Evaluating plant growth-defence trade-offs by modelling the interaction between primary and secondary metabolism

Jan Zrimec[1,†,*], Sandra Correa[2,4,†], Maja Zagorščak[1], Marko Petek[1], Carissa Bleker[1], Katja Stare[1], Christian Schuy[3], Sophia Sonnewald[3], Kristina Gruden[1], Zoran Nikoloski[2,4,*]

[1] Department of Biotechnology and Systems Biology, National Institute of Biology, Večna pot 111, SI1000 Ljubljana, Slovenia

[2] Bioinformatics Department, Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Str. 24-25, DE14476 Potsdam, Germany

[3] Department Biologie, Lehrstuhl für Biochemie, Friedrich-Alexander-Universität Erlangen-Nürnberg, 25 Staudstr. 5, DE91058 Erlangen, Germany

[4] Systems Biology and Mathematical Modeling, Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, DE14476 Potsdam, Germany

[†] shared first authorship

* corresponding author emails: jan.zrimec@nib.si, nikoloski@mpimp-golm.mpg.de

Abstract

Plants respond to stress by redistributing resources from growth- to defence-related processes, often resulting in decreased yields. Understanding the molecular mechanisms behind these growth-defence trade-offs can enhance breeding strategies to help us design crop varieties with improved stress tolerance, yields and quality. To enable the investigation of growth-defence trade-offs in the context of metabolism in an essential tuber crop, potato, here we develop a large-scale compartmentalised genome-scale metabolic model, Potato-GEM. Apart from primary metabolism, the model is curated to include the full potato secondary metabolism spanning over 600 reactions that facilitate the biosynthesis of 182 distinct potato secondary metabolites. Using constraint-based modelling, we find that the activation of the largest amount of secondary (defence) pathways occurs at a decrease of the relative growth rate of potato leaf, due to the costs incurred by defence. To investigate if the model can pinpoint reactions underlying growth-defence trade-offs, we then perform and process transcriptomics experiments by exposing potato leaves to two biotic stress scenarios, a chewing herbivore and a viral pathogen. By constraining the model with the resulting transcriptomics data, we obtain a set of condition-specific models, which predict decrease in optimal growth rates under treatment, in line with experimental observations. Potato-GEM thus presents a highly useful resource to study and broaden our understanding of potato and general plant defence responses under stress conditions.

1. Introduction

The challenge of ensuring a secure supply of food for the rising global population is linked to improving not just the yield and quality but also the stress tolerance of major crops [1,2]. Environmental stresses lead to annual losses amounting to billions of euros per crop. Apart from the detrimental effects of abiotic stresses, such as temperature changes, droughts and floods, biotic stresses lead to yearly losses of up to 80% of crop yield [3–5]. In the case of potato, especially damaging are viral infections and herbivore infestations, such as with Potato virus Y (PVY) [6] and Colorado potato beetle (CPB) [7,8]. Despite these concerns, the molecular processes defining and interlinking crop yields and defence responses under stress conditions are still not well understood [9,10]. Plants attacked by biotic stressors slow down their growth to preserve molecular resources and direct them for defence purposes, including production of signalling as well as defence compounds [9]. Conversely, rapid growth of plants under conditions, such as when seeking light during germination or due to a crowded environment, is often accompanied by increased susceptibility to pests and pathogens, as growth is prioritised over defence [11]. This growth–defence trade-off is a fundamental principle of plant economics, allowing plants to balance growth and defence according to external conditions [9,10]. However, modern agricultural crops including potato have often been bred to maximise growth-related traits at the expense of losing useful defence-related genetic traits [12]. To this end, improved understanding of the molecular mechanisms behind growth-defence trade-offs is a crucial step toward enhancing breeding strategies that could help design superior crops, combining high yields with the ability to defend against stress [1,13].

The stress response is often systemic, in that it occurs throughout the plant and beyond the infected or damaged tissue [14]. It is mediated by complex signalling and regulatory networks which sense and respond to environmental perturbations [15,16]. Consequently, hormones including salicylic acid (SA) and jasmonic acid (JA) induce plant resistance mechanisms to

either biotrophic pathogens, such as PVY [17,18], or herbivores, such as CPB [7], respectively. At the interface of the genetic and environmental factors lies cellular metabolism, comprising the complex network of, often, enzyme-catalysed chemical reactions that synthesise and transform substances into energy and base components necessary for growth, replication and defence, among other cellular tasks [19,20]. Here, plant growth is mediated by primary, biomass-producing processes that include photosynthesis, respiration, and the synthesis and degradation of carbohydrates, amino acids and nucleic acids [21], whereas secondary metabolism facilitates the production of a plethora of signalling and defence compounds (so-called specialised metabolites) [22,23]. Lipid metabolism is also a critical subsystem interlinking the growth and defence processes [24], by providing the backbone for many molecules in secondary metabolism. Although plant metabolism has been thoroughly modelled and studied with constraint-based mathematical modelling approaches based on genome-scale metabolic models (GEMs) [25–29], the lack of secondary metabolism reconstructions hinders the ability of these approaches to study growth-defence trade-offs in the context of metabolism. Therefore, to correctly interpret experimental data and study the effects of plant biotic interactions at the molecular level, it is imperative to refine and expand existing metabolic modelling resources to encompass not merely the primary but also the full secondary metabolism in an essential crop system, such as potato.

In the present study, we present Potato-GEM, a metabolic reconstruction of potato leaf metabolism that spans not only all critical primary and lipid metabolic processes, but also adds a full reconstruction of the known potato secondary metabolism. We then perform a general analysis of how secondary metabolite production is linked to plant growth, determining the models' inherent potential to capture growth-defence trade-offs and to predict how resource limitation affects these trade-offs. We further process and analyse transcriptomic data from real-world biotic stress experiments on potato leaves, capturing both insect pests (chewing herbivore, CPB) and pathogens (intracellular virus, PVY) interaction characteristics. By connecting the catalysis of Potato-GEM reactions with the

underlying genes, we use the transcriptomic data to constrain reaction upper bounds and build a set of condition-specific models. We find that these models indeed reflect experimental observations of decreased growth under stress conditions, serving as excellent starting points for analysis of growth-defence trade-offs.

2. Results

2.1 Constructing Potato-GEM by merging and curating multiple metabolic modules

To ensure an accurate reconstruction of potato metabolism, we first merged the metabolic model of *Arabidopsis thaliana* core metabolism (AraCore, spanning 549 reactions and 407 metabolites) [25] and the basic single-tissue model of tomato metabolism, recently updated and expanded in the Virtual Young TOmato Plant (VYTOP, spanning 2,261 reactions and 2,097 metabolites) [27] (Methods M1). The rationale for initiating Potato-GEM from these models is that tomato is a genetically and metabolically highly related plant from the *Solanaceae* family [30], whereas the Arabidopsis core metabolism constitutes the set of functionally conserved metabolic pathways across the dicot species [25]. During the merging process, we identified an overlap of 332 reactions and 346 metabolites common to both models (Figure 1A). However, the resulting model did not contain a functional secondary metabolism nor was it able to produce vital lipid-related precursors. To resolve this issue, we further integrated the model with the recently developed Plant Lipid Module [24], identifying an overlap of 296 reactions and 403 metabolites present in PLM and the merged AraCore-VYTOP model (Figure 1A). Moreover, we curated 363 reactions from 107 pathways belonging to potato secondary metabolism and 22 related precursor pathways from the MetaCyc-derived Plant Metabolic Network database [31,32], adding an additional 277 reactions and 256 metabolites (Methods M1). This resulted in the Potato-GEM model spanning 8,272 reactions and 4,982 metabolites (Figure 1A), across 16 unique compartments (Figure 1B: only key compartments shown, since the majority are related to Plant Lipid Module, Supp. files S1, S2). Compared to the tomato GEM in VYTOP, an ~5-fold reduction of blocked reactions was achieved (Figure 1C: ratio of blocked reactions was 10.5% with Potato-GEM and 51.4% with VYTOP).

We next defined a leaf biomass growth function to target within Potato-GEM metabolic simulations, quantifying the general leaf cell biomass components by performing an extensive review of the literature [24,27,33–38] that included measurements of protein content, amino acid and sugar levels of potato leaves [39]. The leaf biomass thus comprised 67 compounds and an additional 122 lipid backbones and chains (Figure 1D, Supp. table S1.1, Supp. file S3). To ensure correct optimization, the component values were further calibrated to a total of 1 g/gDW [40] by proportionally increasing their quantities and setting the overall quantity of nucleic acids, for which experimental data was not available, similar to that of proteins (Figure 1D: 14.5% of DW) [27]. We then tested if Potato-GEM can be used to simulate growth under phototrophic (light) and heterotrophic (dark) nutrient-limiting conditions (Methods M2). Under both light and dark regimes, we found that the model is indeed capable of producing all 189 biomass components (Supp. Figure 1.1). Moreover, compared with the existing AraCore and VYTOP models, Potato-GEM includes a larger number of biomass precursors across multiple component classes (Figure 1E: producing in total 71.8% or 11.7% more biomass precursors if lipids are disregarded, respectively), supporting its secondary metabolite-producing functionality.

