage lipids given that the latter two are the main storage forms fueling metamorphosis and that the larvae are filled up with the adipose-tissue like storage organ – called the fat body – which is the main storage site for storage lipids and glycogen. The water content could be easily measured by gravimetrics (see above) and on top of free protein, glucose, triglyceride and glycogen amounts, we also determined the levels of lactate and free glycerol by targeted biochemical assays (see above). Our GC-MS measurements further covered metabolites from the central carbon metabolism as well as almost all free amino acids.

Biomass functions usually cover each amino acid separately. Yet, our free protein measurements did not provide such fine-grained information. The GC-MS based metabolomics measurements resulted in the identification of free amino acid amounts; yet three amino acids (arginine, glutamine and histidine) were missing in our measurements. In order to approximate the levels of the different amino acids, we followed a bioinformatics strategy with the reasoning that the amino acid fractions would relate to the respective amino acid frequency across the *Drosophila* proteome. Thus, we first calculated the frequency of each of the twenty classical amino acids in the complete proteome of *Drosophila melanogaster* (http://www.uniprot.org/uniprot/?query=proteome:UP000000803). Figure S8 provides the calculated amino acid frequencies. We based the coefficients on the differences between the first and last time points investigated. Thus, for each measurement we calculated:

$$m_{168h,Metabolite} - m_{96h,Metabolite} = m_{\Delta,Metabolite} \tag{4}$$

where $m_{\Delta,Metabolite_i}$ represents the weight difference of a metabolite i between 168 h and 96 h in *gram*, $m_{168h,Metabolite_i}$ is the weight of the metabolite i at 168 h in *gram*, and $m_{96h,Metabolite_i}$ is the weight of the metabolite i at 96 h in *gram*. For the amino acids, we could now calculate the individual amino acid weights with the help of the calculated amino acid frequencies:

$$m_{\Delta,Metabolite_i} \triangleq m_{AA_i} = m_{\Delta,Protein} * f_{AA_i} \tag{5}$$

where m_{AA_i} shall be equivalent to $m_{\Delta,Metabolite_i}$ and represents the weight of amino acid i in $gram$, $m_{\Delta,Protein}$ is the difference of weight of the protein in $gram$ which was calculated based on equation Eq. 4, and f_{AAi} which represents the frequency of amino acid i from Fig. S8. Metabolite weights from equations Eqs 4 and 5 enable the calculation of an assay-based coefficient for each metabolite for the biomass objective function:

$$x_{Assay,Metabolite_i} = \frac{\frac{m_{\Delta,Metabolite_i}}{M_{Metabolite_i}}}{m_{\Delta,dry}} = \frac{n_{Metabolite_i}}{m_{\Delta,dry}} \tag{6}$$

where $x_{Assay,Metabolite_i}$ represents the coefficient of the metabolite based on assay data in $\frac{mmol}{g}$, $M_{Metabolite_i}$ is the molar mass of metabolite i in $\frac{g}{mmol}$, $m_{\Delta,dry}$ is the larval dry weight difference from 96 h till 168 h in $gram$, and $n_{Metabolite_i}$ is the amount of metabolite i in $mmol$.

The GC-MS analysis quantified free metabolite amounts. In order to calculate a GC-MS coefficient, we calculated the difference of metabolite amounts between 168 h and 96 h by the following equation:

$$n_{\Delta,Metabolite_i} = n_{168h,Metabolite_i} - n_{96h,Metabolite_i} \tag{7}$$

where $n_{\Delta,Metabolite_i}$ represents the difference of amount between 168 h and 96 h from metabolite i in $mmol$, $n_{168h,Metabolite_i}$ is the amount of the metabolite i at 168 h in $mmol$, and $n_{96h,Metabolite_i}$ is the amount of the metabolite i at 96 h in $mmol$. We used the resulting metabolite amounts from equation Eq. 7 to calculate a GC-MS coefficient with the equation:

$$x_{GCMS,Metabolite_i} = \frac{n_{Metabolite_i}}{m_{\Delta,dry}} \tag{8}$$

where $x_{GCMS,Metabolite_i}$ as the GC-MS coefficient of the metabolite i in $\frac{mmol}{g}$, $m_{\Delta,dry}$ is the difference of dry weight between 168 h and 96 h in $gram$, and $n_{Metabolite_i}$ is the amount of metabolite i in $mmol$.

