Final Publishable Summary

Project Acronym:	MultiBioPro
Project full title:	The development and evaluation of Multipurpose crops biorefining feedstocks for the production of industrial BioProducts and biomass
Grant Agreement no:	311804
Project website:	www.multibiopro.eu

Coordinator:	Dr. Alisdair Fernie
Organisation:	Max Planck Institute of Molecular Plant Physiology
Phone:	+ 49 (0)331 567 8211
Email:	fernie@mpimp-golm.mpg.de

1. Executive summary

Over the past few years, renewable energy sources have moved into the spotlight of modern society. Fossil sources of energy tend to run short and, moreover, cause severe environmental damage. Against this backdrop, MultiBioPro aimed to develop and implement multipurpose crops that will deliver improved biomass extractability as well as new sources of non-food oils and biomaterials. For this purpose two species capable of growth on marginal land, namely the woody species poplar and *Nicotiana glauca* (tree tobacco), were utilized as backgrounds for the evaluation of field growth characteristics, biorefining potential and the potential for the extraction of biopolymers and metabolites of commercial interest. These backgrounds both represent industrially utilized species, albeit the use of poplar currently far exceeds that of *N. glauca*. Given this fact we established new field sites for *N. glauca* in the United Arab Emirates and in Madagascar. By contrast, poplar was grown at established experimental field sites in Belgium. The MultiBioPro project explored already existing poplar and *N. glauca* genotypes. In parallel the project used innovative translational research techniques which leverage research from the model organism *Arabidopsis thaliana* to identify targets for genetic engineering, for improved fiber and long-chain fatty acid quality. An extensive crosstalk between the analytical platforms, biorefining and field trials allowed the provision of the basis for industrial determination of *N. glauca* and poplar varieties as novel multi-purpose feedstocks. Residual biomass, such as suberin in the poplar bark was analysed and demonstrated to yield long-chain fatty acids with potentially important characteristics as biomaterials as well as for biorefinery purposes. Indeed chemical analysis of the entire biorefinery process allowed the identification of several lead compounds of high potential industrial value. Both poplar and *N. glauca* genotypes can readily be grown in areas that are sub-optimal for food-crop species, and field trials of both species contribute knowledge to the environmental sustainability of the studied genotypes. Moreover, attributes which are advancing biorefinery and bio-materials properties were further exploited towards marketing in a manner that is demonstrative of the industry-driven nature of the MultiBioPro project.

Dissemination and training events have constituted an integral part of the MultiBioPro project. All project partners participated in multiple dissemination and exploitation activities targeted at the scientific community, the broad public and interested stakeholders/user groups. Throughout the five year project the consortium was dedicated to make its findings accessible and to share news regarding project progress and public events.

2. Summary description of project context and the main objectives

The finite nature of our fossil energy reserves has rendered investigation into renewable energy resources highly important. When the ever expanding human population and the

need to boost crop yield in the face of environmental deterioration is also considered, the optimal solution would be to cultivate biofuels (dedicated biomass crops) on land which is marginal from an agricultural perspective. Moreover, in addition to components that can be utilized for biofuel, plants contain a vast diversity of secondary metabolites which are compounds that are non-essential to fundamental plant processes, but have an important role in the plant's interaction with its environment. Many secondary metabolites are also essential components of the human diet, can be utilized as phytomedicines and are routinely used as industrial raw materials and high-value fine chemicals. Plant secondary metabolites can be structurally divided into five major groups: polyketides, alkaloids, phenylpropanoids, flavonoids and isoprenoids. In the MultiBioPro project we have performed broad profiling of phenylpropanoids, flavonoids and isoprenoids alongside lipids and primary metabolites utilizing a range of analytical platforms. In poplar this was largely carried out as comparative studies of wild type and lignin deficient genotypes. However, given the relative paucity of available data for *N. glauca,* a deep characterization of the chemical composition in different tissues and across leaf development was performed, as was a characterization of carotenoid and terpenoid modified genotypes. Furthermore, extensive pilot scale refinery trials of both poplar and *N. glauca* were performed and the chemical composition of refined material was tested at multiple time points in the procedure to allow biorefinery engineering in order to maximize the fuel yield but also, in conjuncture with the metabolic profiling of the biostocks themselves in order to search for additional compounds of industrial importance. For both poplar and *N. glauca,* several potential compounds were identified at high enough levels to fit this criterion; and discourse concerning their valorization has been initiated with multinational companies.