Importantly, the reconstruction of secondary metabolism in Potato-GEM captures the complete *Solanum tuberosum* secondary metabolism as detailed in the Plant Metabolic Network database [31,32] (Methods M1). It covers the major classes including: (i) alkaloids, such as: alpha-solanine and alpha-chaconine, calystegines and tropane alkaloids, (ii) phenylpropanoid derivatives, such as: flavonoids, coumarins, cinnamates, lignans and lignins, (iii) terpenoids, incl. carotenoids and mono-, di-, tri- and sesquiterpenoids, (iv) phytoalexins, i.e. resveratrol and capsidiol, and (v) hormones, namely: jasmonic acid (JA) and salicylic acid (SA), respectively, among the total of eight key hormonal types (Figure 2A, Supp. figure S1.2). This enables the modelling of the production of defence compounds related to stress response and immune signalling, allowing us to study the effects and

coupling of growth with secondary metabolite production [19,41]. The complete module spans over 600 reactions in 107 pathways, producing a total of 182 unique secondary metabolites (Supp. figures S1.2, S1.3, Supp. table S1.2, Supp. file S4).

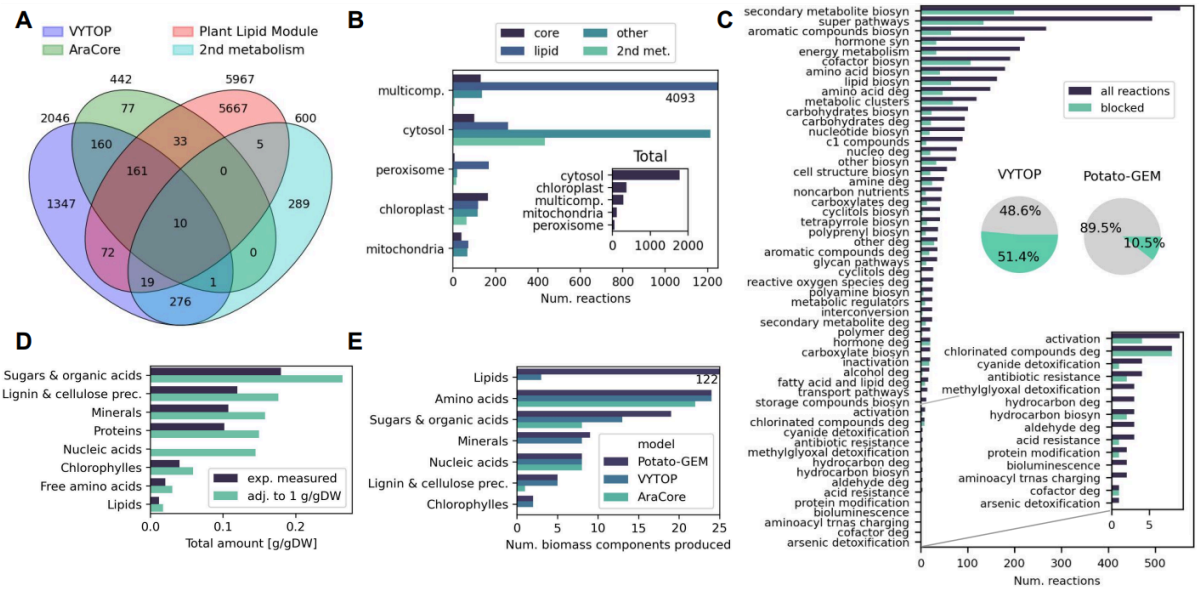

Figure 1. Constructing Potato-GEM by merging and curating multiple modules. (A) Venn diagram of the four models that comprise potato-GEM, including AraCore [25], the basic single-tissue model from Virtual Young TOmato Plant (VYTOP) [27], Plant Lipid Module [24] and additional manual curation of secondary metabolism based on the MetaCyc database [32,42]. (B) Number of core, lipid and other primary as well as secondary metabolism reactions across the model's key cellular compartments including multi-compartment reactions. Inset shows the total number of reactions. (C) Depiction of the total number and blocked reactions across the MetaCyc pathway ontology. Inset pie charts show the ratio of blocked reactions in VYTOP and potato-GEM. Lower inset shows a zoom in on the values of the bottom 14 pathways. (D) Depiction of potato biomass components based on experimentally measured values and after adjustment of total content to 1 g/gDW [40]. (E) Quantification of the capability to produce biomass precursors across potato-GEM and the existing VYTOP [27] and AraCore [25] models.

2.2 Secondary metabolism reconstruction enables quantifying growth-defence trade-offs

The refined Potato-GEM model enabled us to investigate the coupling [43] between the potato basic growth and stress response modes to ascertain the possibility of growth-defence trade-offs. To this end, we investigated if fluxes through secondary pathways and reactions are coupled to biomass production (growth, Figure 2B). Here, the variability of fluxes of the final product-producing reactions in each secondary pathway were evaluated within the optimal biomass space using flux variability analysis (FVA) [44], where different fractions, ranging from 0 to 1, of the optimal relative growth rate were used (Figure 2C). We observed that at the optimal relative growth rate, the majority (96%) of secondary pathways were either inactive, with computed flux ranges equalling 0, or minimally active, with a limited flux range below 1, as occurred with 6 reactions (Figure 2C, Supp. figure S2.1). The exception was cytokinin biosynthesis (PWY-2681) that was positively coupled with growth, as its flux lower bound increased proportionally to the relative growth rate (Supp. figure S2.2). Furthermore, with a decreasing relative growth rate from the optimum, the majority of pathways exhibited proportional increases of their flux ranges (Figure 2C), supported by a significant negative correlation (Spearman ρ = -0.55, p-value < 10^{-16}, Supp. figure S2.3) between growth and flux through secondary metabolite production. This demonstrated that the majority of secondary pathways are thus negatively coupled with the plants' growth objective (Figure 2D: 155 reactions across 86 pathways). In addition, of these negatively coupled pathways, almost all (98%) were found to be strongly negatively correlated (Spearman ρ < -0.99, p-value < 8×10^{-12}, Supp. figure S2.3) with the relative growth rate. The remaining pathways were either positively coupled (1 reaction and pathway), exhibited non-monotonic and thus unstable changes (17 reactions across 9 pathways), or were inactive across the whole range of biomass production (25 reactions across 20 pathways, Figure 2D). Importantly, an optimal secondary metabolite production was observed at a

relative growth rate of 0.4, where the largest amount of secondary reactions and pathways were found to be active (Supp. figure S2.1).

The observed negative coupling between growth and secondary metabolism (Figure 2C) suggested that a growth-defence trade-off is occurring with these pathways. Therefore, we further investigated how much a unit increase of flux through secondary pathway production (defence) decreases the flux through the biomass reaction (growth), which can be interpreted as a growth-defence trade-off factor (Figure 2B, Methods M2). On average, across all coupled secondary pathways (Figure 2D), we observed a 3×10^{-4} decrease of biomass flux with an unit increase of secondary metabolite production (Figure 2E). Interestingly, this growth-defence trade-off factor was ~5-fold and significantly (Wilcoxon rank-sum test *p*-value < 0.0052) higher than a trade-off factor obtained in response to unit increase of flux through primary metabolic reactions (Figure 2E: ~10^{-5}). With primary metabolic response, both the core and lipid metabolism resulted in the lowest trade-offs on average, though lipid reactions also achieved the highest trade-offs of over 10^{-3} (Figure 2E). Moreover, we observed that the growth-defence trade-off factors remain on average relatively equal across different secondary metabolite classes (Figure 2F). Here, the largest factor of 4.7×10^{-4}, ~1.6-fold larger than the overall secondary metabolism average, was found for alkaloids (note that difference to other classes not significant), possibly since these are some of the most resource costly compounds to produce [45]. Overall, the results suggest that under stress, secondary metabolism activation under defence responses requires a relatively higher amount of resources (Figure 2E: > 5-fold more) than the general rewiring of core metabolism.

Finally, we explored the hypothesis that resource constraint is a primary reason for the inverse growth-defence relationship, meaning that providing more resources would reduce growth-defence trade-offs and allow plants to simultaneously grow and defend [10,46]. To this end, we proportionally limited or expanded the model's key resource inputs: CO_2, light,

nitrogen or a combination of all three (Figure 3A, Methods M2). We observed proportional decreases in the predicted growth rates when limiting resources, reaching no growth when resources were completely withdrawn (Figure 3B). Conversely, as expected, no change in growth was observed when increasing resource availability above a factor (resource ratio) of 1 (Supp. figure S2.4). Similarly as with growth, the secondary metabolite production capacity decreased significantly (Wilcoxon rank-sum test *p*-value < 10^{-14}, measured between the resource ratio of 1 and 0) and in proportion to the limited resources (Figure 3C). The exception was with nitrogen limitation, where the defence response decreased by merely 6% when varying nitrogen availability between a factor of 1 and 0. Increasing resource availability, on the other hand, did not lead to an increased secondary production capacity (Supp. figure S2.5), likely since the model's optimal secondary metabolite production was already reached at the relative growth rate of 0.4 within the given constraints.

