**Metabolic changes in leaves of *N. tabacum* and *N. benthamiana* during plant development**

Margit Drapal, Eugenia M. A. Enfissi and Paul D. Fraser\*

Department of Biological Sciences, Royal Holloway University of London, Egham, United Kingdom

\*Corresponding author: P.Fraser@rhul.ac.uk, Tel: +44 1784 443555

# Abstract

Dwindling fossil fuel reserves and poor environmental credentials of chemical synthesis means, new renewable sources for the production and manufacture of valuable chemicals and pharmaceuticals are required. Presently, tobacco is an underutilised non-food crop with the potential to act as a biofactory. In this study, metabolite profiling across vegetative development has been carried out to provide a quantitative baseline of metabolites, their formation and interaction. Two tobacco platforms have been used, *Nicotiana benthamiana* and *Nicotiana tabacum*. Our data generated has provided the quantitative and qualitative baseline levels for exploitable pathways and metabolites, across two complementary *Nicotiana* species. *N. benthamiana* is the chassis of choice for transient expression. The metabolite data obtained for *N. benthamiana* highlighted that before flower emergence, the increased central carbon metabolism and high amino acid levels are available for the biosynthesis of endogenous or heterologous metabolites. In the future, engineering pathways or biocatalysts into *N. benthamiana* could add value to the process presently used to produce low volume, high cost pharmaceuticals. Similar outputs were obtained for *N. tabacum,* which has the advantage of providing a large biomass and hence, high product yield. These data provide an insight into the metabolite pools available in tobacco for future exploitation by emerging New Plant Breeding Techniques.

# Keywords

*N. tabacum*; *N. benthamiana*; metabolite analysis; plant development

# Abbreviations

ANOVA, analysis of variance; DW, dry weight; GABA, γ-amino butyric acid; GC-MS, gas chromatography mass spectrometry; LC-MS, liquid chromatography mass spectrometry; PC, principle component; PCA, principle component analysis; TCA, tricarboxylic acid; UPLC-ESI-QTof; DPG, days post germination.

# Introduction

Plant molecular farming utilises plants to produce a wide array of valuable small molecules as well as pharmaceutically important compounds such as antibodies and recombinant proteins. The use of plants as biofactories has several advantages such as sustainability, safety and scalability. A large number of high-value chemicals are produced from petrochemical-derived precursors and use rare metal catalysts in the process; ketocarotenoids for aquaculture or building blocks for pharmaceutical compounds are good examples (Nogueira et al., 2017; Tramontina et al., 2020). The biosynthesis of these chemicals in plant chassis can be easily up-scaled to open-field cultivation in a sustainable and low-cost manner. Another advantage of plants as biofactories is the absence of human pathogens compared to other production systems e.g. mammalian cell cultures (Molina-Hidalgo et al., 2020).

*Nicotiana* species, in particular *N. benthamiana* and *N. tabacum*, are ‘non-food’ crops amenable to transformation. Previous studies showed successful application of both species for production of artemisinic acid, α-tocopherol, influenza (H1N1 and H5N1) vaccine, anti-SARS-CoV-2 antibodies and ZMapp, an anti-Ebola antibody cocktail (Chen and Davis, 2016; Diego-Martin et al., 2020; Fuentes et al., 2016; Landry et al., 2014; Sathish et al., 2018). The phenotypes of the two *Nicotiana* species differ greatly in leaf shape, stem branching and flowers. *N. benthamiana* comprises ovate, petiolate leaves, branches in axillary buds and solitary flowers in axils or internodes compared to approx. ten times larger, sessile leaves along one main stem with dense panicles at the terminal bud of *N. tabacum* (Iizuka et al., 2012). Furthermore, the *N. benthamiana* LAB strain has a mutation in the RNA-dependent RNA polymerase gene (*Rdr1*), which causes hyper-susceptibility to infection by viruses and bacteria and therefore is an ideal host for transient expression with agro-infection (Bally et al., 2018; Wylie et al., 2015). However, the large biomass of *N. tabacum* is of great advantage for stable transformation and increased product yield. Despite these phenotypic difference, these two *Nicotiana* species share metabolic pathways, which have been studied since 1800 (Rodgman and Perfetti, 2013).

Since the development of metabolomics techniques, different platforms have been developed and detection has been improved for a more comprehensive coverage of the metabolome (Pinu et al., 2019). These techniques can be used to elucidate chemotypes resulting from gene modification, stress responses or answer more fundamental questions such as metabolite profiles of senescing leaves and leaves at different stalk positions (Li et al., 2016; Zhang et al., 2018). Approaches similar to the latter two studies enable a better understanding of the metabolic baseline for plant modifications. They provide information of the relationship between central carbon metabolism and end products and how modifications or additions of pathways change this relationship (Enfissi et al., 2021; Fraser et al., 2020). Additionally, the metabolite data can indicate which stage or age of plant development is more suitable for the harvest of specific metabolites or presents the most precursors for introducing heterologous pathways by transformation.

Hence, the present study was focused on the metabolic changes occurring in *N. benthamiana* and *N. tabacum* during plant development from juvenile plants to the emergence of seed pods. To avoid measuring metabolites associated with leaf expansion and senescence, the most fully expanded leaves at the respective time points were harvested. Sampling commenced 29 days post germination (DPG) to guarantee survival and minimal stress after the first leaf harvest. The consecutive sampling was performed weekly until emergence of seed pods. The resulting data set showed great metabolic variability over the development of the plant, characterised by distinct metabolite profiles for certain developmental phases. *N. benthamiana* and *N. tabacum* followed expected metabolic adaptation of primary metabolism at the juvenile stages and secondary metabolism at the adult stages. Furthermore, content of nicotine, phenylpropanoid, central carbon metabolites and amino acids indicated that the plants in vegetative phases are more advantageous for both transient and stable transformation.