As a starting point for the design of the genetic perturbation strategies for poplar and *N. glauca,* we drew on the vast experimental resources of (and prior research carried out in) the model plant *Arabidopsis thaliana*. In this species, we were able to demonstrate that the currently accepted model of lignin biosynthesis required revision to reflect the importance of the enzyme caffeoyl shikimate esterase (CSE) as an enzyme central to the lignin biosynthetic pathway (Vanholme et al., 2013; Science 341: 1103-1106), whilst the importance of lignocellulosic biomass in pathogen resistance was also characterized in this species (Miedes et al., 2014; Frontiers in Plant Science 5:358) and mutants of the cytochrome P450 reductase 2 of Arabidopsis alters lignin composition and improves saccharification (the process of breaking a complex carbohydrate into its monosaccharide components; Sundin et al., 2014 Plant Physiol. 166: 166: 1956-1971). Similarly, it was used as a test plant for the introduction of a chemically labile substructure in lignin (Tsuji et al., 2015; Plant Biotech J. 13: 821-832), analysis of the consequences of the downregulation of the aforementioned CSE (Vargas et al., 2016; Biotechnol for Biofuels: 9: 134) and

characterization of a recently evolved novel flavonol phenylacyltransferase gene. The latter of which produces a novel form of flavonol with enhanced UV-B absorbant properties (Tohge et al., 2016; Nature Communications:7: 12399). We have additionally used the model Solanaceous species tomato to offer insight into *N. glauca* focusing studies here on glycoalkaloids (Schwann et al., 2014; J Integr Plant Biol 56: 864-875; Alseekh et al., 2015; Plant Cell 27: 485-512) and flavonoids (Tohge et al., 2015; Plant J.83: 686-709; Alseekh et al., 2015; Plant Cell 27: 485-512) with the aim of providing a means to identify regulatory genes which were likely conserved in tomato and *N. glauca*. In addition to these experimental analyses, a highly innovative aspect of the MultiBioPro project was the development, implementation and mining of a suite of bioinformatics tools utilizing co-expression analysis of publically available gene expression including the analysis of tobacco and poplar datasets as means of identifying key target genes for the improvement of *N. glauca* and poplar respectively (Ruprecht et al., 2016; Plant Physiol. 170: 1878-1894). In the MultiBioPro project, we demonstrated our genetic engineering approaches using two examples: lignin production in poplar and carotene and terpene production in *N. glauca*. Lignin is a structural polymer that is essential for the support of vertical growth but often unwanted in biofuels since it renders extraction of soluble biomass more problematic whereas carotenoids and terpenes are of very high commercial value.

Considerable advances have been made in the biorefinery processes over the last decade. One of the aims of the MultiBioPro project was to build on these applied foundations and advance our generic knowledge base to a new era in our understanding and hence utilization of multipurpose crops. To this end, the key objectives for the project were:

i) Generation of new tools and strategies for improved biomass exploitability and non-food oils.
ii) Molecular breeding approaches to improve selected multipurpose crops.
iii) The development of technology platforms for the assessment of biomass extractability and chemical composition.
iv) Evaluation of the potential of using poplar biomass for the production of high performance biomaterials.
v) Establishment of suitable biorefinery conditions for poplar and *N. glauca*.
vi) Agronomic and environmental assessment of MultiBioPro products.