Through equation Eqs 6 and 8 the biomass function coefficient can be calculated for all metabolites with the following equation:

$$x_{Metabolite_i} = x_{Assay,Metabolite_i} + x_{GCMS,Metabolite_i} \tag{9}$$

where $x_{Metabolite_i}$ is the biomass function coefficient of metabolite i, $x_{Assay,Metabolite_i}$ as the assay-based coefficient of metabolite i, and $x_{GCMS,Metabolite_i}$ as the GC-MS-based coefficient of metabolite i. All coefficients of Eq. 9 are in $\frac{mmol}{g}$. *Drosophila melanogaster* is cholesterol auxotroph[27,68]. To simulate the cholesterol auxotrophy of *Drosophila*, we included cholesterol in the biomass function based on measurements of cholesterol levels from larvae reared on HD (Fig. S9). The biomass coefficient of cholesterol was calculated according to the equation:

$$x_{Cholesterol} = \frac{\frac{c_{Cholesterol,Protein} * m_{\Delta,Protein}}{M_{Cholesterol}}}{m_{\Delta,dry}} \tag{10}$$

where $x_{Cholesterol}$ is the biomass coefficient for cholesterol in $\frac{mmol}{g}$, $c_{Cholesterol,Protein}$ is the cholesterol mass per protein mass in $\frac{g}{g\ Protein}$, $m_{\Delta,Protein}$ is the difference of weight of protein between 168 h and 96 h in $gram$, $M_{Cholesterol}$ is the molar mass of cholesterol in $\frac{g}{mmol}$, $m_{\Delta,dry}$ is the difference of dry weight between 168 h and 96 h in $gram$. The mean value of the cholesterol mass per protein mass of all 3 time points (96 h, 132 h and 168 h) is $c_{Cholesterol,Protein} \approx 5.46\ \frac{ng}{\mu g\ Protein}$.

Approximation of food intake. Given that the absolute quantification of the uptake of solid food by larvae is difficult, and that the measurement of the absorption rate and organismic distribution for each nutrient is close to impossible, we used a theoretical approximation of the food intake as a starting point for our modeling experiments. First, we calculated the maximum volume of the mouth cavity by approximating a cylindrical shape and taking length and diameter measurements of the differently aged larvae into account. We calculated the oral cavity volume according to equation:

$$V_{mouth} = \pi * r^2 * h \tag{11}$$

where V_{mouth} is the volume of the oral cavity in mm^3, π is the mathematical constant, r is the radius of the oral cavity in mm, and h is the height of the oral cavity in mm.

Because the larva grows over time, we estimated the diameter of the larva as the mean of the width of time points 96 h and 168 h and assumed that the diameter of the oral cavity is about half of the larva diameter (the radius of the oral cavity thus is $r = 0.118\ mm$). We took the height h of the oral cavity from a publication from Alpatov[35]. Here, the height of the oral cavity is the average of the mean length from larval stages II and III with a value of $h = 0.25\ mm$. Thus, the resulting oral cavity volume according to equation Eq. 11 is $v_{Mouth} = 0.011\ mm^2$.

We approximated the feeding rate using measures of the sclerite reactions per minute from[69] $\left(f_{Feed} = \frac{110}{min}\right)$, with the assumptions that each sclerite reaction completely fills the oral cavity and that all food ingredients are homogeneously distributed. The dietary intake can thus be calculated using the following equation:

$$m_{Dietary} = v_{Mouth} * f_{Feed} * \rho_{holidic} \tag{12}$$

where $m_{Dietary}$ is the dietary intake in $\frac{g}{h}$, v_{Mouth} is the oral cavity volume in mm^3, f_{Feed} is the sclerite reactions per h, and $\rho_{holidic}$ is the sum of the mass concentrations of the holdic diet ingredients in $\frac{g}{\mu L}$. Our calculated dietary intake is $m_{Dietary} = 0.064 \frac{g}{h}$.