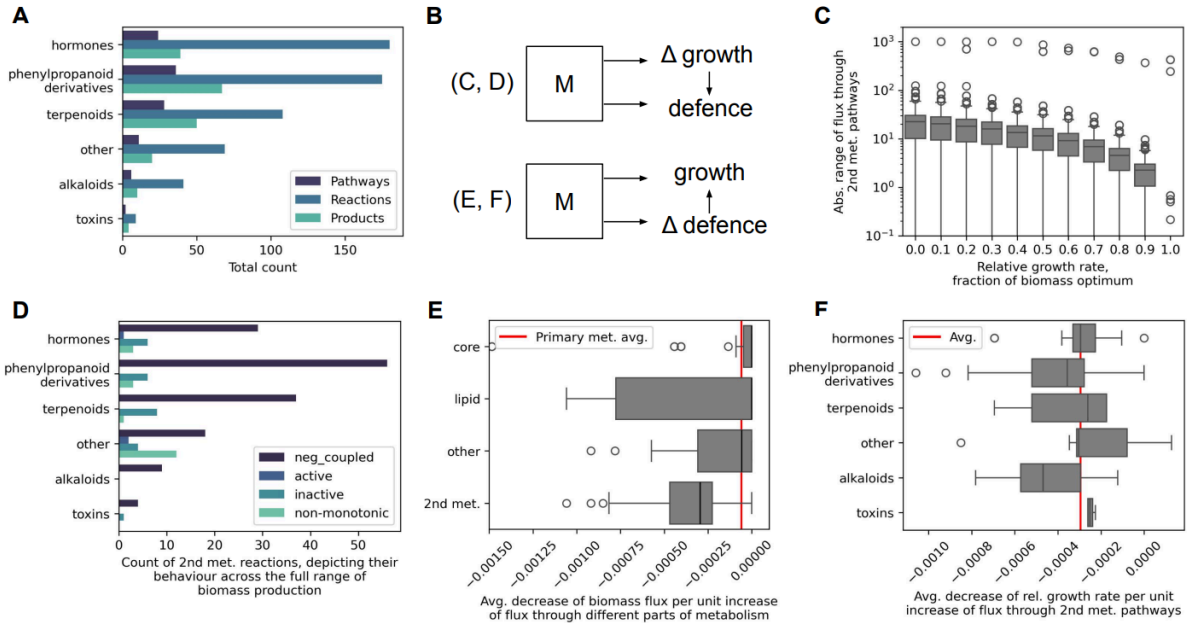

Figure 2. Secondary metabolism reconstruction enables quantifying growth-defence trade-offs. (A) The number of reactions, pathways and products across secondary metabolism classes in the reconstruction. (B) Schematic depiction of the studied growth-defence trade-off relationships by either observing the effects of growth limitation on defence activation (as depicted in panels C, D) or vice versa (panels E, F). (C) Depiction of predicted growth-defence trade-off

capabilities as the absolute range of flux through secondary metabolic pathways in response to the varying ratio of optimal biomass production in FVA. (D) Segregation of secondary pathway reactions across their coupling behaviour with the growth objective, showing that the majority of these reactions are negatively coupled with growth and exhibit growth-defence trade-offs. (E) Distribution of the decrease of biomass flux per unit increase of flux through different parts of metabolism including core, lipid and secondary metabolism. Red line denotes the median value of primary metabolism. (F) Distribution of the decrease of biomass flux per unit increase of flux through pathways across secondary metabolism classes. Red line denotes median value across the whole secondary metabolism.

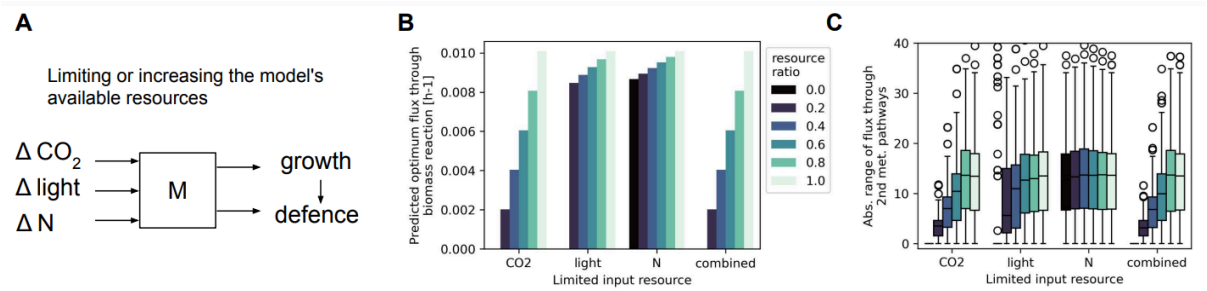

Figure 3. Exploring the effects of resource limitation on growth and defence.

(A) Schematic depiction of the analysis of resource limitation or expansion, where the key resources CO2, light and nitrogen were varied within a range of 0 to 1 relative to the resource consumption flux obtained at an optimal growth rate of 1. (B) Effect of resource limitation on the predicted relative growth rate (biomass optimum). (C) Effect of resource limitation on the range of secondary metabolite production (defence), computed at the fraction of 0.4 of the relative growth rate to observe the largest amount of possible active secondary metabolism.

2.3 Capturing biotic stress responses with transcriptomics

Our next aim was to investigate a range of common potato biotic stress scenarios and the growth-defence trade-offs that they elicit. To this end, we performed a transcriptomics experiment based on the herbivore attack of potato leaves with the Colorado potato beetle (CPB) [47] and complemented it with previously published data on the pathogen interaction with the potato virus Y (PVY) [48] (Figure 4A). Briefly, within the presently performed CPB experiment, plants were exposed to 2 beetles per leaf for 30 minutes and 24 hours post infection, and the leaf region surrounding the damaged part was sampled in parallel with non-infested leaves (control) and processed for RNA-Seq (Methods M3,4). Within the previously published PVY experiment [18], leaves were inoculated with PVY and tissue immediately surrounding the site of viral multiplication was sampled when the hypersensitive resistance response was fully established (4 days after inoculation) in parallel with control non-inoculated leaves tissue and processed for RNA-Seq (Methods M4). In both experiments, biological triplicates were analysed per treatment.

Non-metric multidimensional scaling (NMDS) analysis of transcript counts indicated that clear separation trends among treated and control samples were obtained in both experiments (Figure 4B, Supp. figure S3.1, Methods M4). We also observed significant correlation (Spearman ρ = 0.775, p-value < 10^{-16}) of transcript counts among the control samples of both experiments (Supp. Figure S3.2). However, whereas correlation analysis among all 12 samples showed a clear distinction between PVY control and treatment samples, this was not observed with CPB samples (Figure 4B). The CPB experiment also captured a smaller number of differentially expressed genes (325 DEGs), compared to the PVY experiment (9,289 DEGs) (see Methods M4). These differences were likely due to the shorter CPB exposure time and thus milder observed defence response than with PVY [18]. Nevertheless, further DEG and enrichment analysis using plant specific MapMan ontology terms [49] showed (i) an increase in hormonal production, specifically salicylic acid (SA) with

PVY [17] and jasmonic acid (JA) with CPB [8,50], which are known signalling cascade regulators of the specific biotic stresses, (ii) general upregulation of multiple secondary pathways, including those belonging to glycoalkaloid [51], phenylpropanoid [52,53] and terpenoid classes [54], (iii) lowered photosynthetic activity with PVY [55], (iv) upregulated sucrose degradation factors with PVY likely due to increased sucrose accumulation [53], as well as (v) downregulation of biotic stress response-related factors due to an overactivated signalling response [16]. This suggested that known responses as a consequence of biotic stress and growth-defence trade-offs were indeed captured within both experiments (Supp. files S5,6) [18].