Wild type poplar and *N. glauca* and varieties with perturbed lignin and carotene or terpene content have been subjected to concurrent multi-level omic based analyses over the life time of the project. For example, leaf development in *N. glauca* was followed using an integrative omic based approach. Indeed via the integration of transcriptomic and metabolomic datasets the consortium has identified several key features of the *N. glauca* lifestyle including its high oil content, the production of vitamin D precursors in high amounts and hints towards its

extreme water use efficiency. These findings were enabled by access to the genome sequence via the establishment of an *N. glauca* sequencing consortium including partners within and external to the project. These data and those emanating from the above described co-expression analysis were utilized to identify targets for the genetic improvement of poplar and *N. glauca* for enhanced biomass extractability and altered oil content. To this end, functional homologs of glycospingolipid synthase and lipid transfer proteins were overexpressed and/or downregulated in *N. glauca* or poplar. In parallel field trials with poplar engineered for low lignin content via downregulation of cinnamoyl coenzyme A reductase (CCR) or cinnamyl alcohol dehydrogenase were carried out and wood was analysed for fiber quality and suberin amount and composition. In addition, novel lignin structures were engineered using genes that modify the monolignols themselves. In all instances the cell wall properties were studied as were metabolite profiles and the potential for improved saccharification. In addition, MultiBioPro focused on enhancing non-food oils in *N. glauca*. For this purpose, transgenic c3-hydroxy-3-methylgluraryl coenzyme A reductase approaches were taken and oil content as well as general growth characteristics were determined.

Another important component of the MultiBioPro project was the use of state-of-the-art approaches and technologies to assess fundamental properties of the fiber structures and hydrocarbon-based compounds of the above-mentioned plant materials. Knowledge thus obtained was then used for the design of advanced biomaterials and their subsequent fabrication. Fibers from lignin altered poplar trees were evaluated using a battery of techniques. Those fibers identified to be weaker than normal were subjected to saccharification analysis which revealed that they exhibited higher yields whilst those that were stronger than normal were evaluated for biomaterial development. One component of the poplar bark in particular, suberin, a complex polyester based on glycerol and long chain fatty acids and hydroxyacids, and the terpenoid solanesol of *N. glauca* were evaluated intensively with regard to their biomaterial characteristics. Disintegration of woody materials of poplar was tested as a resource of cellulosic building blocks with unique mechanical properties, facilitating the development of foams, hydrogels and cellulose sheets in combination with a range of polymers.

In parallel, multifunctional biorefinery analysis was carried out on both poplar and *N. glauca* biostocks with studies revealing that both biostocks were largely comparable to the industrial standard feedstock maize. A comprehensive chemical analysis across multiple biorefinery strategies was made for both feedstocks, which on the one hand, identified the most appropriate conditions for hydrocarbon extraction, and on the other resulted in the identification of industrially valuable components, either biomaterial polymers or small molecule metabolites being present in high amounts. Discussions with multinationals regarding the valorization of these findings are currently in progress.

The MultiBioPro project has disseminated its findings at the earliest opportunity to the scientific community, general public and government agencies. For example, 21 peer reviewed publications have been generated, several in high impact journals such as Science, Nature Communications and Plant Cell. Over 100 dissemination activities have been delivered and at least two patents generated. In addition, two informative promotional videos of the project have been generated. Training has constituted a cornerstone of the MultiBioPro project with many training events being provided both by academic and industrial partners. The work has additionally strengthened the position and sustainable foundations of several SME partners, which in turn has contributed to rural development. In addition, strong working links with the multinational companies Philip Morris International and Pierre Fabre were established. The project website http://multibiopro.eu acted as a highly successful focal point for information and will continue beyond the RTD activity of the project.

3. Description of main Scientific &Technological results/foregrounds

The RTD activities of MultiBioPro were divided into six work packages (WPs 2 to 7).

Work package 2. Generation of new tools and strategies for improved biomass exploitability and non-food oils.

