Figures and figure supplements

A new synthetic biology approach allows transfer of an entire metabolic pathway from a medicinal plant to a biomass crop

Paulina Fuentes et al
The canonical pathway of artemisinin synthesis starts with the conversion of IPP/DMAPP (C5 isoprenoids produced by the MVA pathway in the cytosol or the MEP pathway in the chloroplast) into farnesyl pyrophosphate (FPP), catalyzed by farnesyl pyrophosphate synthase (FPS). Amorpha-4,11-diene synthase (ADS) converts FPP into amorpha-4,11-diene in the first committed step of the pathway. Amorpha-4,11-diene is then successively oxidized to artemisinic alcohol, artemisinic aldehyde and artemisinic acid by the cytochrome P450 monooxygenase CYP71AV1 (CYP) and its redox partner, the cytochrome P450 reductase (CPR). In A. annua, artemisinic aldehyde is converted to dihydroartemisinic aldehyde by DBR2, and then to dihydroartemisinic acid by ALDH1. Artemisinin is generated by the spontaneous oxidation of dihydroartemisinic acid in planta, and can be produced by chemical conversion of artemisinic acid in vitro. Enzymes depicted in red improve the efficiency of different oxidation steps in yeast (Paddon et al., 2013; Paddon and Keasling, 2014). See text for details.

DOI: 10.7554/eLife.13664.003
Figure 2. Implementation of the canonical pathway of artemisinic acid biosynthesis in chloroplasts. Synthetic codon-optimized genes for the four enzymes required to produce artemisinic acid (Figure 1) were introduced into the tobacco plastid genome by stable genetic transformation with four...
different synthetic operon constructs (pAO1-4). The constructs differ in gene arrangement and in the translation signals that drive synthesis of the key pathway enzyme (ADS) catalyzing the first committed step. (A) Physical map of the plastid genome region (ptDNA) used for integration of the synthetic artemisinic acid operons and maps of the transgenic loci in the generated transplastomic tobacco lines (Nt-AO1-4). The artemisinic acid operon genes are depicted as light blue boxes. Chloroplast promoters and terminators are shown in green, the aadA selectable marker gene for chloroplast transformation is represented as a white box, and genes in flanking plastid sequences used for transgene targeting via homologous recombination are in black. Genes above the line are transcribed from left to right, genes below the line are transcribed in the opposite direction. The four transgenes are arranged in two dicistronic operons. FPS and CYP are driven by the *Chlamydomonas reinhardtii* plastid ribosomal RNA operon promoter (*Cr* *P*rrn) and the *g10* leader sequence from phage T7 (*T7* *Lg10*). The second operon containing ADS and CPR is driven by the *C. reinhardtii* psbA promoter (*Cr* *P*psbA) and either the *T7* Lg10 or the psbA leader sequence from *C. reinhardtii* (*Cr* LpsbA). This operon is arranged either in sense and downstream of the first operon (AO1, 3) or in antisense, downstream of the aadA cassette (AO2, 4). The genes in each operon are separated by an intercistronic expression element (IEE) conferring intercistronic RNA processing and, in this way, enhancing expression of downstream cistrons of the operon (*Zhou et al., 2007; Drechsel and Bock, 2010*). The BamHI restriction sites used in RFLP analyses and the expected fragment sizes are indicated. The location of the hybridization probe is shown as a black bar. *Cr*: *C. reinhardtii*; *Nt*: *N. tabacum*; *T7*: bacteriophage T7; *P*: promoter; *L*: leader sequence; *T*: terminator; *SD*: Shine-Dalgarno sequence. (B) RFLP analysis of transplastomic plants. Two independently isolated transplastomic lines are shown for constructs pAO1-3 and one for pAO4. (C) Seed assays confirming the homoplasmic state of the transplastomic plants. Seeds were germinated on medium containing 500 mg/L spectinomycin (Nt-AO2-1, Nt-AO3-1, Nt-wt) or antibiotic-free medium (Nt-wt). DOI: 10.7554/eLife.13664.004
Figure 3. Phenotype of transplastomic tobacco plants and accumulation of artemisinic compounds. (A) Transplastomic lines Nt-AO1-1 and Nt-AO3-1 display a slightly pale and growth-delayed phenotype at the juvenile stage. WAT: weeks after transfer from tissue culture to soil; scale bars: 10 cm. (B) Amorpha-4,11-diene is synthesized in all transplastomic lines, but accumulates to lower levels in the lines displaying an altered phenotype (purple bars). (C) Artemisinic alcohol is detected in similar amounts in all transplastomic plants. (D) Accumulation of artemisinic acid correlates with the altered phenotype of Nt-AO1-1 and Nt-AO3-1. Relative accumulation of amorpha-4,11-diene was profiled by GC-MS analysis of volatile organic compounds (VOCs). Relative accumulation of the sum of free and conjugated artemisinic alcohol and artemisinic acid were determined by GC-MS analysis of the soluble metabolite fraction after saponification (see Materials and methods; Figures 6 and 7). In agreement with previous reports (van Herpen et al., 2010), these compounds were found to be present mainly as conjugates. Expanding leaves of 5–6 plants per line were used for each measurement. Error bars represent the SD. Different letters above the bars indicate significant differences as determined by One-way ANOVA (p<0.001) and the Holm-Sidak post-hoc test.