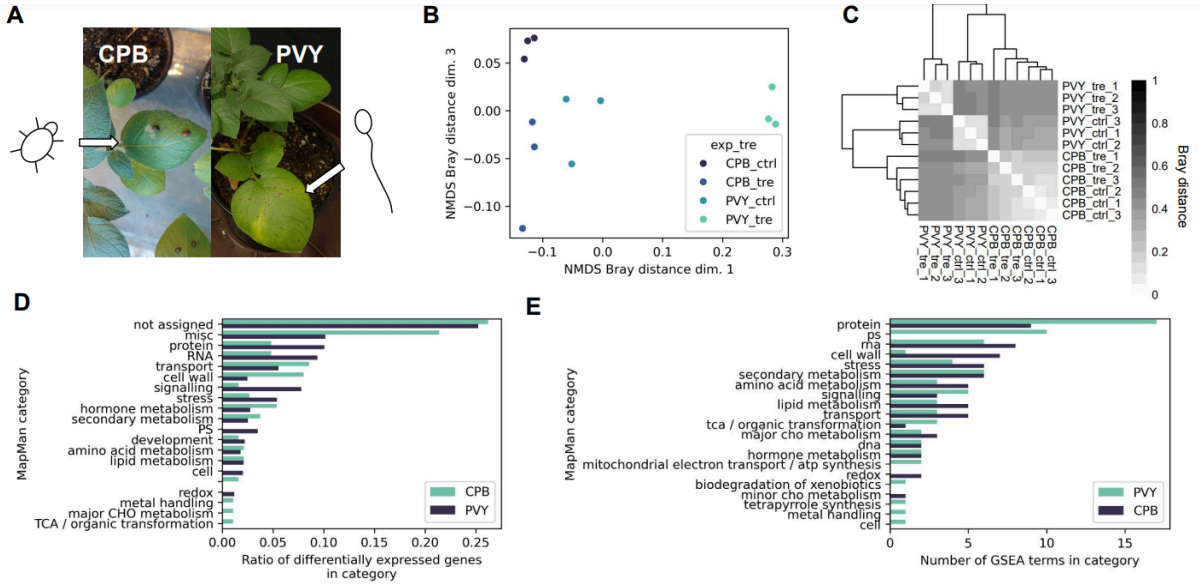

Figure 4. Capturing biotic stress responses with transcriptomics. (A) Schematic depiction of biotic stress transcriptomics experiments with Colorado potato beetle (CPB) and Potato virus Y (PVY). (B) Principal-coordinate analysis of Bray-Curtis dissimilarity between transcript counts across experiments and treatments. (C) Heatmap and dendrogram visualisation of hierarchical clustering on Bray-Curtis dissimilarity between sample replicates across experiments and treatments. (D) Ratio of differentially expressed genes (DEGs) (BH-corrected *p*-value < 0.05 and abs(log2FC) > 2) belonging to a MapMan [49,56] category versus the total number of DEGs across experiments. Ratios above 0.025 shown. (E) Transcriptomics responses on the level of pathways. MapMan-defined pathways [49,56] enriched with differentially regulated genes were calculated using GSEA [57] (BH-corrected p-value < 0.05).

2.4 Transcriptome-constrained models recapitulate observed stress responses

To further study the characteristics of biotic interactions, we next constructed models constrained by the transcriptomics data (Figure 5A). To this end, we first annotated model reactions with gene protein reaction (GPR) rules, which were obtained from the MetaCyc-derived Plant Metabolic Network database [31] (*Solanum tuberosum* subset), as well as by translating GPRs with Arabidopsis gene identifiers from the Plant Lipid Module [24] and AraCore [25] modules (Methods M5). Here, orthologous genes between Arabidopsis and potato were identified according to Plaza v5.0 [58] orthologs using clustering algorithms as well as BLAST reciprocal best hit search [59,60]. This yielded a total of 3,134 unique potato gene identifiers and resulted in 3,950 reactions (48%) annotated with GPR rules comprising, on average, 2 unique gene identifiers (Figure 5B). Unsurprisingly, better coverage of GPR-annotated reactions of 65% was achieved with the core metabolism compared to a 50% coverage of both the lipid and secondary metabolism (Figure 5B inset). However, secondary metabolic reactions were annotated with GPR rules that comprised significantly (Wilcoxon rank-sum test p-value $< 3.3\times10^{-13}$) more unique gene identifiers than primary metabolic ones, with a 1.5-fold higher number of genes on average (Figure 5C). Within the secondary metabolism, apart from precursors exhibiting the highest coverage (83%), terpenoid, toxin and hormone classes achieved a better than average coverage of over 48% (Figure 5D).

Next, we integrated the transcriptomics data with the Potato-GEM model, by evaluating GPR rules based on the provided transcriptomic measurements in order to constrain the model's upper bounds across reactions. This resulted in 12 transcriptomics-constrained models, one for each replicate of the data treatment (control and treated) and experiment type (CPB and PVY, see Figure 4). An initial 14,404 and 19,367 transcripts from the CPB or PVY

experiment, respectively, mapped to 3,662 and 3,716 unique protein complexes in GPR rules across 3,561 and 3,467 reactions, respectively (Supp. figure S4.1). Of the GPR-annotated reactions, ~46% (Supp. figure S4.1: 1,725 and 1,692, respectively) were essential. We observed that the variability of the upper bounds of the CPB constrained models differed significantly (Wilcoxon rank-sum test *p*-value < 10^{-16}) across treatments, where models corresponding to treated samples exhibited an average 2.5-fold increase in variability of upper bounds compared to controls (Figure 5E). However, this was not the case with PVY, where the variability remained approximately equal between treated and control models. Similarly, no significant difference in upper bound variability was observed for essential in comparison to non-essential reactions (Supp. figure S4.2). Moreover, the transcriptomics-constrained reactions were found to comprise the full range of pathways covering 96% of the key cellular processes as existing in the original unconstrained model (Supp. figure S4.3, see Figure 1C).

Finally, we observed that the predicted biomass production with the constrained models (Methods M2: phototrophic regime as applied in the experimental setup) recapitulated our experimental observations, where under both biotic stresses, treated plants displayed lower growth rates on average compared to controls (Figure 5F: 0.39 of the optimal growth rate observed with CPB and 0.14 with PVY, respectively, note that significance not observed due to low amount of replicates). The overall 73% decrease in growth rate under biotic stress treatments was also in line with published observations under field conditions, where yield losses of up to 80% were observed both under CPB and PVY induced stresses [4,5]. This suggested that the transcriptomics-constrained models represent accurate metabolic replicas of the *in vivo* stress scenarios, supporting their usefulness to further study biotic stress-induced growth-defence trade-offs in the context of metabolism.

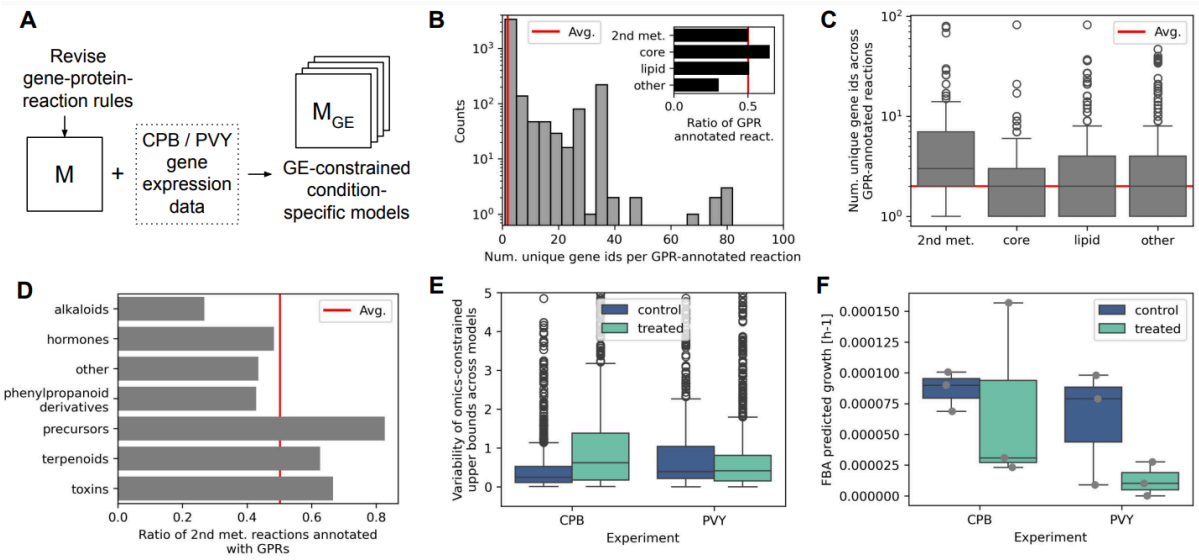

Figure 5. Transcriptome-constrained models recapitulate observed stress responses. (A) Schematic depiction of the procedure to construct a set of CPB and PBY-transcriptome constrained models. (B) Number of unique gene identifiers present in GPR rules across GPR-annotated reactions. Red line denotes an average value of 2. Inset depicts the ratio of GPR-annotated reactions across the general metabolic types. (C) Number of unique gene identifiers present in GPR rules of GPR-annotated reactions across the general metabolic types. Red line denotes the average value. (D) Ratio of GPR-annotated reactions across secondary metabolism classes. Red line denotes the average value. (E) Variability of model upper bounds of the transcriptomics-constrained reactions across treatments and experiments. (F) Predicted biomass production with the CPB and PVY transcriptome-constrained models under a phototrophic regime.