WP2 Objectives

O2.1. The capture of transcriptomic and metabolomic datasets associated with the biosynthesis, and deposition of lipids in *Nicotiana glauca*.

O2.2. Generation of Nicotiana genome-level co-expression networks, and comparative co-expression analyses across eight plant species (http://aranet.mpimp-golm.mpg.de/; Arabidopsis, barley, Medicago, soybean, poplar, wheat, rice and Brachypodium), to identify genes associated with cell wall synthesis/metabolism, suberin and lipid synthesis and deposition.

O2.3. Functional testing and characterization of Arabidopsis candidate gene(s) identified from the network having the potential to enhance lipid and suberin formation/modification, and alter cell wall properties.

O2.4. Validation of the network based approach as a discovery pipeline for optimisation of lipid and suberin production/secretion, and for improving chemical composition for biomass extractability and biomaterial characteristics.

Activities in WP2 used state-of-the-art approaches and technologies to advance our fundamental understanding of key genetic and metabolic processes essential for the development of advanced multipurpose crops. Since transcriptionally associated genes tend

to be involved in related biological processes, transcriptional co-expression has been used extensively to infer gene functions in many model plants. Recently, several reports have touched upon the notion that co-expressed gene neighborhoods also appear multiple times in an organism. For instance, the primary wall cellulose synthesis neighborhood contains several genes for which close homologs appear in the secondary wall cellulose synthesis neighborhood (Ruprecht et al., 2011; Front Plant Sci. 2:23). Such recent cross-species studies of coefficient correlation based network analysis have indicated that not only genes, but also pathways can be multiplied and diversified to perform related functions in different parts of an organism. Identification of such diversified pathways, or modules, is needed to expand our knowledge of biological processes in plants and to understand how biological functions evolve. Additionally, co-expression patterns which may be conserved across species barriers can be used to transfer knowledge obtained from a well-investigated model species to other organisms. We therefore developed transcriptional co-expression network analysis of Arabidopsis, tomato and tobacco as a starting point for the design of the genetic perturbation strategies for poplar and *N. glauca.* WP2 thus acted as a discovery pipeline delivering resources which fed into the more translational activities of WP3 and the technology platforms of WP4 and 5.

As an example of the genetic perturbation strategies for poplar and *N. glauca,* we drew on the vast experimental resources of (and prior research carried out in) the model plant *Arabidopsis thaliana*. Co-expression network analysis was successfully applied to identify new candidate genes with a potential role in lignin/secondary cell wall (SCW) biosynthesis. A total of 20 genes were selected based on their tight link with lignin/SCW biosynthesis in terms of gene expression and their respective T-DNA mutants were ordered from T-DNA line databases. These mutants were screened based on their phenolic composition. The best candidates (i.e. the genes whose T-DNA mutants showed clear shifts in phenolic content) were selected for further functional characterization using wet chemistry and cell wall characterization. The knockout mutants of two genes showed not only a higher saccharification yield, but also significant changes in lignin content and/or composition. In addition, the changes in the phenolic profile of their respective mutants and overexpression lines were analysed by UPLC-MS/MS. Using this approach in tandem with Arabidopsis transcriptome data, we were able to demonstrate that the currently accepted model of lignin biosynthesis required revision to reflect the importance of the enzyme caffeoyl shikimate esterase (CSE) as an enzyme central to the lignin biosynthetic pathway (Vanholme et al., 2013; Science 341:1103-1106). The importance of lignocellulosic biomass in pathogen resistance was also characterized in this species (Miedes et al., 2014; Frontiers in Plant Science 5:358). Furthermore, mutation of the cytochrome P450 reductase 2 alters lignin composition and improves saccharification (the process of breaking a complex carbohydrate

into its monosaccharide components; Sundin et al., 2014 Plant Physiol. 166:1956-1971). Similarly, it was used as a test plant for the introduction of chemically labile substructure in lignin (Tsuji et al., 2015; Plant Biotech J. 13:821-832.) We additionally focused on other polyphenolic metabolites and characterized a recently evolved novel flavonol phenyl-acyltransferase gene which produces a novel form of flavonol with enhanced UV-B absorbent properties (Tohge et al., 2016; Nature Communications. 7:12399). Furthermore, in order to characterize the function of candidate genes involved in lipid and/or hydrocarbon formation, we constructed a co-expression network of Arabidopsis genes using the tool generated. In total, we selected almost 50 genes and took eight of these forward to functional characterization in WP3.