DOI: 10.7554/eLife.13664.005
Figure 3—figure supplement 1. Phenotypes of Nt-AO2-1 and Nt-AO3-1 plants throughout development. Six plants per line (Nt-wt, Nt-AO2-1 and Nt-AO3-1) were grown under standard greenhouse conditions and photographs were taken of one representative plant per line at different time points: young plants (before flowering, stage 1), flowering plants (stage 2) and old plants (seed capsules formed, stage 3). Light-green leaves and slightly delayed growth of line Nt-AO3-1 are more evident at the young stage. At later stages, all Nt-AO lines display a wild type-like phenotype and produce viable seeds in indistinguishable amounts. y: young leaf; i: expanding (intermediate) leaf; m: fully expanded (mature) leaf.

DOI: 10.7554/eLife.13664.006
Figure 3—figure supplement 2. Isoprenoids levels throughout development in wild-type Nicotiana tabacum plants (Nt-wt) and the transplastomic lines Nt-AO2-1 and Nt-AO3-1. Plants were grown under standard greenhouse conditions and samples were taken from young (y), expanding (i), and fully expanded (m) leaves at three developmental stages (1–3; cf. Figure 3—figure supplement 1). Metabolite levels were determined by UPLC analysis. The values represent the peak height for each compound divided by 10,000 and normalized to the fresh weight (FW), resulting in the normalized response/FW. Error bars represent the SD (n = 3 plants per line). Different letters above the bars indicate significant differences as determined by One-way ANOVA (p<0.05) and the Holm-Sidak post-test.

DOI: 10.7554/eLife.13664.007
Production of artemisinic acid is maintained throughout plant development and correlates with a high CYP/CPR expression ratio. Artemisinic compounds and expression levels of the transgenes were measured in young (stage 1), flowering (stage 2) and old plants (stage 3; see Figure 3—Figure 4 continued on next page)
Figure 4 continued

**figure supplement 1.** (A) Amorpha-4,11-diene accumulates to higher levels in line Nt-AO2-1 than in Nt-AO3-1. (B) Artemisinic alcohol is present at similar levels in early (1) and late stages (3) of development in lines Nt-AO2-1 and Nt-AO3-1, but it is slightly higher in the flowering stage (2) of line Nt-AO3-1. (C) Artemisinic acid accumulates to high levels during all developmental stages of line Nt-AO3-1, whereas in line Nt-AO2-1, it is detectable only in mature leaves of young and flowering plants. Relative accumulation of amorpha-4,11-diene and artemisinic alcohol was profiled, the tissue content of artemisinic acid was quantified using an authenticated reference standard (n = 5–6 plants per line, Figures 6 and 7). The sum of free and conjugated artemisinic alcohol and artemisinic acid were determined: y: young leaf; i: expanding (intermediate) leaf; m: fully expanded (mature) leaf. Error bars represent the SD. (D–G) Northern blot analysis of the expression of the four transgenes. Total RNA samples from *N. tabacum* wild-type (Nt-wt) plants and the transplastomic lines Nt-AO2-1 and Nt-AO3-1 (at the developmental stages 1–3) were separated in denaturing 1.5% agarose gels, blotted and hybridized to strand-specific RNA probes. Below each blot, the rRNA-containing region of the ethidium bromide-stained gel prior to blotting is shown as a control for RNA integrity and equal loading. The Nt-wt sample corresponds to RNA extracted from a fully expanded leaf of a *N. tabacum* wild-type plant at developmental stage 2. The smallest labeled band in each blot corresponds to the monocistronic mRNA. Larger bands represent unprocessed polycistronic precursor transcripts and read-through transcripts (which are common in plastids; e.g., Elghabi et al., 2011; Lu et al., 2013). CYP transcripts accumulate to higher levels in line Nt-AO3-1, while CPR transcripts accumulate to higher levels in line Nt-AO2-1, resulting in a higher CYP/CPR expression ratio in line Nt-AO3-1.