3. Discussion

In the present study, we asked whether we can "interpret" plant growth-defence trade-offs in the context of differences in metabolic fluxes, by constructing a sufficiently comprehensive and accurate model of primary and secondary metabolism in the major crop plant, potato. To this end, we developed and applied potato-GEM, a genome-scale metabolic model constructed by merging multiple models [25,27] and modules [24] (Figure 1A), and includes a reconstruction of the full potato secondary metabolism according to the Plant metabolic network [31] MetaCyc database [32] (Figure 2A). We demonstrated how, due to its comprehensiveness and the secondary metabolism reconstruction, the model enables exploring the principles of plant growth-defence trade-offs in the context of metabolism and under biotic stress.

Similarly as normal growth [25,27], the plant stress, or defence response comprises a large and complicated set of cellular processes [15,16,41]. In contrast to experimental research that may be limited to studying merely a subset of different molecular responses at once [9], *in silico* mathematical modelling of cellular processes such as metabolism [24,25] and signalling [15,16], enables us to capture and examine the full range of cellular responses en masse. The potato-GEM model thus allowed us, for the first time, to observe, explore and interpret the rewiring of plant metabolism under decreasing growth to a defence response. The analysis of growth-defence trade-offs was performed in two ways. First, by directly probing the model using flux balance and flux variability analysis (FBA and FVA, respectively) (Figure 2B and 3A), either by (i) decreasing the relative growth rate (fraction of biomass optimum) and observing how defence pathways are unlocked (Figure 2C,D), (ii) varying the secondary metabolite production and determining its effects on growth (Figure 2E,F), or (iii) testing the effect of limiting different key input resources on the growth-defence trade-offs (Figure 3B,C). Second, by constraining the model according to experimental gene expression

measurements under biotic stress (Figure 4), and then analysing model predictions by performing FBA (Figure 5F).

Following that, on average, the effects of decreasing the relative growth rate (flux through the biomass reaction) were significantly negatively correlated with increasing flux through secondary pathways (Figure 2C), the initial, direct approach enabled us to quantify specific general principles of growth-defence trade-offs. Firstly, this included determining that the largest number of secondary (defence) pathways are active at a decrease of ~60% of the potato relative growth rate (Figure 2C: fraction of biomass optimum of 0.4). Secondly, compared to primary metabolism, we find that the secondary metabolism (defence) response is generally the most costly set of metabolic processes in the plant cell, based on measuring the effect of perturbing secondary metabolite production (Figure 2E: unit change flux through secondary pathways) on growth (flux through biomass reaction). The cost of diverted resources is also, on average, approximately equal across secondary metabolite classes (Figure 2F). Indeed, it is known that due to the costliness of diverting energy and resources away from key processes such as growth and reproduction [61], cells employ multiple strategies to lower the high metabolic costs of defence [23]. Apart from fine-tuning secondary metabolite production via gene expression, metabolite multifunctionality and effective recycling [23], balancing the objectives between biomass or defence production is a key cellular strategy defined by the spectrum of resistance and tolerance of the underlying plant genotype [61]. Thirdly, we measure that resource limitation has a profound effect on growth-defence trade-offs, effectively decreasing defence rewiring due to a lack of resources (Figure 3C). This supports the existing resource allocation hypothesis [46,62], which states that resource constraints are a primary reason for the inverse growth-defence relationship. Namely, since plants typically do not optimise merely the growth objective and thus do not necessarily grow at the metabolic optimum, they can find themselves somewhere between the minimum and maximum modelled growth rate (Figure 3B: 0 to 1 relative growth rate). Thus, providing more resources, such as increasing nutrients, CO2 and/or light (Figure 3C),

would reduce growth-defence trade-offs, resulting in more secondary metabolism activation at a higher growth rate and thus allowing plants to simultaneously grow and defend [10].

To apply Potato-GEM to study growth-defence trade-offs under biotic stress, we first performed and processed transcriptomics experiments exposing potato leaves to CPB (this study) and PVY [18], respectively (Figure 4A). According to differential gene expression and pathway enrichment analyses, both experiments captured the full expected response of their corresponding stresses (Figure 4D,E), despite the CPB experiment characterising a more early stress response than was the case with PVY [18]. By revising the Potato-GEM gene-protein-reaction (GPR) rules that link gene expression and reactions (Figure 5A,B), we were able to constrain reaction upper bounds according to the measured gene expression levels [63,64]. This resulted in a set of condition-specific models, which, when predicting optimal growth rates using FBA, were capable of recapitulating the experimentally observed decreases in growth rates under treatment (Figure 5F). Especially with CPB, the ~60% decrease in growth rate corresponded well with our prior findings of optimal growth-defence trade-offs at a similar decrease in relative growth rate (Figure 2C).

To our knowledge, this is presently the largest and most capable metabolic model of potato [29], the third most consumed food crop globally. The secondary metabolism reconstruction is also useful across the *Solanaceae* family, which includes tomato, tobacco and pepper. These are all economically highly-important crops, with large amounts of published data on biotic stress available, and close relatives of potato, meaning that they share the majority of their metabolism including secondary metabolism with potato [30,31,65,66]. Therefore, repurposing potato-GEM to study related organisms requires merely (i) adapting the secondary metabolism, (ii) reconfiguring the biomass function and (iii) redefining gene-protein reaction rules, by finding orthologs from potato or Arabidopsis or obtaining them from online databases. Besides repurposing potato-GEM across related species, there are potentially multiple further avenues of future development with this model, such as

multi-tissue [27,67] and multi-species [68,69] modelling. To conclude, we demonstrate that Potato-GEM presents a highly useful resource to study and broaden our understanding of potato as well as general plant defence responses in biotic interactions and under environmental stress.

4. Methods

M1. Metabolic model construction

To develop a reconstruction of potato metabolism, we used three existing genome scale metabolic models (GEMs): (i) model of *Arabidopsis thaliana* core metabolism (AraCore, spanning 549 reactions and 407 metabolites) [25], (ii) the basic single-tissue model of tomato metabolism from the Virtual Young TOmato Plant (VYTOP, spanning 2,261 reactions and 2,097 metabolites) [27], and (iii) Plant Lipid Module (spanning 5,956 reactions and 3,108 metabolites) [24]. To enable model intercompatibility and merging, the different reaction and metabolite identifiers of the older AraCore and VYTOP models were updated, including BioCyc [32] ids, Enzyme Commission (EC) numbers, KEGG [70] ids, International Chemical Identifier (InChI) strings [71] and MetaNetX [72] MNXref ids, using manual searching and custom Python scripts. Prior to model merging, VYTOP was further curated to remove reactions not supported by present knowledge or located in erroneous compartments.

To merge the AraCore and VYTOP models, each reaction was described as a vector of its metabolites, where each metabolite was represented using its MNXref and compartment identifiers that enable maximum uniqueness. Then, the cosine distance (Python package scipy v1.9.3) [73] among the metabolite vectors was used to determine the overlapping reactions between the models that were merged. A cosine distance threshold of 0.2 was used and all non-perfectly overlapping reaction pairs were manually checked to remove spurious merges. Next, the AraCore-VYTOP merged model was extended by integrating a lipid metabolism reconstruction from the Plant Lipid Module [24]. Prior to this, the merged model was verified according to several criteria, including (i) having a COBRA model structure [74]; (ii) including metabolites with neutral and/or charged molecular formulas, accompanied by the respective charges and identifiers (e.g., KEGG, ChEBI, PubChem and/or Lipid Maps); and (iii) being able to be imported and exported with standard software

packages/toolboxes (e.g., via COBRA Toolbox) [75]. Next, the accompanying software tool that allowed semi-automated integration of the lipid network was used [24], which was followed by the removal of reactions that were found to be duplicated.

After merging the three metabolic models, we manually curated the potato secondary metabolism of the resulting model according to the Plant Metabolic Network database v15 (PotatoCyc v6.0.1) [31,32]. The database was accessed both via its web interface (pmn.plantcyc.org) as well as the BioCyc software Pathway-tools v25.5 [42,76] and the PythonCyc v2.02 package. Additionally, the KEGG [70] database (/www.genome.jp) and published literature were queried for specific information on phenylpropanoid precursors [77], glycoalkaloid [78,79] and calystegine metabolism [80] as well as cross-compartmental biosynthesis and transports in relation to shikimate [81] and hormones biosynthesis [82–85]. For all secondary pathways the capacity to produce their precursors was verified, where alternatively 6 compounds that at present cannot be produced by the model (cytosol L-quinate and D-glucarate for caffeoylglucarate biosynthesis, UDP-D-apiose for apigenin glycosides biosyn., sinapoyl-CoA for hydroxycinnamic acid tyramine amides biosyn., tRNA-Adenosines-37 for cis-zeatin biosyn. and CDP-diacylglycerol for D-myo-inositol(1,4,5)-trisphosphate biosyn.) are fed to the model externally as sinks. In total, 363 reactions were curated, where an additional 277 reactions and 256 metabolites were added to the model. The final secondary metabolic module thus spans 600 reactions from 107 secondary pathways and 22 related precursor pathways and can produce 182 distinct secondary metabolites.