Our calculation is of course an overestimation given that each sclerite reaction most likely does not fill the mouth volume completely and that the uptake from the gut is not 100% efficient. To account for this limitation, we introduced a correction factor χ based on our experimental data and simulations by an iterative process. In brief, we calculated all uptake rates with increasing values for the correction factor χ (from 0 to 0.20 in 0.001 steps) and used the different uptake rates to calculate the corresponding growth rate. We selected for χ the value where the calculated growth rate fitted the experimentally determined growth rate best (Fig. S4). We determined the experimental growth rate based on the dry weight measurements during the three time points. We calculated the growth rate between the first and last time point as:

$$\mu_{BIOMASS,Exp} = \frac{\frac{\overline{m}_{t2} - \overline{m}_{t1}}{\Delta t}}{\overline{m}_{t1}} \tag{13}$$

As a result we obtained the experimental growth rate $\mu_{BIOMASS,Exp} = \frac{0.0882}{h}$ and for $\chi = 0.122$ the predicted growth rate $\mu_{BIOMASS,Pred} = \frac{0.088}{h}$. Thus, the corrected dietary intake calculation is:

$$m_{Dietary,Metabolite_i} = m_{Dietary} * \chi * p_{Metabolite_i} \tag{14}$$

where $m_{Dietary,Metabolite_i}$ is the dietary intake of metabolite i in $\frac{g}{h}$, $m_{Dietary}$ is the dietary intake in $\frac{g}{h}$, χ is the correction factor, and $p_{Metabolite_i}$ is the proportion of each metabolite i in the holidic medium.

For the network modelling, all dietary internalizations have to be flux rates or uptake rates, which we calculated as follows:

$$v_{uptake,i} = \frac{m_{Dietary,Metabolite_i}}{M_{Metabolite_i} * m_{\Delta,dry}} \tag{15}$$

where $v_{uptake,i}$ is the uptake rate of metabolite i in $\frac{mmol}{g * h}$, $m_{Dietary,Metabolite_i}$ is the dietary intake of metabolite i in $\frac{g}{h}$, $M_{Metabolite_i}$ is the molar mass of metabolite i in $\frac{g}{mmol}$, and $m_{\Delta,dry}$ is the difference of dry weight between 168 h and 96 h in $gram$.

Summary statement.   FlySilico, a flux balance analysis suitable metabolic network of *Drosophila melanogaster* is presented, and its use for the investigation of larval growth and metabolism is demonstrated.

Data availability
The supplementary material consists of the metabolic network reconstruction (Table S1), the raw data of all measurements (Table S2), the result of the flux variability analysis for the different diet simulations (Table S3) and the corresponding flux changes in relation to the standard HD (Table S4). Additionally, we added interactive versions of the metabolic network map (Interactive Supplementary Fig. 1) and the GAM/NGAM plot (Interactive Supplementary Fig. 2), which corresponds Fig. S6. Further, we provide all raw data, the custom python scripts used for the calculation of the weights of the biomass function and the uptake rates, the model comparison and the main scripts for the model reconstruction and analysis in a zip file for use with the Anaconda Project function (https://anaconda-project.readthedocs.io/en/latest/). Once unzipped, the folder contains all information to create an Anaconda Python environment with the required packages in the appropriate version to run our code and all necessary information to rerun or modify our analyses. A readme file within the folder should guide users through the procedure to get the environment operational. We additionally provide all scripts via GitLab: https://gitlab.com/Beller-Lab/flysilico.

References
1. Davison, R., Boggs, C. L. & Baudisch, A. Resource allocation as a driver of senescence: life history tradeoffs produce age patterns of mortality. *J. Theor. Biol.* **360**, 251–262 (2014).
2. Dmitriew, C. M. The evolution of growth trajectories: what limits growth rate? *Biol Rev Camb Philos Soc* **86**, 97–116 (2011).
3. Dashty, M. A quick look at biochemistry: Carbohydrate metabolism. *Clin Biochem* **46**, 1339–1352 (2013).
4. Hood & Terjung. Amino acid metabolism during exercise and following endurance training. *Sports Medicine Auckl N Z* **9**, 23–35 (1990).
5. Votruba, S. & Jensen, M. Regional Fat Deposition as a Factor in FFA Metabolism. *Annu Rev Nutr* **27**, 149–163 (2007).
6. Feist, A. & Palsson, B. The biomass objective function. *Curr Opin Microbiol* **13**, 344–349 (2010).
7. Moed, D., Kruitwagen, C., Jong, D. & Scharloo Critical weight for the induction of pupariation in Drosophila melanogaster: genetic and environmental variation. *Journal of Evolutionary Biology* **12**, 852–858 (1999).