# Material and methods

## Plant cultivation and leaf sampling

*N. benthamiana* (LAB strain)and *N. tabacum* (variety K326) seeds were germinated on Levington® Advance Pot&Bedding Compost F2+Sand (ICL Specialty Fertilizers, UK). Six days post germination (DPG), plantlets reached a height of ~3cm with three to four leaves and six biological replicates per species were transferred to individual pots with Levington® Advance Pot&Bedding Compost M3 (ICL Specialty Fertilizers, UK). Plants were grown in a glasshouse at 24°C under supplementary lighting (16h light/8h dark cycle) and supplemented with Universol® Blue (ICL Specialty Fertilizers, UK) once a week. From 29 DPG, leaves were sampled every seven days until the emergence of flowers and one additional leaf sample was taken at the emergence of seed pots. The leaf samples comprised a fully expanded leaf of *N. benthamiana* and half of a fully expanded leaf of *N. tabacum*. Leaves were immediately frozen in liquid nitrogen and stored at -80°C until all plant stages were sampled. Frozen leaves were then lyophilised and ground to a fine powder with a TissueRuptor (Qiagen).

## Metabolite extraction and analysis

A portion of the ground powder (10-11mg) was weighed out and extracted with a methanol/water and chloroform separation method as previously described (Drapal et al., 2020b). For metabolite profiling by UPLC-ESI-QTof 6560 (Agilent, UK), an aliquot of the polar phase (100µL) was filtered and an internal standard genistein (5µL of a 0.2mg/ml stock) added. The samples (1µL injection) were analysed with a water/acetonitrile gradient as previously described without split analysis by DAD (Drapal et al., 2020b).

For metabolite profiling by 7890A GC on-line with a 5975C MSD (Agilent Technologies, Palo Alto, California, US), an aliquot of the polar (150µL) and non-polar (700µL) phase were dried down with internal standards d4-suiccinic acid (10µg) and d27-myristic acid (10µg), respectively. The dried samples were derivatised with methoxyamine hydrochloride and *N-*methyl-*N*-(trimethylsilyl)trifluoroacetamide and analysed in splitless mode as previously described (Drapal et al., 2019).

## Data analysis

Data processing was performed with AMDIS (v2.71, NIST) and an in-house library for GC-MS files and with Agilent Profinder (v10.0 SP1, Agilent Technologies, Inc.) for LC-MS files. The molecular features of LC-MS files were then compared to an in-house library based on retention time and mass spectrum. Database NIST11 (<http://chemdata.nist.gov/mass-spc/ms-search/>) was used for identification of GC-MS compounds not present in the in-house library. The data matrix was subjected to relative quantification with the respective internal standard and sample weight. PCA analysis was performed with Simca P (13.0.3.0, Umetrics), heatmaps and correlation analysis were produced with Metaboanalyst including pareto-scaling and ANOVA (Xia and Wishart, 2016), bar charts were created with GraphPad Prism (v.9.1.0, GraphPad Software, LLC) and pathway displays created with in-house software.

**Table 1. Description of leaf sample and of the plant at the time of sampling.** Descriptors listed include the height of the plant, the number of leaves present on the plant and the size of the most fully expanded leaf. The sampling times are listed as days post germination (DPG). Emergence of flowers and seed pots are highlighted for *N. benthamiana* (\* and \*\*, respectively) and *N. tabacum* (^ and ^^, respectively) in the DPG column. Number of leaves and plant height represents an average of six biological replicates.

|  |  |  |
| --- | --- | --- |
|  | ***N. benthamiana*** | ***N. tabacum*** |
| **Days post germination (DPG)** | **Time point** | **Fully ex-panded leaf** | **No. of leaves** | **Plant height** | **Time point** | **Fully ex-panded leaf** | **No. of leaves** | **Plant height** |
| 29 | 1 | 11.5cm | 16 | 8cm | 1 | 26cm | 4 | 6cm |
| 36 | 2 | 15cm | 20 | 15cm | 2 | 32cm | 7 | 12cm |
| 40\* | 3 | 15cm | 20 | 16cm |  | - | - | - |
| 43 | 4 | 15cm | 30 | 28cm | 3 | 32cm | 9 | 19cm |
| 50 | 5 | 11cm | 36 | 35cm | 4 | 30cm | 13 | 21cm |
| 57 | 6 | 7cm | 36 | 35cm | 5 | 30cm | 14 | 23cm |
| 64 |  | - | - | - | 6 | 30cm | 16 | 36cm |
| 71 |  | - | - | - | 7 | 30cm | 15 | 36cm |
| 78\*\* | 7 | 10cm | 38 | 39cm | 8 | 30cm | 18 | 40cm |
| 85^ |  | - | - | - | 9 | 30cm | 26 | 68cm |
| 92^^ |  | - | - | - | 10 | 30cm | 28 | 68cm |

# Results

## Metabolic variability ofleaves during plant development

*N. tabacum* and *N. benthamiana* were grown and leaf samples taken every seven days from the first month after germination, until the emergence of flowers and then sampled once more at the emergence of seed pods. This resulted in seven time points for *N. benthamiana* and ten time points for *N. tabacum*. Each sampling consisted of the most recent, fully expanded leaf for *N. tabacum* and a pool of the three most recent, fully expanded leaves for *N. benthamiana*. The polar extracts of these samples were subjected to metabolite profiling, which detected >4000 molecular features in the sample set. The PCA score plot of the metabolite profiling showed a consecutive grouping in order of the sampling time (Fig. 1). The exception was time point 2 for *N. tabacum*, which grouped with time point 6 (Fig. 1B). The grouping of the score plot would suggest that time points 1-2 in *N. benthamiana* and 1 and 3 in *N. tabacum* represent the initial/juvenile stage based on phenotypic data (Orlando et al., 2011). The other time points for *N. tabacum* correspond to “development” (time points 4-8), “med-season” (time point 9 – flower emergence) and “late season” (seed pod emergence). The highest biological variation was detected at the emergence of seed pods for both *Nicotiana* species. Statistical analysis (one-way ANOVA for time series) highlighted that 70% and 80% of metabolites were significantly different between all time points for *N. benthamiana* and *N. tabacum,* respectively. A heatmap display was chosen to visualise the difference in metabolite composition between the time points. For *N. benthamiana*, the heatmap showed opposite trends for time points 1-2 compared to 5-7, whereas time points 3-4 showed metabolite trends similar to both earlier and later time points (Fig. A.1). For *N. tabacum*, the heatmap highlighted four different metabolic compositions for time points 1 and 3, time points 4 and 5, time points 2 and 6 and time points 7-10 (Fig. A.2). Time points 2 and 6 seemed to be intermediary stages with trends similar to earlier and later time points. Both *Nicotiana* species display the highest levels of the majority of molecular features (~75%) at the last time point representing seed pod emergence.