Finally, we revised the model's mass and charge balance [86]. Metabolite chemical formulas and charges were updated according to the MetaNetX resource [72] and InChI strings using automated procedures, including custom Python scripts and the Matlab COBRA *getChargeFromInChI* function, as well as manual curation. Custom Matlab scripts were used for full reaction mass and charge balancing, with the exception of 131 reactions that remain charge imbalanced due to chemical formula and charge discrepancies with specific

metabolites. The complete procedure resulted in the Potato-GEM model spanning 8,272 reactions and 4,982 metabolites across 16 unique compartments.

M2. Constraint-based modelling procedures

For constraint-based analysis of metabolic models, the Matlab COBRA v2.42.0 [75] and Python COBRApy v0.26.2 [87] toolboxes were used. The IBM CPLEX optimization studio v12.10.0 and Gurobi Optimizer v9.5.1 were used with the Matlab and Python toolbox versions, respectively. For the phototrophic (light) regime, the metabolites NO3, NH4, SO4, Pi, H, H2O, CO, Ca2, Cl, Fe2, K, Mg2, Na, H2S and photon were set to sink reactions, whereas in the heterotrophic (dark) regime, starch and O2 replaced photon and CO2 from the phototrophic regime. To compute how much a unit increase of flux through secondary pathway production (defence) decreases the flux through the biomass reaction (growth) (see Figure 2B), we first computed the flux bounds of each secondary pathway final reaction at the relative growth rate of 0.4 using flux variability analysis (FVA) [44]. The growth-defence trade-off factor for each secondary pathway was then obtained as the difference between the growth rate (optimal biomass flux obtained using flux balance analysis, FBA) [88] obtained by overstepping the maximum secondary metabolite production (upper bound) by one unit and the growth flux obtained at the maximum production (upper bound). The analysis of resource limitation (Figure 3A) was performed by proportionally limiting (within a range of 0 to 1) or expanding (within a range of 1 to 10) resource inputs, including CO2, light (photon flux) and nitrogen or a combination of all three, according to the fluxes of these resources obtained at the optimal growth rate of 1 (Figure 2C). Growth rates (optimal biomass fluxes) were then computed at the different ratios of resources using FBA. Lower and upper bounds of flux through secondary pathways were computed using FVA at the secondary metabolism-optimal growth rate of 0.4 to ensure that the largest amount of possible active secondary metabolism was observed. Note that nitrogen resources refer to the molecules NO3 and NH4.

M3. Experimental procedures

To capture potato biotic stress responses with the Colorado Potato Beetle (CPB), potato cv. Desiree plants were propagated from stem node tissue cultures and transferred to soil 2 weeks after node segmentation. The plants were then kept in growth chambers under controlled environmental conditions at 22/20°C with a long-day (16 h) photoperiod of light (light intensity 4,000 lm/m2) and 45–60% relative humidity. For the CPB treatment, two starved 3rd instar larvae were placed on the potato on the first fully developed leaf and left feeding for 30 minutes (Figure 4A). The damaged part of the leaf was excised using a scalpel and immediately frozen in liquid nitrogen. Total RNA was isolated and processed as described in Petek et al. (2014) [8]. Poly-A selected RNA-seq library prep and paired-end 150 nt Illumina sequencing was performed by Novogene Inc. (www.novogene.com).

M4. Transcriptomics data analysis

To process the transcriptomics data obtained with CPB experiments (see Methods M3), we performed quality control of reads, merging of overlapping pairs, mapping of reads to the potato genome (PGSC v4.03 genome model) [89] and read counting using the CLC genomics workbench v12.0 (Qiagen). Differential expression analysis was performed using a standard R package DESeq2 v1.42.1 [90] workflow. Genes with a Benjamini Hochberg (BH)-adjusted *p*-value below 0.05 were considered differentially expressed (Figure 4D, Supp. file S5). Gene set enrichment analysis (GSEA) [57] was performed using MapMan ontology [49] as the source of the gene sets (obtained from GoMapMan database) [56] and processed with custom R scripts (Supp. file S6). Principal-coordinate analysis of Bray-Curtis dissimilarity and hierarchical clustering of sample replicates across experiments and treatments was performed using the R packages vegan v2.6-4 [91], ggordiplots v.0.4.3 [92], and pheatmap v1.0.12 [93]. Processed transcriptomic data of potato biotic stress responses with Potato Virus

Y was used from Lukan et al. 2020 [18], where the RNA-seq analysis of the tissue within the lesion section A was used as treatment and mock as control.

M5. Revision of gene-protein-reaction associations

To link the reconstructed model's reactions with the genes responsible for expressing the underlying enzymatic components (termed gene-protein-reaction, GPR associations) [86] we applied the following procedure. Arabidopsis gene identifiers were translated to potato gene ids for reactions from AraCore [25] and Plant Lipid Module [24] with GPRs assigned. Of the 1,014 unique Arabidopsis gene ids in total contained within GPRs from these modules, 980 ids were successfully translated to potato gene ids. The Arabidopsis to potato gene id translation was based on a hierarchical procedure of translations, where initially the PLAZA v5.0 [58] comparative genomics resource was used as a source of orthologous gene families in Arabidopsis and potato, serving as the basis to translate Arabidopsis gene ids to potato ids. Of a starting pool of 6,511 Arabidopsis gene ids, present in the combined AraCyc v17.1.1 (Plant Metabolic Network v15) database [31] as well as AraCore [25] and Plant Lipid Module [24], we identified 6,467 present in PLAZA, 5,872 occurring across 2,122 families with at least one potato gene (and 595 Arabidopsis genes occur in families without any potato gene). We extracted Arabidopsis (Araport11) [94], potato (JGI v4.03) [89], tomato (ITAG4.0) [30], and poplar (JGI v4.1) [95] gene sequences from the families with a sufficient number of genes (1,708 families), and performed all to all computations of sequence Levenshtein distances (R package stringdist v.0.9.10) [96] between the genes within each family. Three of the families had high similarity between all genes (distances of less than 0.08), and we assumed all-to-all translations for these. For the others, we transformed the Levenstein distance to a similarity score (using e^(-x^3), scaled between 0 and 1), and applied a “whole-graph” similarity threshold per family (the maximum threshold whereby each gene still has at least one similarity to another gene above the threshold) [97]. For the families with 100 or fewer members, we performed Louvain clustering (Python package louvain v0.8.2, incrementally

increasing the threshold if clustering did not converge) [98], resulting in translation groups for 3,347 of the Arabidopsis genes. For the remaining 32 families with more than 100 members (too large to cluster using Louvain) we performed Markov clustering (inflation parameter of 5) [99], resulting in translation groups for 624 further genes. Each of these translation groups contained at least one Arabidopsis and at least one potato identifier for 3,975 Arabidopsis genes, which rendered translations to potato for 712 of the 1,014 Arabidopsis GPR gene ids. Next, the remaining gene identifiers were translated using BLAST v2 [59,100] reciprocal best hit search (RBB) on Arabidopsis vs. potato PGSCv4 proteomes, where the highest scoring hits were used as translations. Default parameters were used with both approaches, with the top three potato gene ids per Arabidopsis id considered. ITAG potato identifiers were additionally consolidated to PGSC using an *in-house* generated translation table (files available from https://github.com/NIB-SI/DiNAR/tree/master/TranslationTables) [101]. In case of many-to-many relationships, the translation pair was prioritised based on the differential expression and raw count values, selecting the most up- or down-regulated gene.

Finally, the model was revised with potato-translated GPR associations by preferentially including translated GPRs from AraCore and then from the Plant Lipid Module, for 315 and 3036 reactions, respectively, and finally GPRs from the PotatoCyc v6.0.1 (Plant Metabolic Network v15) database [31] for a further 806 reactions.

M6. Software and data

Python v3.8.16 (www.python.org), R v4.3.1 (www.r-project.org/) and Matlab v2019b (www.mathworks.com) were used for computations. For statistical hypothesis testing, Scipy [73] v1.9.3 and Statsmodels v0.14.0 were used with default settings. All statistical tests were two-tailed except where stated otherwise. For correlation analysis, the Spearman correlation coefficient is reported and statistical significance was tested using Welch's t-test. For general data analysis and visualisation, the Python packages Pandas v1.5.2, Biopython v1.78, and

Seaborn v0.13.2 were used. The model and related files were deposited to the Github repository and are available at https://github.com/NIB/Potato-GEM.

Acknowledgements

We thank Živa Ramšak for technical support, as well as Phillip Wendering and Anže Županič for technical discussions and critical comments. The study was supported by the Slovenian Research and Innovation Agency (ARIS) grants no. J2-3060 (J.Z.), Z4-50146 (C.B.) and P4-0165 (K.G.), the Public Scholarship, Development, Disability and Maintenance Fund of the Republic of Slovenia grant no. 11013-9/2021-2 (J.Z.), the European Union Horizon 2020 Framework Programme under grant agreement no. 862858 (https://adapt.univie.ac.at/, K.G.). The European Union's Horizon 2020 research and innovation program also financially supported this work through grant 862201 (to Z.N.). Z.N and S.C.C. acknowledge the financial support by the Collaborative Research Center 1644 funded by the German Research Foundation.