8. Schuetz, R., Zamboni, N., Zampieri, M., Heinemann, M. & Sauer, U. Multidimensional Optimality of Microbial Metabolism. *Science* **336**, 601–604 (2012).
9. Shoval *et al*. Evolutionary Trade-Offs, Pareto Optimality, and the Geometry of Phenotype Space. *Science* **336**, 1157–1160 (2012).
10. Szekely, P., Korem, Y., Moran, U., Mayo, A. & Alon, U. The Mass-Longevity Triangle: Pareto Optimality and the Geometry of Life-History Trait Space. *PLOS Computational Biology* **11**, e1004524 (2015).
11. Shoval, O. *et al*. Response to Comment on 'Evolutionary Trade-Offs, Pareto Optimality, and the Geometry of Phenotype Space'. *Science* **339**, 757–757 (2013).
12. Edelaar, P. Comment on 'Evolutionary Trade-Offs, Pareto Optimality, and the Geometry of Phenotype Space'. *Science* **339**, 757–757 (2013).
13. Merkey, A., Wong, C., Hoshizaki, D. & Gibbs, A. Energetics of metamorphosis in Drosophila melanogaster. *Journal of insect physiology* **57**, 1437–45 (2011).
14. Zinke, I., Kirchner, C., Chao, L. C., Tetzlaff, M. T. & Pankratz, M. J. Suppression of food intake and growth by amino acids in Drosophila: the role of pumpless, a fat body expressed gene with homology to vertebrate glycine cleavage system. *Development* **126**, 5275–84 (1999).
15. Beadle, T. & Bulletin, C. Food level in relation to rate of development and eye pigmentation in Drosophila melanogaster (1938).
16. Koyama, T., Rodrigues, M., Athanasiadis, A., Shingleton, A. & Mirth, C. Nutritional control of body size through FoxO-Ultraspiracle mediated ecdysone biosynthesis. *Elife* **3**, e03091 (2014).
17. Levin. Metabolic imprinting: critical impact of the perinatal environment on the regulation of energy homeostasis. *Philosophical Transactions of the Royal Society B: Biological Sciences* **361**, 11071121 (2006).
18. Martínez, D. *et al*. In utero undernutrition in male mice programs liver lipid metabolism in the second-generation offspring involving altered Lxra DNA methylation. *Cell metabolism* **19**, 941–51 (2014).
19. Vijendravarma, R., Narasimha, S. & Kawecki, T. Chronic malnutrition favours smaller critical size for metamorphosis initiation in Drosophila melanogaster. *Journal of Evolutionary Biology* **25**, 288–292 (2012).
20. Gluckman, P. *et al*. Metabolic plasticity during mammalian development is directionally dependent on early nutritional status. *Proc National Acad Sci* **104**, 12796–12800 (2007).
21. Kuzawa, C. W. Developmental origins of life history: growth, productivity, and reproduction. *Am. J. Hum. Biol.* **19**, 654–61 (2007).
22. Aktipis, C. A., Boddy, A. M., Gatenby, R. A., Brown, J. S. & Maley, C. C. Life history trade-offs in cancer evolution. *Nat. Rev. Cancer* **13**, 883–92 (2013).
23. Limones, M. *et al*. Metabolic alterations associated with maternal undernutrition during the first half of gestation lead to a diabetogenic state in the rat. *Eur J Nutr*, https://doi.org/10.1007/s00394-018-1805-z (2018).
24. Tarry-Adkins, J. L. *et al*. A suboptimal maternal diet combined with accelerated postnatal growth results in an altered aging profile in the thymus of male rats. *FASEB J.* fj201701350RR, https://doi.org/10.1096/fj.201701350RR (2018).
25. Bischoff, A. R. *et al*. Low birth weight is associated with increased fat intake in school-aged boys. *Br. J. Nutr.* **119**, 1295–1302 (2018).
26. Piper, M. D. *et al*. A holidic medium for Drosophila melanogaster. *Nat. Methods* **11**, 100–5 (2014).
27. Carvalho, M. *et al*. Survival strategies of a sterol auxotroph. *Development (Cambridge, England)* **137**, 3675–85 (2010).
28. Feala, J., Coquin, L., McCulloch, A. & Paternostro, G. Flexibility in energy metabolism supports hypoxia tolerance in Drosophila flight muscle: metabolomic and computational systems analysis. *Molecular Systems Biology* **3**, 99 (2007).