## Adaptations of primary and secondary metabolism during plant development

Analysis of polar and non-polar extracts by GC-MS identified 143 metabolites including 52 class 3 and 4 unknowns (Halket et al., 2005). A further 80 metabolites were identified from the metabolite profiling data by LC-MS. The identified metabolites were grouped into primary and secondary metabolites. The PCA score plots and heatmaps of the respective data sets showed the same trends as observed for the metabolite profiling data. This suggests a sufficient representation of the identified metabolites for the whole leaf metabolism (Fig. 2).

Part of the present study was to measure metabolites which are known to interfere in the down-stream processing or are adverse medical products produced in *Nicotiana* species, e.g. phenylpropanoid derived compounds and nicotine. *N. benthamiana and N. tabacum* showed different trends in regard to nicotine content and total phenylpropanoid levels. Nicotine was increased after the first time point in both *Nicotiana* species and decreased in quantity until the last time point for *N. benthamiana*. *N. tabacum* maintained the nicotine levels for two more time points, increased the levels by ~3-fold at time point 5 and 6, followed by a reduction to 4mg/g DW which was slowly increased again to ~6mg/g DW at the last time point. The total phenylpropanoid levels in *N. benthamiana* did not change significantly over time, whereas *N. tabacum* showed and initial increase and then decrease over time. The first three time points showed similar phenylpropanoid levels, followed by two-fold higher levels at time point 4 and 5 and a decrease back to the original levels at time point 7. At time points 8 and 9 the levels increase again to the same level as time point 6 and decreased again to the original levels at the last time point. The changes of individual phenylpropanoid derived compounds and other metabolites detected will be discussed individually for *N. benthamiana* and *N. tabacum*.

### Metabolic alterations specific to *N. benthamiana*

The metabolite data of *N. benthamiana* was analysed with repeated one-way ANOVA, which showed 78% of the identified compounds were significantly different between the time points. The heatmap based on the statistical analysis displayed that the seven time points could be grouped into three distinguishable metabolite profiles (Fig. A.3). The groupings included time points 1 and 2, time points 3 and 4 and time points 5 to 7. However, time point 7, at the emergence of seed pods, showed metabolite levels significantly higher (on average 2.1-fold) than the previous two time points, 5 and 6, and was therefore regarded as a separate metabolite profile. The information from the heatmap was summarised as a pathway display highlighting at which developmental stage each metabolite was present with the highest levels (Fig. 3). Correlation analysis was incorporated to establish connections between single metabolites, metabolite classes and/or metabolite pathways (Fig. A.4). The most obvious observation from the metabolite data for *N. benthamiana* was that the metabolite profiles for the first and last time points showed opposite trends and primary and secondary metabolites were in general negatively correlated.

The first two time points (stage 1) showed the highest levels of most amino acids, sugars (e.g. glucose, fructose and sucrose), ascorbic acid, *trans*-caffeic acid, phytol, sterols, hydrocarbons and precursors for glycerolipids (e.g. glycerol-phosphate and linolenic acid). The levels of these metabolites were on average 1.5-fold higher for sterols and sugars, ~2-fold higher for glycerolipid precursors, ~2.5-fold higher for hydrocarbons and ~4-fold higher for amino acids. The highest fold difference was detected for alanine, which was 11-fold higher in stage 1 compared to the other developmental stages. The metabolites highest in stage 1 showed a positive correlation to each other and to even-chain fatty acids detected in the present study. Interestingly, the fatty acids, with the exception of C20:0 and linolenic acid, were not significantly different between the developmental stages. Ascorbic acid was only detected in the first three time points and decreased by half from stage 1 to stage 2. *trans*-Caffeic acid, the key intermediate for lignin biosynthesis, was 2.4-fold higher in stage 1 and showed a positive correlation with primary metabolites such as sugars and amino acids and a negative correlation with all phenylpropanoid derived compounds except chlorogenic acids and their glycosides. The latter two metabolite groups showed no significant correlation to any metabolites detected in the present study.

Stage 2, including time points 3 and 4, had the highest levels of intermediates of the TCA cycle (fumaric acid, malic acid and maleic acid) by ~2-fold, GABA (1.6-fold), amino acids aspartic acid, tryptophan and glutamic acid (~2.3-fold), glycerol (1.2-fold) and sedoheptulose and other unidentified sugars (~3-fold). Neither the intermediates of the TCA cycle nor GABA showed any significant correlation to other metabolites. All other metabolites significantly higher in stage 2 showed a positive correlation to metabolites of stage 1, with the exception of tryptophan, which was negatively correlated to primary metabolites and positively correlated to phenylpropanoid derived secondary metabolites.

Stage 3 was comprised of time points 5 and 6 and as mentioned before, showed a similar profile as stage 4 (time point 7). Hence, stage 3 showed only eight metabolites with higher levels compared to other stages, which is described as fold change for each metabolite. These metabolites included phenylpropanoid precursors shikimic acid (1.3-fold) and quinic acid (3-fold), quercetin glucosides (2.2-fold), pregnane (1.8-fold) and sugar acids erythronic (1.9-fold) and glucuronic acid (1.1-fold). All metabolites of stage 3 were positively correlated to phenylpropanoid derived compounds.

The last time point (stage 4) had the highest levels of the majority of phenylpropanoids including chlorogenic acids, coumaroyl- and feruloyl-quinates, hexose and shikimic acid esters of coumaric acid and kaempferol- and quercetin-glycosides, as well as C20:0 and glycerolipids of C18:0 and C20:0, citric acid, *myo*-inositol and six unidentified sugars, β-sitosterol and hydrocarbons (e.g. hentriacontane and docosanol). The phenylpropanoids were on average ~4.5-fold higher than stage 1 and 2 and ~1.7-fold higher than stage 3. Fatty acids and *myo-*inositol were ~1.2-fold higher than all other stages, whereas glycerol ester of C20:0 and C18:0 were ~10-fold higher compared to stage 1 and 2 and ~2.5-fold higher than stage 3. *myo*-Inositol showed the opposite trend to all other lipid precursors with a positive correlation to phenylpropanoids and a negative correlation to intermediates of glycolysis, pentose phosphate pathway, amino acids and all other precursors for glycerolipids. Citric acid was the only intermediate of the TCA cycle, which showed a positive correlation to phenylpropanoid derived compounds, and was present 3.5-fold higher in stage 4 compared to the earlier developmental stages.