References

1. Hickey, L. T. *et al.* Breeding crops to feed 10 billion. *Nat. Biotechnol.* **37**, 744–754 (2019).
2. Murphy, D. *Plants, Biotechnology and Agriculture*. (CABI, 2011).
3. Savary, S. *et al.* The global burden of pathogens and pests on major food crops. *Nat Ecol Evol* **3**, 430–439 (2019).
4. Maharijaya, A. & Vosman, B. Managing the Colorado potato beetle; the need for resistance breeding. *Euphytica* **204**, 487–501 (2015).
5. Petrov, N. M., Stoyanova, M. I. & Gaur, R. K. Biodiversity and characterization of economically important viruses on potato cultivars. in *Plant RNA Viruses* 245–270 (Elsevier, 2023).
6. Baebler, Š., Coll, A. & Gruden, K. Plant Molecular Responses to Potato Virus Y: A Continuum of Outcomes from Sensitivity and Tolerance to Resistance. *Viruses* **12**, (2020).
7. Chung, S. H., Rosa, C., Hoover, K., Luthe, D. S. & Felton, G. W. Colorado potato beetle manipulates plant defenses in local and systemic leaves. *Plant Signal. Behav.* **8**, e27592 (2013).
8. Petek, M. *et al.* Potato virus Y infection hinders potato defence response and renders plants more vulnerable to Colorado potato beetle attack. *Mol. Ecol.* **23**, 5378–5391 (2014).
9. Huot, B., Yao, J., Montgomery, B. L. & He, S. Y. Growth-defense tradeoffs in plants: a balancing act to optimize fitness. *Mol. Plant* **7**, 1267–1287 (2014).
10. He, Z., Webster, S. S. & He, S. Growth–defense trade-offs in plants. *Curr. Biol.* **32**, R634–R639 (2022).
11. Herms, D. A. & Mattson, W. J. The Dilemma of Plants: To Grow or Defend. *Q. Rev. Biol.* **67**, 283–335 (1992).
12. Gao, M., Hao, Z., Ning, Y. & He, Z. Revisiting growth-defence trade-offs and breeding strategies in crops. *Plant Biotechnol. J.* **22**, 1198–1205 (2024).
13. Yıldırım, K. *et al.* Genome editing for healthy crops: traits, tools and impacts. *Front. Plant Sci.* **14**, 1231013 (2023).
14. Zandalinas, S. I. *et al.* Systemic signaling during abiotic stress combination in plants. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 13810–13820 (2020).
15. Bleker, C. *et al.* Stress Knowledge Map: A knowledge graph resource for systems biology analysis of plant stress responses. *bioRxiv* 2023.11.28.568332 (2023)

doi:10.1101/2023.11.28.568332.

16. Delplace, F., Huard-Chauveau, C., Berthomé, R. & Roby, D. Network organization of the plant immune system: from pathogen perception to robust defense induction. *Plant J.* **109**, 447–470 (2022).
17. Baebler, Š. *et al.* Dynamics of responses in compatible potato-Potato virus Y interaction are modulated by salicylic acid. *PLoS One* **6**, e29009 (2011).
18. Lukan, T. *et al.* Precision transcriptomics of viral foci reveals the spatial regulation of immune-signaling genes and identifies RBOHD as an important player in the incompatible interaction between potato virus Y and potato. *Plant J.* **104**, 645–661 (2020).
19. Töpfer, N. Environment-coupled models of leaf metabolism. *Biochem. Soc. Trans.* **49**, 119–129 (2021).
20. Nielsen, J. Systems Biology of Metabolism. *Annu. Rev. Biochem.* **86**, 245–275 (2017).
21. Rojas, C. M., Senthil-Kumar, M., Tzin, V. & Mysore, K. S. Regulation of primary plant metabolism during plant-pathogen interactions and its contribution to plant defense. *Front. Plant Sci.* **5**, 17 (2014).
22. Zaynab, M. *et al.* Role of secondary metabolites in plant defense against pathogens. *Microb. Pathog.* **124**, 198–202 (2018).
23. Panda, S., Kazachkova, Y. & Aharoni, A. Catch-22 in specialized metabolism: balancing defense and growth. *J. Exp. Bot.* **72**, 6027–6041 (2021).
24. Córdoba, S. C. *et al.* Identification of gene function based on models capturing natural variability of Arabidopsis thaliana lipid metabolism. *Nat. Commun.* **14**, 4897 (2023).
25. Arnold, A. & Nikoloski, Z. Bottom-up Metabolic Reconstruction of Arabidopsis and Its Application to Determining the Metabolic Costs of Enzyme Production. *Plant Physiol.* **165**, 1380–1391 (2014).
26. Hashemi, S., Laitinen, R. & Nikoloski, Z. Models and molecular mechanisms for trade-offs in the context of metabolism. *Authorea Preprints* (2022) doi:10.22541/au.166019951.10377598/v1.
27. Gerlin, L., Cottret, L., Escourrou, A., Genin, S. & Baroukh, C. A multi-organ metabolic model of tomato predicts plant responses to nutritional and genetic perturbations. *Plant Physiol.* **188**, 1709–1723 (2022).
28. Yuan, H., Cheung, C. Y. M., Poolman, M. G., Hilbers, P. A. J. & van Riel, N. A. W. A

genome-scale metabolic network reconstruction of tomato (Solanum lycopersicum L.) and its application to photorespiratory metabolism. *Plant J.* **85**, 289–304 (2016).

29. Botero, K., Restrepo, S. & Pinzón, A. A genome-scale metabolic model of potato late blight suggests a photosynthesis suppression mechanism. *BMC Genomics* **19**, 863 (2018).
30. Tomato Genome Consortium. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485**, 635–641 (2012).
31. Hawkins, C. *et al.* Plant Metabolic Network 15: A resource of genome-wide metabolism databases for 126 plants and algae. *J. Integr. Plant Biol.* **63**, 1888–1905 (2021).
32. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic Acids Res.* **48**, D445–D453 (2020).
33. Schauer, N., Zamir, D. & Fernie, A. R. Metabolic profiling of leaves and fruit of wild species tomato: a survey of the Solanum lycopersicum complex. *J. Exp. Bot.* **56**, 297–307 (2005).
34. Pęksa, A., Miedzianka, J., Nemś, A. & Rytel, E. The Free-Amino-Acid Content in Six Potatoes Cultivars through Storage. *Molecules* **26**, (2021).
35. Hastilestari, B. R. *et al.* Deciphering source and sink responses of potato plants (Solanum tuberosum L.) to elevated temperatures. *Plant Cell Environ.* **41**, 2600–2616 (2018).
36. Santacruz, S., Koch, K., Andersson, R. & Aman, P. Characterization of potato leaf starch. *J. Agric. Food Chem.* **52**, 1985–1989 (2004).
37. Chea, L., Meijide, A., Meinen, C., Pawelzik, E. & Naumann, M. Cultivar-Dependent Responses in Plant Growth, Leaf Physiology, Phosphorus Use Efficiency, and Tuber Quality of Potatoes Under Limited Phosphorus Availability Conditions. *Front. Plant Sci.* **12**, 723862 (2021).
38. Mengel, K. & Kirkby, E. A. *Principles of Plant Nutrition*. (Springer Science & Business Media, 2012).
39. Zagorščak, M. *et al.* Integration of multi-omics and deep phenotyping provides novel insights into multiple abiotic stress responses in potato. *bioRxiv* (2024) doi:10.1101/2024.07.18.604140.
40. Chan, S. H. J., Cai, J., Wang, L., Simons-Senftle, M. N. & Maranas, C. D. Standardizing biomass reactions and ensuring complete mass balance in genome-scale metabolic models. *Bioinformatics* **33**, 3603–3609 (2017).
41. Isah, T. Stress and defense responses in plant secondary metabolites production. *Biol. Res.* **52**, 39 (2019).