DOI: 10.7554/eLife.13664.008
Figure 4—figure supplement 1. Quantitation of the expression of CYP and CPR by qRT-PCR analysis. Samples from young (y), expanding (intermediate; i) and fully expanded (mature; m) leaves were measured in three technical replicates for each line, in early, flowering and late developmental stages (1–3; cf. Figure 3—figure supplement 1). The wild-type sample (wt) corresponds to a sample from a fully expanded leaf of N. tabacum cv. Petit Havana at the flowering stage. The qRT-PCR data confirm the northern blot analyses that had revealed a higher CYP to CPR expression ratio in transplastomic line Nt-AO3-1 than in line Nt-AO2-1.

DOI: 10.7554/eLife.13664.009
Isolation of combinatorially supertransformed transplastomic lines with a strong increase in artemisinic acid accumulation. Transplastomic lines Nt-AO2-1 and Nt-AO3-1 were combinatorially supertransformed with genes for additional enzymes of the pathway (Figure 1) to facilitate large-scale screening for increased artemisinic acid production. (A) Relative artemisinic acid levels (given in response/FW; see Materials and methods) in the T1 generation of two combinatorially supertransformed lines obtained with transplastomic recipient line Nt-AO2-CS and eight lines obtained with transplastomic recipient line Nt-AO3-CS. An up to 77-fold increase in artemisinic acid was achieved (line Nt-AO3-CS180) in comparison to recipient line Nt-AO3-1. For a complete list of screened supertransformed lines, see Figure 5—source data 1. (B) Inverse relationship.

Figure 5 continued on next page
between artemisinic acid accumulation and artemisinic alcohol accumulation in supertransformed lines. Fully expanded leaves of 5–6 plants per line (at the flowering stage) were used for metabolite profiling. The sum of free and conjugated artemisinic alcohol and artemisinic acid were determined. (C) qRT-PCR analysis of transgene expression suggests a predominant role of \textit{ALDH1} in boosting artemisinic acid synthesis. 2–3 plants per line were measured, and the expression levels were ranked after One-way ANOVA comparison (p<0.05). Brown color indicates absence of gene expression. (D) Combinatorially supertransformed lines with a high increase in artemisinic acid (\textit{Nt-AO2-CS132} and \textit{Nt-AO3-CS180}) display a similar phenotype as the corresponding transplastomic recipient line. WAT: weeks after transfer from tissue culture to soil; scale bars: 10 cm.

\textbf{Source data 1.} Metabolic screening (phenotyping) and genotyping of the T0 generation of combinatorially supertransformed \textit{Nt-AO-CS} lines.

\textbf{DOI:} 10.7554/eLife.13664.011
Figure 5—figure supplement 1. Amorpha-4,11-diene, dihydroartemisinic alcohol and dihydroartemisinic acid accumulation in the T1 generation of combinatorially supertransformed plants. Levels of amorpha-4,11-diene in
Figure 5—figure supplement 1 continued

Nt-AO-CS lines do not correlate with the strong increase in artemisinic acid (Figure 5). Dihydroartemisinic alcohol accumulates to higher levels in lines Nt-AO2-CS95 and Nt-AO3-CS145. Dihydroartemisinic acid is detectable only in two of the best-performing COSTREL lines. Amorpha-4,11-diene levels were determined by GC-MS of VOCs. Dihydroartemisinic alcohol and dihydroartemisinic acid were measured by GC-MS analysis of soluble saponified metabolites. The sum of free and conjugated artemisinic compounds was determined. Error bars represent the SD (n = 5–6 plants per line).