Nicotine showed a positive correlation to intermediates of the TCA cycle, amino acids of the nitrogen shunt (GABA and putrescine) and citramalic acid and a negative correlation to sugars and amino acids such as glycine, threonine, serine, alanine and valine.

### Metabolic alterations specific to *N. tabacum*

One-way repeated ANOVA showed 85% identified metabolites were significantly different between the time points of *N. tabacum*. The heatmap based on this statistical analysis highlighted five distinguishable metabolite profiles (Fig. A.5). Time points 1 and 3 were grouped together, time points 4 and 5, time points 2 and 6 and time points 7 to 9. As observed for *N. benthamiana*, (i) the last time point showed a metabolite profile similar to the earlier time points 7 to 9 with significantly higher (1.7-fold) metabolite levels, (ii) metabolite levels showed an opposite trend between earlier and later time points and (iii) primary metabolites including precursors for glycerolipids were negatively correlated to phenylpropanoid metabolism (Fig. A.6). The highest levels of metabolites at each developmental stage were summarised in a pathway display (Fig. 4).

Stage 1 (time points 1 and 3) had the highest levels of sugars (e.g. glucose, fructose and sucrose), glycerolipid precursors glycerol and -phosphate, C16:0, C18:0 and C18:1, hydrocarbons (nonacosane and hexadecene), campesterol and stigmasterol, one phytol isomer, intermediates of the TCA cycle (from α-ketoglutaric acid to malic acid), amino acids glutamic acid, aspartic acid, putrescine, valine, alanine, serine and glycine and almost all chlorogenic acid glycosides. These metabolites were on average ~1.3-fold for sugars, ~3-fold for intermediates of the TCA cycle, >2-fold for amino acids, ~2-fold for glycerolipid precursors, hydrocarbons and sterols and ~4.7-fold for chlorogenic glycosides higher compared to the other developmental stages. All of these metabolites were positively correlated and showed a positive correlation to intermediates of the pentose phosphate pathway and α-tocopherol.

Time points 4 and 5 (stage 2) had the highest levels of amino acids threonine, proline and glutamine, three unidentified sugars and trehalose (~2-fold), α-tocopherol (~2.8-fold), *trans*-caffeic acid and two shikimate esters, chlorogenic acids, one chlorogenic acid glycoside and two di-caffeoyl quinates (>3-fold), two coumaroyl quinates, one kaempferol-glucoside, quercetin glucoside and rutin (~1.8-fold). The sugars at stage 2 showed no correlation and α-tocopherol showed a negative correlation to the phenylpropanoid derived compounds. *trans*-Caffeic acid showed almost no correlation with the exception of a positive correlation to chlorogenic acids.

Time points 2 and 6 (stage 3) had the highest levels of shikimic acid (~2.1-fold), quinic acid esters of coumaric acid and ferulic acid, kaempferol-rutinosides, one sinapoyl malate and di caffeoyl quinate, one phytol isomer, α-linolenic acid, cadaverine, xylose and erythronic acid, glyceric acid and glycerol-glycoside. The phenylpropanoid derived compounds and glycerolipid precursors were ~1.5-fold higher at this stage compared to the other time points and showed a negative correlation as described for the previous two stages.

Similar to *N. benthamiana*, the time points most similar to seed pod emergence had the least amount of highest metabolites. In *N. tabacum,* these time points were 7 to 9 (stage 4) and had the highest levels of itaconic and citric acid (1.3 and 2.1-fold), two coumaroyl shikimates (~2.3-fold), ferulic acid (~1.4-fold) and two unidentified sugars (>2-fold).

The last time point at the emergence of seed pods (stage 5) had the highest levels of amino acids phenylalanine, tryptophan and GABA, myo-inositol, fatty acids C14:0 and C17:0, cholesterol, β-sitosterol, prostaglandins, unidentified steroids, most of the detected hydrocarbons, malic acid esters of ferulic acid, sinapic acid and quercetin hexose, shikimic acid esters of *trans*-caffeic acid and coumaric acid, coumaroyl hexosides, di caffeoyl quinates and quercetin malonyl glucosides. Phenylpropanoids were ~2.5-fold higher than the other stages and the other compounds were ~1.4-fold (*myo*-inositol, phenylalanine and sterols) to 5.5-fold (tryptophan) higher. The phenylpropanoid derived compounds were positively correlated to myo-inositol, GABA and shikimic acid and negatively correlated to all other primary metabolites detected.

Several unidentified compounds of the non-polar extract showed similarities to the backbone structure of cembranoids. These compounds were present in the highest amounts at stage 5 and were positively correlated with sterols, fatty acids and their glycerol esters and a few phenylpropanoids (e.g. quinic and ferulic acid) and negatively correlated with intermediates of glycolysis and TCA cycle.

Contrary to *N. benthamiana*, nicotine showed no correlation to the TCA cycle or nitrogen shunt in *N. tabacum*. However, it showed a positive correlation to cadaverine, citramalic acid, sucrose, xylose and phenylpropanoids except chlorogenic acids and a negative correlation to fructose, fatty acids and their glycerol esters and amino acids glycine, valine, alanine.

# Discussion

## Plant development of *N. tabacum*

The present study highlighted distinct metabolite profiles before, at and after flowering for both *N. benthamiana* and *N. tabacum*. For *N. tabacum*, these profiles followed the physical stages (initial, crop development, pre-flowering and flowering) described by Orlando et al. (2011). Primary metabolites were more prevalent at initial stages and crop development I. However, the initial stages had higher levels of amino acids, the main sugars sucrose, fructose and glucose and intermediates of the TCA cycle. This was consistent with previous findings in tobacco,banana and yacόn (Drapal et al., 2020a; Padilla-González et al., 2019; Zhang et al., 2018). In general, juvenile plants significantly increase their photosynthetic activity for energy conversion and their ammonium assimilation for amino acid biosynthesis to support cellular processes necessary for plant growth (Huijser and Schmid, 2011; Zheng, 2009).