42. Karp, P. D. *et al.* Pathway Tools version 23.0 update: software for pathway/genome informatics and systems biology. *Brief. Bioinform.* **22**, 109–126 (2021).
43. Burgard, A. P., Nikolaev, E. V., Schilling, C. H. & Maranas, C. D. Flux coupling analysis of genome-scale metabolic network reconstructions. *Genome Res.* **14**, 301–312 (2004).
44. Gudmundsson, S. & Thiele, I. Computationally efficient flux variability analysis. *BMC Bioinformatics* **11**, 489 (2010).
45. Pereira, A. G. *et al.* Plant Alkaloids: Production, Extraction, and Potential Therapeutic Properties. in *Natural Secondary Metabolites: From Nature, Through Science, to Industry* (eds. Carocho, M., Heleno, S. A. & Barros, L.) 157–200 (Springer International Publishing, Cham, 2023).
46. Monson, R. K., Trowbridge, A. M., Lindroth, R. L. & Lerdau, M. T. Coordinated resource allocation to plant growth-defense tradeoffs. *New Phytol.* **233**, 1051–1066 (2022).
47. Schoville, S. D. *et al.* A model species for agricultural pest genomics: the genome of the Colorado potato beetle, Leptinotarsa decemlineata (Coleoptera: Chrysomelidae). *Sci. Rep.* **8**, 1931 (2018).
48. Baebler, Š., Coll, A. & Gruden, K. Plant Molecular Responses to Potato Virus Y: A Continuum of Outcomes from Sensitivity and Tolerance to Resistance. *Viruses* **12**, (2020).
49. Schwacke, R. *et al.* MapMan4: A Refined Protein Classification and Annotation Framework Applicable to Multi-Omics Data Analysis. *Mol. Plant* **12**, 879–892 (2019).
50. Pieterse, C. M. J., Van der Does, D., Zamioudis, C., Leon-Reyes, A. & Van Wees, S. C. M. Hormonal modulation of plant immunity. *Annu. Rev. Cell Dev. Biol.* **28**, 489–521 (2012).
51. Milner, S. E. *et al.* Bioactivities of glycoalkaloids and their aglycones from Solanum species. *J. Agric. Food Chem.* **59**, 3454–3484 (2011).
52. Dixon, R. A. *et al.* The phenylpropanoid pathway and plant defence-a genomics perspective. *Mol. Plant Pathol.* **3**, 371–390 (2002).
53. Kogovšek, P. *et al.* Primary Metabolism, Phenylpropanoids and Antioxidant Pathways Are Regulated in Potato as a Response to Potato virus Y Infection. *PLoS One* **11**, e0146135 (2016).
54. Boncan, D. A. T. *et al.* Terpenes and Terpenoids in Plants: Interactions with Environment and Insects. *Int. J. Mol. Sci.* **21**, (2020).
55. Iqbal, Z., Iqbal, M. S., Hashem, A., Abd Allah, E. F. & Ansari, M. I. Plant Defense Responses to Biotic Stress and Its Interplay With Fluctuating Dark/Light Conditions. *Front. Plant Sci.* **12**,

631810 (2021).

56. Ramsak, Ž. *et al.* GoMapMan: integration, consolidation and visualization of plant gene annotations within the MapMan ontology. *Nucleic Acids Res.* **42**, D1167–75 (2014).
57. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).
58. Van Bel, M. *et al.* PLAZA 5.0: extending the scope and power of comparative and functional genomics in plants. *Nucleic Acids Res.* **50**, D1468–D1474 (2022).
59. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
60. Moreno-Hagelsieb, G. & Latimer, K. Choosing BLAST options for better detection of orthologs as reciprocal best hits. *Bioinformatics* **24**, 319–324 (2008).
61. Cipollini, D., Walters, D. & Voelckel, C. Costs of resistance in plants: From theory to evidence. *Annual Plant Reviews online* 263–307 (2017) doi:10.1002/9781119312994.apr0512.
62. Coley, P. D., Bryant, J. P. & Chapin, F. S., 3rd. Resource availability and plant antiherbivore defense. *Science* **230**, 895–899 (1985).
63. Reed, J. L. Shrinking the metabolic solution space using experimental datasets. *PLoS Comput. Biol.* **8**, e1002662 (2012).
64. Machado, D. & Herrgård, M. Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. *PLoS Comput. Biol.* **10**, e1003580 (2014).
65. Kurotani, K.-I. *et al.* Genome Sequence and Analysis of Nicotiana benthamiana, the Model Plant for Interactions between Organisms. *Plant Cell Physiol.* **64**, 248–257 (2023).
66. Qin, C. *et al.* Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 5135–5140 (2014).
67. Shaw, R. & Cheung, C. Y. M. A Dynamic Multi-Tissue Flux Balance Model Captures Carbon and Nitrogen Metabolism and Optimal Resource Partitioning During Arabidopsis Growth. *Front. Plant Sci.* **9**, 884 (2018).
68. Schäfer, M. *et al.* Metabolic interaction models recapitulate leaf microbiota ecology. *Science* **381**, eadf5121 (2023).

69. Mataigne, V., Vannier, N., Vandenkoornhuyse, P. & Hacquard, S. Multi-genome metabolic modeling predicts functional inter-dependencies in the Arabidopsis root microbiome. *Microbiome* **10**, 217 (2022).

70. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

71. Heller, S. R., McNaught, A., Pletnev, I., Stein, S. & Tchekhovskoi, D. InChI, the IUPAC International Chemical Identifier. *J. Cheminform.* **7**, 23 (2015).

72. Moretti, S., Tran, V. D. T., Mehl, F., Ibberson, M. & Pagni, M. MetaNetX/MNXref: unified namespace for metabolites and biochemical reactions in the context of metabolic models. *Nucleic Acids Res.* **49**, D570–D574 (2021).

73. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

74. Schellenberger, J. *et al.* Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.* **6**, 1290–1307 (2011).

75. Heirendt, L. *et al.* Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat. Protoc.* **14**, 639–702 (2019).

76. Karp, P. D. *et al.* Pathway Tools version 24.0: Integrated software for Pathway/genome informatics and systems biology. *arXiv [q-bio.GN]* (2015).

77. Nishizaki, Y. *et al.* p-Hydroxybenzoyl-glucose is a zwitter donor for the biosynthesis of 7-polyacylated anthocyanin in Delphinium. *Plant Cell* **25**, 4150–4165 (2013).

78. Baur, S. *et al.* Biosynthesis of α-solanine and α-chaconine in potato leaves (Solanum tuberosum L.) - A 13CO2 study. *Food Chem.* **365**, 130461 (2021).

79. Cárdenas, P. D. *et al.* The bitter side of the nightshades: Genomics drives discovery in Solanaceae steroidal alkaloid metabolism. *Phytochemistry* **113**, 24–32 (2015).

80. Kohnen-Johannsen, K. L. & Kayser, O. Tropane alkaloids: Chemistry, pharmacology, biosynthesis and production. *Molecules* **24**, 796 (2019).

81. Hu, S. *et al.* Rerouting of the lignin biosynthetic pathway by inhibition of cytosolic shikimate recycling in transgenic hybrid aspen. *Plant J.* **110**, 358–376 (2022).

82. Liu, W. & Park, S.-W. 12-oxo-Phytodienoic acid: A fuse and/or switch of plant growth and defense responses? *Front. Plant Sci.* **12**, 724079 (2021).

83. Li, M., Yu, G., Cao, C. & Liu, P. Metabolism, signaling, and transport of jasmonates. *Plant Commun.* **2**, 100231 (2021).

84. Peng, Y., Yang, J., Li, X. & Zhang, Y. Salicylic acid: Biosynthesis and signaling. *Annu. Rev. Plant Biol.* **72**, 761–791 (2021).

85. Finkelstein, R. Abscisic Acid synthesis and response. *Arabidopsis Book* **11**, e0166 (2013).

86. Thiele, I. & Palsson, B. Ø. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* **5**, 93–121 (2010).

87. Ebrahim, A., Lerman, J. A., Palsson, B. O. & Hyduke, D. R. COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.* **7**, 74 (2013).

88. Orth, J. D., Thiele, I. & Palsson, B. Ø. What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248 (2010).

89. Potato Genome Sequencing Consortium *et al.* Genome sequence and analysis of the tuber crop potato. *Nature* **475**, 189–195 (2011).

90. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

91. Oksanen, J. *et al.* Vegan: Community ecology package. *CRAN: Contributed Packages* The R Foundation https://doi.org/10.32614/cran.package.vegan (2001).

92. Quensen, J., Simpson, G. & Oksanen, J. ggordiplots: Make ‘ggplot2’ Versions of Vegan’s Ordiplots. *CRAN: Contributed Packages* The R Foundation https://doi.org/10.32614/cran.package.ggordiplots (2023).

93. Kolde, R. pheatmap: Pretty Heatmaps. *CRAN: Contributed Packages* The R Foundation https://doi.org/10.32614/cran.package.pheatmap (2010).

94. Cheng, C.-Y. *et al.* Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *Plant J.* **89**, 789–804 (2017).

95. Tuskan, G. A. *et al.* The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). *Science* **313**, 1596–1604 (2006).

96. Loo, M. J. vander. The stringdist Package for Approximate String Matching. *R J.* **6**, 111 (2014).

97. Bleker, C., Grady, S. K. & Langston, M. A. A comparative study of gene co-expression thresholding algorithms. *J. Comput. Biol.* **31**, 539–548 (2024).

98. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in

large networks. *arXiv [physics.soc-ph]* (2008).

99. Enright, A. J., Van Dongen, S. & Ouzounis, C. A. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**, 1575–1584 (2002).
100. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
101. Petek, M. *et al.* Cultivar-specific transcriptome and pan-transcriptome reconstruction of tetraploid potato. *Sci Data* **7**, 249 (2020).