DOI: 10.7554/eLife.13664.012
Measurements of plant height and total leaf biomass of COSTREL line Nt-AO3-CS180 (progeny of four different T1 lines), its transplastomic progenitor line Nt-AO3-1 and the wild type (wt). (A) Comparison of plant height at the same age. Plants were measured when the wild type started to flower. (B) Comparison of plant height at the same developmental stage. Transplastomic line Nt-AO3-1 and COSTREL line Nt-AO3-CS180 were measured five days later than the wild type to compensate for their slightly delayed onset of flowering. (C) Comparison of total leaf biomass (fresh weight, FW) at the same plant age. Plants were measured when the wild type started to flower. (D) Comparison of total leaf biomass at the same developmental stage. Transplastomic line Nt-AO3-1 and COSTREL line Nt-AO3-CS180 were measured five days later than the wild type to compensate for their delayed flowering. Note that, at the same plant age, transplastomic line Nt-AO3-1 and COSTREL line Nt-AO3-CS180 are slightly shorter and produce less total leaf biomass than the wild type. Once the transplastomic plants and the COSTREL lines flower (same developmental stage, five days later), the COSTREL plants reach a height and a total leaf biomass that is close to the values measured for the wild type. On average, the COSTREL plants are 7% shorter and produce 13% less leaf biomass than wild-type plants. No significant difference in either height or total leaf FW was observed between COSTREL plants and their transplastomic progenitor line Nt-AO3-1. Error bars represent the SD (n = 6). Different letters above the bars indicate significant differences as determined by One-way ANOVA (p<0.05) and the Holm-Sidak post-hoc test.

DOI: 10.7554/eLife.13664.013
Figure 6. Chromatograms and mass spectra of amorpha-4,11-diene and artemisinic alcohol. Characteristic peaks for one specific fragment at the expected retention time or index are displayed for each compound. (A) Amorpha-4,11-diene-specific mass feature 119 at a retention time of 1564 s. This metabolite is present in all Nt-AO lines, and at slightly higher levels in lines Nt-AO2-1 and Nt-AO4-1. (B) Artemisinic alcohol-specific mass feature 202 at a retention index of 1784. The compound is present at similar levels in all Nt-AO lines. Both compounds are absent from the wild-type sample. In addition to the chromatograms, the characteristic mass spectrum (m/z) of each compound is shown for the standard and for one of the artemisinic acid operon lines. EPY224: yeast strain that produces amorpha-4,11-diene (Ro et al., 2006). One representative plant per line is depicted. Mass spectra and mass features of trimethylsilylated artemisinic alcohol are shown.

 DOI: 10.7554/eLife.13664.016

Fuentes et al. eLife 2016;5:e13664. DOI: 10.7554/eLife.13664

Research article
Biochemistry | Plant biology
Figure 7. Chromatograms and mass spectra of the artemisinic compounds artemisinic acid, dihydroartemisinic alcohol and dihydroartemisinic acid. Characteristic peaks for one specific fragment at the expected retention index are shown for each compound. (A) Artemisinic acid-specific mass feature.
Figure 7 continued

216 is shown at a retention index of 1850. This compound accumulates to higher levels in lines Nt-AO1-1 and Nt-AO3-1. (B) Dihydroartemisinic alcohol-specific mass feature 162 at a retention index of 1789. The compound is present at high levels in COSTREL line Nt-AO2-CS95, but is absent from transplastomic line Nt-AO2-1. (C) Dihydroartemisinic acid-specific mass feature 163 at a retention index of 1859. This compound accumulates in COSTREL line Nt-AO3-CS180, but is absent from transplastomic line Nt-AO3-1. All compounds are absent from the wild-type sample. In addition to the chromatograms, the characteristic mass spectrum of each compound is shown for the standard and for one of the artemisinic acid operon lines. Mass spectra and mass features of trimethylsilylated artemisinic compounds are shown.

DOI: 10.7554/eLife.13664.017