In the next growth stage, crop development I, a significant increase of phenylpropanoid derived compounds for lignin synthesis was observed. This corresponds to the increase of stem height and leaf number, two developmental processes which require lignification (Barros et al., 2015; Laskar et al., 2010; Schlimme et al., 2002). The stem height increase of *N. tabacum* at time point 6 indicated the new growth stage crop development II, which showed metabolic traits of both the initial stage and crop development I. The combined assimilation of primary metabolites (glycolysis, pentose phosphate pathway and TCA cycle) and precursors for lignin synthesis suggests that the plant had enough leaf surface to support both processes simultaneously.

The biggest metabolic change could be observed between the vegetative stage (time point 1-8) and the reproductive stages (time points 9-10). Based on the present metabolite data, *N. tabacum* showed a significant increase of the specialised pathways for flavonoids, hydroxy cinnamic acids, terpenoids, hydrocarbons and alkaloids. These pathways provide precursors for lignification and flower colouration, components for the leaf surface gum, anti-oxidants and protection against herbivores (Heemann et al., 1983; Nishihara et al., 2005; Nugroho and Verpoorte, 2002; Song et al., 2020; Zenkner et al., 2019). The positive correlation of nicotine to phenylpropanoids, hydrocarbons and potential cembranoids, which in turn were negatively correlated to the TCA cycle, suggests that in *N. tabacum* several specialised pathways compete for the same precursors. The changes of nicotine and total phenylpropanoid levels (Fig. 2) support this hypothesis, e.g. phenylpropanoids remain unchanged from time point 1 to 2 whereas nicotine content increased, followed by the reverse trend at time point 3. Interestingly, at time point 7 nicotine and total phenylpropanoid levels as well as plant height and leaf number remained unchanged despite a significant change of the metabolite profile. These changes included an increase of sugar acids and dibasic acids, suggesting an induction of the pentose phosphate pathway and the N shunt. Furthermore, a simultaneous decrease of amino acids indicates protein biosynthesis for the initiation of reproductive structures, which could not be observed for another fourteen days (Pajoro et al., 2014).

## Plant development of *N. benthamiana*

The phenotypic differences between *N. benthamiana* and *N. tabacum*, in particular the early flowering, complicate a direct comparison of growth stages. However, the present metabolite data highlighted similarities between the clearly definable developmental stages: vegetative and reproductive phase and emergence of seed pods. The metabolite profiles of these time points follow the same trends detected in *N. tabacum* with primary metabolites more prevalent in the younger plants and secondary metabolites more prevalent in the adult plants. The pentose phosphate pathway and TCA cycle seemed to be important throughout the plant development of *N. benthamiana* as individual intermediates of both pathways were present at highest level in all four stages and the majority of TCA cycle intermediates showed no correlation to other metabolites. The exception was citric acid which showed a positive correlation to phenylpropanoids and was significantly increased at the last growth stage. This trend could be related to antioxidant protection of citric acid for phenylpropanoids, also present at the highest levels at the last stage. Citric acid derives its antioxidant properties through lowering the pH as well as chelating proteins such as polyphenol oxidases, which oxidise phenolic compounds. The increase of citric acid would therefore prevent polyphenol oxidases from reducing the content of lignin precursors and at the same time prevent production of reactive quinones (Mallhi et al., 2019; Moon et al., 2020).

The last growth stage of *N. benthamiana* also showed more prevalent levels of *myo*-inositol and glycerol ester of fatty acids, which are components of phosphatidylinositol lipids. Previous studies have highlighted the importance of this minor lipid group for normal plant growth including signalling functions and pollen development (Heilmann, 2020; Lee et al., 2008). Physiological functions of phosphatidylinositol such as vesicle and membrane trafficking and cytoskeletal remodelling could be of great relevance for the correct development of reproductive organs flowers and seed pods (Bennett et al., 2011; Sousa et al., 2008). *N. tabacum* only showed an increase of *myo*-inositol at the last stage, which suggests a different lipid or signalling regulation between the *Nicotiana* species.

## Correlation of primary metabolism and nicotine

The nicotine production in *N. tabacum* and *N. benthamiana* seems to be differently regulated throughout plant development based on the nicotine content measured (Fig. 2E). The correlation of nicotine to primary metabolites supports this hypothesis further. As discussed earlier, nicotine in *N. tabacum* showed little correlation to the TCA cycle and N shunt due to several competing secondary pathways. However, the positive and negative correlation of nicotine to sucrose and fructose, respectively, indicates the importance of glycolysis for the production of nicotine. In *N. benthamiana* the correlation of nicotine to its precursors was displayed as a clear positive correlation to the TCA cycle and nitrogen shunt and a negative correlation to glycolysis and several amino acids. This suggests that the TCA cycle and nitrogen shunt were fuelled by the catabolism of sugars and amino acids to provide a carbon and nitrogen source for the production of nicotine without compromising the pool of TCA cycle intermediates. Interestingly, both *Nicotiana* species showed a negative correlation of nicotine to specific amino acids: alanine, glycine and valine. Previous studies have elucidated that both alanine and glycine are precursors for nicotine as part of the glutamic acid cycle and α-carbon donor for the *N*-methyl group of nicotine, respectively (Byerrum et al., 1954; Xu et al., 2017). Literature highlighted that the negative correlation to valine was based on the function of the two metabolites. Valine catabolism is activated at the reproductive stage as it is important for normal seed development (Gipson et al., 2017). At this same developmental stage, the biosynthesis of the toxic defence compound nicotine is increased to protect the reproductive organs from herbivores as seen through the significant increase of nicotine after “topping” (Zenkner et al., 2019). Based on the present metabolites data, both *Nicotiana* species have similar pathways for nicotine production. The consistent nicotine levels in *N. benthamiana*, after flower emergence, would indicate that nicotine plays a less important role in the defence of reproductive organs compared to *N. tabacum*. The reduction of alkaloids in *N. tabacum* would present more intermediates of the central carbon metabolism for other pathways, which is of interest for the use of tobacco for biofactories. However, the reduction of the natural defence mechanism might lead to crop loss in the field due to increased herbivore feeding.