We have additionally used the model Solanaceous species tomato to offer insight into *N. glauca* in order to focus studies here with respect to glycoalkaloid (Schwann et al., 2014; J Integr Plant Biol. 56:864–875; Alseekh et al., 2015; Plant Cell 27:485-512) and flavonoid (Tohge et al., 2015; Plant J. 83:686-704; Alseekh et al., 2015; Plant Cell 27:485-512) metabolism with the aim of providing a means to identify regulatory genes which were likely conserved in tomato and *N. glauca*. In order to generate a genome-level co-expression network for Nicotiana, we downloaded publicly available microarray datasets for tobacco, calculated mutual rank-based co-expression relationships between genes and constructed a genome-wide co-expression network (http://aranet.mpimp-golm.mpg.de/tobnet) (Ruprecht et al., 2016; Plant Physiol. 170:1878-1894) to enable a cross-species comparison with other plant species (Arabidopsis, barley, medicago, poplar, rice, soybean, wheat and Brachypodium). We also extended this platform with a new bioinformatics tool that analyses transcriptional associations of gene families and finds conserved recurring co-expression vicinities within and across species. We have developed a completely new algorithmic pipeline for this tool, which can improve identification of candidate genes based on comparative co-expression analyses (Famnet http://aranet.mpimp-golm.mpg.de/pfamnet2/). After construction of these network frameworks in model plants and Solanaceous species, a developmental series for leaf material of *N. glauca* was harvested and transcriptome data (RNAseq) and multi-platform metabolite profiling approaches (data provided by WP4 included; general primary metabolites, phenylpropanoids, alkaloids, oligolignols, isoprenoids, carotenoids, tocopherols, solanesol, phytol, phytosterols) were combined to allow integrative “omics” analysis. All data is currently being combined and will be published in tandem with the *N. glauca* genome sequence.

Future work will include publishing the genome sequence together with an integrative analysis of transcriptome and metabolic data obtained in WP2 and WP4. The function of candidate genes obtained from *N. glauca* genome-wide co-expression network combined with multi-platform metabolite profiling approaches for “omics” integration analysis are

additionally being tested experimentally. A greater understanding of the key genes compounds in feed stock materials is very important for designing multipurpose bio-fuel materials and its application in metabolic engineering.

Conclusions

- Genomic, transcriptomic, metabolic data from *N. glauca* leaf developmental series for genome-wide "omics" integration analysis capturing both oil synthesis and deposition, lignin and other beneficial compounds such as phenylpropanoids have been collated and will shortly be disseminated (Objective 2.1)
    - A web-based platform that can explore and visualize multiplied modules in co-expression networks of eight plant species with multiplied modules has been published (Famnet http://www.gene2function.de/famnet.html) (Objective 2.2).
- An algorithmic pipeline for this co-expression network analysis tool has been developed and demonstrated for Arabidopsis, tomato and tobacco focusing on lipid and/or hydrocarbon formation and biosynthesis of lignin, oligolignols, flavonoids and alkaloids (Objective 2.3 and Objective 2.4).
- Candidate genes identified in WP2 have been sent to and evaluated in WP3 and WP4.
- The algorithmic pipeline and the networks developed for Arabidopsis, tomato, tobacco and *N. glauca* in WP2 also form a considerable resource that could be exploited to enhance fundamental understanding of plant metabolism in research areas outside of the focus of the MultibioPro project.