29. Coquin, L., Feala, J., McCulloch, A. & Paternostro, G. Metabolomic and flux-balance analysis of age-related decline of hypoxia tolerance in Drosophila muscle tissue. *Molecular Systems Biology* **4**, 233 (2008).
30. Feala, J. *et al*. Metabolism as means for hypoxia adaptation: metabolic profiling and flux balance analysis. *BMC Systems Biology* **3**, 1–15 (2009).
31. Jehrke, L., Stewart, F., Droste, A. & Beller, M. The impact of genome variation and diet on the metabolic phenotype and microbiome composition of Drosophila melanogaster. *Scientific Reports* **8**, 6215 (2018).
32. Mo, M. L., Palsson, B. O. & Herrgård, M. J. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *BMC Syst Biol* **3**, 37 (2009).
33. Ryuda, M. *et al*. Identification of a Novel Gene, Anorexia, Regulating Feeding Activity via Insulin Signaling in Drosophila melanogaster. *J Biol Chem* **286**, 38417–38426 (2011).
34. Musselman, L. *et al*. A high-sugar diet produces obesity and insulin resistance in wild-type Drosophila. *Dis Model Mech* **4**, 842–849 (2011).
35. Alpatov, W. Growth and variation of the larvae of Drosophila melanogaster. *J Exp Zool* **52**, 407–437 (1929).
36. Waldbauer, G. P. The Consumption and Utilization of Food by Insects. *Adv Insect Physiol* **5**, 229–288 (Academic Press, 1968).
37. Carman, G. M. & Han, G.-S. S. Roles of phosphatidate phosphatase enzymes in lipid metabolism. *Trends Biochem. Sci.* **31**, 694–9 (2006).
38. Martins Conde, P., do, R., Sauter, T. & Pfau, T. Constraint Based Modeling Going Multicellular. *Front Mol Biosci* **3**, 3 (2016).
39. Douglas, A. E. Simple animal models for microbiome research. *Nat. Rev. Microbiol*, https://doi.org/10.1038/s41579-019-0242-1 (2019).
40. Landmann, F. The Wolbachia Endosymbionts. *Microbiol Spectr* **7** (2019).
41. Pietri, J. E., DeBruhl, H. & Sullivan, W. The rich somatic life of Wolbachia. *Microbiologyopen* **5**, 923–936 (2016).
42. Geoghegan, V. *et al*. Perturbed cholesterol and vesicular trafficking associated with dengue blocking in Wolbachia-infected Aedes aegypti cells. *Nat Commun* **8**, 526 (2017).
43. Da-Ré, C. *et al*. UCP4C mediates uncoupled respiration in larvae of Drosophila melanogaster. *EMBO Rep.* **15**, 586–91 (2014).
44. Besson, A., Lagisz, M., Senior, A., Hector, K. & Nakagawa, S. Effect of maternal diet on offspring coping styles in rodents: a systematic review and meta-analysis. *Biol Rev* **91**, 1065–1080 (2016).
45. Caraballo, S., Comhair, T., Dejong, C., Lamers, W. & Köhler, E. A high-protein diet is anti-steatotic and has no pro-inflammatory side effects in dyslipidaemic APOE2 knock-in mice. *Brit J Nutr* **112**, 1251–1265 (2014).
46. Reis, T. Effects of Synthetic Diets Enriched in Specific Nutrients on Drosophila Development, Body Fat, and Lifespan. *PLoS ONE* **11**, e0146758 (2016).
47. Piper, M. *et al*. Matching Dietary Amino Acid Balance to the In Silico-Translated Exome Optimizes Growth and Reproduction without Cost to Lifespan. *Cell metabolism* **25**, 1206 (2017).
48. Shao, D. *et al*. Glucose promotes cell growth by suppressing branched-chain amino acid degradation. *Nature communications* **9**, 2935 (2018).
49. Chen, J. *et al*. KLHL22 activates amino-acid-dependent mTORC1 signalling to promote tumorigenesis and ageing. *Nature* **557**, 585–589 (2018).
50. Knott, S. *et al*. Asparagine bioavailability governs metastasis in a model of breast cancer. *Nature* **554**, 378 (2018).
51. Tennessen, J., Barry, W., Cox, J. & Thummel, C. Methods for studying metabolism in Drosophila. *Methods* **68**, 105–115 (2014).
52. Werthebach, M. *et al*. Control of Drosophila Growth and Survival by the Lipid Droplet-Associated Protein CG9186/Sturkopf. *Cell reports* **26**, 3726–3740.e7 (2019).
53. Fiehn, O. *et al*. Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* **18**, 1157–61 (2000).
54. Cardoso, J. *et al*. Cameo: A Python Library for Computer Aided Metabolic Engineering and Optimization of Cell Factories. *Acs Synth Biol* **7**, 1163–1166 (2018).

55. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* **44**, D471–80 (2016).
56. Chang, A., Scheer, M., Grote, A., Schomburg, I. & Schomburg, D. BRENDA, AMENDA and FRENDA the enzyme information system: new content and tools in 2009. *Nucleic Acids Res.* **37**, D588–92 (2009).
57. Hastings, J. *et al.* ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Res.* **44**, D1214–9 (2016).
58. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361 (2017).
59. Wheeler, D. L. *et al.* Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **36**, D13–21 (2008).
60. King, Z. A. *et al.* BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.* **44**, D515–22 (2016).
61. Gramates, L. S. *et al.* FlyBase at 25: looking to the future. *Nucleic Acids Res*, https://doi.org/10.1093/nar/gkw1016 (2016).
62. Thiele, I. & Palsson, B. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature protocols* **5**, 93–121 (2010).
63. Orth, J., Thiele, I. & Palsson, B. What is flux balance analysis? *Nature biotechnology* **28**, 245–8 (2010).
64. Kauffman, K., Prakash, P. & Edwards, J. Advances in flux balance analysis. *Curr Opin Biotech* **14**, 491–6 (2003).
65. Schellenberger, J., Lewis, N. E. & Palsson, B. Ø. Elimination of thermodynamically infeasible loops in steady-state metabolic models. *Biophysical journal* **100**, 544–553 (2011).
66. Price, N. D., Famili, I., Beard, D. A. & Palsson, B. Ø. Extreme pathways and Kirchhoff's second law. *Biophysical journal* **83**, 2879–82 (2002).
67. Gudmundsson, S. & Thiele, I. Computationally efficient flux variability analysis. *BMC Bioinformatics* **11**, 489 (2010).
68. Sang, J. & King, R. Nutritional requirements of axenically cultured Drosophila melanogaster adults. *Journal of Experimental Biology* **38**, 793–809 (1961).
69. Fellowes, M., Kraaijeveld, A. & Godfray, H. Association between feeding rate and parasitoid resistance in Drosophila melanogaster. *Evolution* **53**, 1302–1305 (1999).

Acknowledgements

Author contributions

J.S., L.J. and M.B. designed the study. J.S. and L.J. established the metabolic network model and performed the modeling experiments. L.J. performed the *Drosophila* experiments and most of the wet lab procedures. T.M.A. performed the GC-MS experiments. J.S., L.J., T.M.A. and M.B. analyzed the data and prepared graphs. J.S., L.J. and M.B. wrote the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-019-53532-4.

Correspondence and requests for materials should be addressed to M.B.

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.