# Conclusions

The present study elucidated the chemotypes of *N. benthamiana* and *N. tabacum* at different developmental stages. In our quest to develop new renewable sources of chemicals and pharmaceuticals, the data provides important quantitative and qualitative baseline levels for exploitable pathways and metabolites, across two complementary *Nicotiana* species. Presently, the established transient expression in *N. benthamiana* is widely used for the production of therapeutics. The metabolite data highlighted that before flower emergence, increased central carbon metabolism and high amino acid levels are available for the biosynthesis of endogenous or heterologous metabolites. Engineering pathways or biocatalysts into *N. benthamiana* that could utilise these precursor pools has the potential to add value to the process presently used to produce low volume, high cost pharmaceuticals. The data also showed that the total phenylpropanoid levels do not change significantly over time and thus are unlikely to adversely affect down-stream processing. For stable transformation, the leaves of fully grown *N. tabacum* provide a large biomass and hence, high product yield. Similar to *N. benthamiana*, the leaves of *N. tabacum* should be harvested before transition to the reproductive phase to avoid high levels of nicotine and phenylpropanoids. The process of “topping” might revert the metabolism back to the vegetative state, but bears the disadvantage of precursors being diverted to the production of defence metabolites instead of the desired engineered pathway. These data can be utilised to ensure optimal outputs and exploitation of the *N. benthamiana* and *N. tabacum* platforms by emerging New Plant Breeding Techniques.

# Acknowledgments

The authors thank Chris Gerrish for excellent technical assistance.

# Funding

This work was funded by European Union Funding for Research and Innovation “Horizon 2020 - NEWCOTIANA – Project no. 760331”.

# Supplementary data

All processed data sets are available as appendices. Unprocessed data can be accessed at DOI: 10.17632/76r6j4n2y5.1.

# References

Bally, J., Jung, H., Mortimer, C., Naim, F., Philips, J.G., Hellens, R., Bombarely, A., Goodin, M.M., Waterhouse, P.M., 2018. The Rise and Rise of Nicotiana benthamiana: A Plant for All Reasons. Annu. Rev. Phytopathol. 56, 405–426. https://doi.org/10.1146/annurev-phyto-080417-050141

Barros, J., Serk, H., Granlund, I., Pesquet, E., 2015. The cell biology of lignification in higher plants. Ann. Bot. 115, 1053–1074. https://doi.org/10.1093/aob/mcv046

Bennett, E.J., Roberts, J.A., Wagstaff, C., 2011. The role of the pod in seed development: strategies for manipulating yield. New Phytol. 190, 838–853. https://doi.org/https://doi.org/10.1111/j.1469-8137.2011.03714.x

Byerrum, R.U., Hamill, R.L., Ball, C.D., 1954. THE INCORPORATION OF GLYCINE INTO NICOTINE IN TOBACCO PLANT METABOLISM. J. Biol. Chem. 210, 645–650. https://doi.org/https://doi.org/10.1016/S0021-9258(18)65390-6

Chen, Q., Davis, K., 2016. The potential of plants as a system for the development and production of human biologics. F1000Research 5, 912. https://doi.org/10.12688/f1000research.8010.1

Diego-Martin, B., González, B., Vazquez-Vilar, M., Selma, S., Mateos-Fernández, R., Gianoglio, S., Fernández-del-Carmen, A., Orzáez, D., 2020. Pilot Production of SARS-CoV-2 Related Proteins in Plants: A Proof of Concept for Rapid Repurposing of Indoor Farms Into Biomanufacturing Facilities. Front. Plant Sci. 11, 2101. https://doi.org/10.3389/fpls.2020.612781

Drapal, M., Amah, D., Schöny, H., Brown, A., Swennen, R., Fraser, P.D., 2020a. Assessment of metabolic variability and diversity present in leaf, peel and pulp tissue of diploid and triploid Musa spp. Phytochemistry 176, 112388. https://doi.org/https://doi.org/10.1016/j.phytochem.2020.112388

Drapal, M., Barros De Carvalho, E., Ovalle Rivera, T.M., Becerra Lopez-Lavalle, L.A., Fraser, P.D., 2019. Capturing Biochemical Diversity in Cassava (Manihot esculenta Crantz) through the Application of Metabolite Profiling. J. Agric. Food Chem. 67. https://doi.org/10.1021/acs.jafc.8b04769

Drapal, M., Ovalle Rivera, T.M., Becerra Lopez-Lavalle, L.A., Fraser, P.D., 2020b. Exploring the chemotypes underlying important agronomic and consumer traits in cassava (Manihot esculenta crantz). J. Plant Physiol. 251, 153206. https://doi.org/10.1016/j.jplph.2020.153206

Enfissi, E.M.A., Drapal, M., Perez-Fons, L., Nogueira, M., Berry, H.M., Almeida, J., Fraser, P.D., 2021. New plant breeding techniques and their regulatory implications: An opportunity to advance metabolomics approaches. J. Plant Physiol. 258–259, 153378. https://doi.org/10.1016/j.jplph.2021.153378

Fraser, P.D., Aharoni, A., Hall, R.D., Huang, S., Giovannoni, J.J., Sonnewald, U., Fernie, A.R., 2020. Metabolomics should be deployed in the identification and characterization of gene-edited crops. Plant J. 102, 897–902. https://doi.org/https://doi.org/10.1111/tpj.14679

Fuentes, P., Zhou, F., Erban, A., Karcher, D., Kopka, J., Bock, R., 2016. A new synthetic biology approach allows transfer of an entire metabolic pathway from a medicinal plant to a biomass crop. Elife 5. https://doi.org/10.7554/eLife.13664

Gipson, A.B., Morton, K.J., Rhee, R.J., Simo, S., Clayton, J.A., Perrett, M.E., Binkley, C.G., Jensen, E.L., Oakes, D.L., Rouhier, M.F., Rouhier, K.A., 2017. Disruptions in valine degradation affect seed development and germination in Arabidopsis. Plant J. 90, 1029–1039. https://doi.org/10.1111/tpj.13538

Halket, J.M., Waterman, D., Przyborowska, A.M., Patel, R.K., Fraser, P.D., Bramley, P.M., 2005. Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS. J Exp Bot 56, 219–243. https://doi.org/10.1093/jxb/eri069

Heemann, V., Brümmer, U., Paulsen, C., Seehofer, F., 1983. Composition of the leaf surface gum of some Nicotiana species and Nicotiana tabacum cultivars. Phytochemistry 22, 133–135. https://doi.org/https://doi.org/10.1016/S0031-9422(00)80073-4

Heilmann, I., 2020. Phosphoinositide Signaling in Plants.

Huijser, P., Schmid, M., 2011. The control of developmental phase transitions in plants. Development 138, 4117 LP – 4129. https://doi.org/10.1242/dev.063511

Iizuka, T., Kuboyama, T., Marubashi, W., Oda, M., Tezuka, T., 2012. Nicotiana debneyi has a single dominant gene causing hybrid lethality in crosses with N. tabacum. Euphytica 186, 321–328. https://doi.org/10.1007/s10681-011-0570-3

Landry, N., Pillet, S., Favre, D., Poulin, J.-F., Trépanier, S., Yassine-Diab, B., Ward, B.J., 2014. Influenza virus-like particle vaccines made in Nicotiana benthamiana elicit durable, poly-functional and cross-reactive T cell responses to influenza HA antigens. Clin. Immunol. 154, 164–177. https://doi.org/https://doi.org/10.1016/j.clim.2014.08.003

Laskar, D.D., Corea, O.R.A., Patten, A.M., Kang, C., Davin, L.B., Lewis, N.G., 2010. 6.17 - Vascular Plant Lignification: Biochemical/Structural Biology Considerations of Upstream Aromatic Amino Acid and Monolignol Pathways, in: Liu, H.-W. (Ben), Mander, L.B.T.-C.N.P.I.I. (Eds.), . Elsevier, Oxford, pp. 541–604. https://doi.org/https://doi.org/10.1016/B978-008045382-8.00647-X

Lee, Yuree, Kim, E.-S., Choi, Y., Hwang, I., Staiger, C.J., Chung, Y.-Y., Lee, Youngsook, 2008. The Arabidopsis phosphatidylinositol 3-kinase is important for pollen development. Plant Physiol. 147, 1886–1897. https://doi.org/10.1104/pp.108.121590

Li, L., Zhao, J., Zhao, Y., Lu, X., Zhou, Z., Zhao, C., Xu, G., 2016. Comprehensive investigation of tobacco leaves during natural early senescence via multi-platform metabolomics analyses. Sci. Rep. 6, 37976. https://doi.org/10.1038/srep37976

Mallhi, Z.I., Rizwan, M., Mansha, A., Ali, Q., Asim, S., Ali, S., Hussain, A., Alrokayan, S.H., Khan, H.A., Alam, P., Ahmad, P., 2019. Citric Acid Enhances Plant Growth, Photosynthesis, and Phytoextraction of Lead by Alleviating the Oxidative Stress in Castor Beans. Plants . https://doi.org/10.3390/plants8110525

Molina-Hidalgo, F.J., Vazquez-Vilar, M., D’Andrea, L., Demurtas, O.C., Fraser, P., Giuliano, G., Bock, R., Orzáez, D., Goossens, A., 2020. Engineering Metabolism in Nicotiana Species: A Promising Future. Trends Biotechnol. https://doi.org/https://doi.org/10.1016/j.tibtech.2020.11.012

Moon, K.M., Kwon, E.-B., Lee, B., Kim, C.Y., 2020. Recent Trends in Controlling the Enzymatic Browning of Fruit and Vegetable Products. Mol. . https://doi.org/10.3390/molecules25122754

Nishihara, M., Nakatsuka, T., Yamamura, S., 2005. Flavonoid components and flower color change in transgenic tobacco plants by suppression of chalcone isomerase gene. FEBS Lett. 579, 6074–6078. https://doi.org/https://doi.org/10.1016/j.febslet.2005.09.073

Nogueira, M., Enfissi, E.M.A., Martínez Valenzuela, M.E., Menard, G.N., Driller, R.L., Eastmond, P.J., Schuch, W., Sandmann, G., Fraser, P.D., 2017. Engineering of tomato for the sustainable production of ketocarotenoids and its evaluation in aquaculture feed. Proc. Natl. Acad. Sci. 201708349. https://doi.org/10.1073/pnas.1708349114

Nugroho, L.H., Verpoorte, R., 2002. Secondary metabolism in tobacco. Plant Cell. Tissue Organ Cult. 68, 105–125. https://doi.org/10.1023/A:1013853909494

Orlando, F., Napoli, M., Dalla Marta, A., Natali, F., Mancini, M., Zanchi, C., Orlandini, S., 2011. Growth and Development Responses of Tobacco (Nicotiana tabacum L.) to Changes in Physical and Hydrological Soil Properties Due to Minimum Tillage. Am. J. Plant Sci. 02. https://doi.org/10.4236/ajps.2011.23038

Padilla-González, G.F., Frey, M., Gómez-Zeledón, J., Da Costa, F.B., Spring, O., 2019. Metabolomic and gene expression approaches reveal the developmental and environmental regulation of the secondary metabolism of yacón (Smallanthus sonchifolius, Asteraceae). Sci. Rep. 9, 13178. https://doi.org/10.1038/s41598-019-49246-2

Pajoro, A., Biewers, S., Dougali, E., Leal Valentim, F., Mendes, M.A., Porri, A., Coupland, G., Van de Peer, Y., van Dijk, A.D.J., Colombo, L., Davies, B., Angenent, G.C., 2014. The (r)evolution of gene regulatory networks controlling Arabidopsis plant reproduction: a two-decade history. J. Exp. Bot. 65, 4731–4745. https://doi.org/10.1093/jxb/eru233

Pinu, F.R., Goldansaz, S.A., Jaine, J., 2019. Translational Metabolomics: Current Challenges and Future Opportunities. Metabolites 9, 108. https://doi.org/10.3390/metabo9060108

Rodgman, A., Perfetti, T., 2013. The Chemical Components of Tobacco and Tobacco Smoke, Second Edition.

Sathish, S., Preethy, K.S., Venkatesh, R., Sathishkumar, R., 2018. Rapid enhancement of α-tocopherol content in Nicotiana benthamiana by transient expression of Arabidopsis thaliana Tocopherol cyclase and Homogentisate phytyl transferase genes. 3 Biotech 8, 485. https://doi.org/10.1007/s13205-018-1496-4

Schlimme, M., Blaschke, L., Lagrimini, L.M., Polle, A., 2002. Growth Performance and Lignification in Tobacco with Suppressed Apoplastic Anionic Peroxidase Activity under Ambient and Elevated CO2 Concentrations. Int. J. Plant Sci. 163, 749–754. https://doi.org/10.1086/341515

Song, Z., Luo, Y., Wang, W., Fan, N., Wang, D., Yang, C., Jia, H., 2020. NtMYB12 Positively Regulates Flavonol Biosynthesis and Enhances Tolerance to Low Pi Stress in Nicotiana tabacum. Front. Plant Sci. 10, 1683. https://doi.org/10.3389/fpls.2019.01683

Sousa, E., Kost, B., Malhó, R., 2008. &lt;em&gt;Arabidopsis&lt;/em&gt; Phosphatidylinositol-4-Monophosphate 5-Kinase 4 Regulates Pollen Tube Growth and Polarity by Modulating Membrane Recycling. Plant Cell 20, 3050 LP – 3064. https://doi.org/10.1105/tpc.108.058826

Tramontina, R., Galman, J.L., Parmeggiani, F., Derrington, S.R., Bugg, T.D.H., Turner, N.J., Squina, F.M., Dixon, N., 2020. Consolidated production of coniferol and other high-value aromatic alcohols directly from lignocellulosic biomass. Green Chem. 22, 144–152. https://doi.org/10.1039/C9GC02359C

Wylie, S.J., Zhang, C., Long, V., Roossinck, M.J., Koh, S.H., Jones, M.G.K., Iqbal, S., Li, H., 2015. Differential Responses to Virus Challenge of Laboratory and Wild Accessions of Australian Species of Nicotiana, and Comparative Analysis of RDR1 Gene Sequences. PLoS One 10, e0121787.

Xia, J., Wishart, D.S., 2016. Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis, in: Current Protocols in Bioinformatics. John Wiley & Sons, Inc. https://doi.org/10.1002/cpbi.11

Xu, Z., Ma, J., Qu, C., Hu, Y., Hao, B., Sun, Y., Liu, Z., Yang, H., Yang, C., Wang, H., Li, Y., Liu, G., 2017. Identification and expression analyses of the alanine aminotransferase (AlaAT) gene family in poplar seedlings. Sci. Rep. 7, 45933. https://doi.org/10.1038/srep45933

Zenkner, F.F., Margis-Pinheiro, M., Cagliari, A., 2019. Nicotine Biosynthesis in Nicotiana: A Metabolic Overview. Tob. Sci. 56, 1–9. https://doi.org/10.3381/18-063

Zhang, L., Zhang, X., Ji, H., Wang, W., Liu, J., Wang, F., Xie, F., Yu, Y., Qin, Y., Wang, X., 2018. Metabolic profiling of tobacco leaves at different growth stages or different stalk positions by gas chromatography–mass spectrometry. Ind. Crops Prod. 116, 46–55. https://doi.org/https://doi.org/10.1016/j.indcrop.2018.02.041

Zheng, Z.-L., 2009. Carbon and nitrogen nutrient balance signaling in plants. Plant Signal. Behav. 4, 584–591. https://doi.org/10.4161/psb.4.7.8540

# Figure legends

**Fig. 1.** Metabolite profiling of *N. benthamiana* (A) and *N. tabacum* (B) from juvenile stage to emergence of seed pods. Sampling time points are indicated in consecutive order 1-10, see legend. Score plots are based on >4000 molecular features detected in polar leaf extracts. Each biological replicate is displayed individually.

**Fig. 2.** PCA analysis of metabolites detected in leaf tissue of *N. benthamiana* and *N. tabacum.* Metabolites were grouped into primary (A, B) and secondary (C, D) metabolites for *N benthamiana* (A, C) and *N. tabacum* (B, D). Sampling time points are indicated in consecutive order 1-10, see legend. Biological replicates are displayed individually. Relative quantification of nicotine (E) and total phenylpropanoid derived compounds (F) at each time point for *N. benthamiana* and *N. tabacum*.

**Fig. 3.** Metabolite pathway display highlighting metabolites associated with developmental stages in *N. benthamiana*. Association was based on the presence of the metabolite at the highest levels compared to the other stages. Stage 1 included time points 1 and 2 (yellow), stage 2 included time points 3 and 4 (light green), stage 3 included time points 5 and 6 (dark green) and stage 4 included time point 7 (olive green). Metabolites not significantly different between the time points are highlighted in grey and metabolites not detected in the present data set are highlighted white. Metabolite name with asterisk (\*) indicate level 2 identification based on MS/MS fragmentation. Abbreviations: CoQA – coumaroyl quinate; CoSh – coumaryl shikimate; CoHex – coumaroyl hexoside; CGA – chlorogenic acid; CGAgly – chlorogenic acid glycoside; FQA – feruloyl quinate; HC – hydrocarbon; Unp – polar unknown.

**Fig. 4.** Metabolite pathway display highlighting metabolites associated with developmental stages of *N. tabacum*. Association was based on the presence of the metabolite at the highest levels compared to the other stages. Stage 1 included time points 1 and 3 (yellow), stage 2 included time points 4 and 5 (light green), stage 3 included time points 2 and 6 (dark green), stage 4 included time points 7 to 9 (pink) and stage 5 included time point 10 (olive green). Metabolites not significantly different between the time points are highlighted in grey and metabolites not detected in the present data set are highlighted white. Metabolite name with asterisk (\*) indicate level 2 identification based on MS/MS fragmentation. Abbreviations: CoQA – coumaroyl quinate; CoSh – coumaryl shikimate; CoHex – coumaroyl hexoside; CGA – chlorogenic acid; CGAgly – chlorogenic acid glycoside; FQA – feruloyl quinate; HC – hydrocarbon; Unp – polar unknown.