Work Package 3. Molecular breeding approaches to improve selected multipurpose crops

WP3 Objectives

O3.1. Use modern molecular breeding approaches to improve biomass extractability and non-food oil content in poplar and the tobacco tree (*N. glauca*).

O3.2. Create poplar and tobacco tree genotypes with combined biomass, oil and agronomic properties (bioremediation) to increase their potential as a marketable multipurpose feedstock for biorefining.

Nicotiana glauca. The main objective of work WP3 is the generation of poplar and tobacco tree genotypes with improved traits associated with biomass and non-food lipids, improving their potential as feedstocks for the biorefining industry.

The transformation system for *Nicotiana glauca* was optimised and several transformations were carried out with multiple constructs to increase biomass and non-food oil in *N. glauca t*o produce commercially valuable ketocarotenoids, using gene stacking approaches. Stable

lines were generated producing ketocarotenoids and the transformants lines showed a visible phenotype. Several transformations using transcription factors identified in WP2 were performed as well as genes altering phytohormone levels. However, they were lethal, despite several attempts to rescue the transformants in tissue culture.

In order to increase high value phytosterols in *N. glauca,* the *hmgr-1* of *Arabidopsis* was used to transformed *N. glauca* wild-type (WT) plants. HMGR encodes a 3-hydroxy-3-methylgluraryl coenzyme A reductase involved in melavolate biosynthesis and it performs the first committed step in isoprenoid biosynthesis. Stable homozygous lines were produced and phytosterols levels were determined. The leaf tissue of these transformants contained and increased total phytosterol content (up to 1.1-fold) but an unchanged composition. All of the end-product phytosterols (campesterol, stigmasterol, and β-sitosterol) were significantly increased ($P < 0.05$) (Table 1). Some changes in the carotenoid composition in the transformants were observed by the increased phytosterols level. No changes in biomass were observed. These high sterols producing lines represent a valuable renewable resource of these compounds. An additional commercial value of these lines is their use as a source for the bioconversion of natural sterols to the commercially important steroids. Currently, protocols are being optimised with different strains of *Mycobacterium* spp. to test the level of biotransformation of phytosterols present in *N. glauca* high sterols lines into steroids. This activity adds value to the projects outputs.

A range of environmental assessments using WT *N. glauca* and transgenic ketocarotenoid producing plants were performed. These included environmental assessment treatments for

i) Heavy metals: 5mM $ZnSO_4$, , 2mM $NiSO_4$, 0.5 mM $MCuSO_4$
ii) Salinity: 200 mM NaCl
iii) Drought

Guidelines from literature were taken to determine concentrations for salinity and heavy metal (treatments that are typically not well handled by the majority of crops grown). However, no lethality was reached for both lines at two different concentrations. Over the course of 20 days, plants treated with salinity showed some loss of rigidity, while heavy metal treatment showed leaf browning. Moreover, drought treatment showed gradual leaf wilting and eventually death of the plant at around 15 days post the initiation of the experiment.

WT plants at day 5 of drought showed more wilting than the transgenic ketocaroenoid producing line. By day 10, both lines showed a similar level of wilting, but the WT had higher amount of wilt visible phenotypically. By day 15, both lines were found to be completely wilted. An attempt of plant recovery by introducing watering showed no recovery to either of *Nicotiana glauca* plant lines. It was concluded that severe drought affects *Nicotiana glauca* in 15 days to no recovery and it prefers well-watered soil environment to a dry environment.

Poplar. Poplar activities in WP3 involved the translation of the knowledge obtained previously with *Arabidopsis thaliana* in WP2 to the bioenergy tree poplar. The CCR gene was knocked-out by CRISPR/Cas9. After confirmation of the presence of the construct, the selected first transformants have been propagated and evaluated for their cell wall properties and growth. Further results of the poplar trials will be published soon.

Conclusions

- Molecular breeding approaches were implemented for the genetic engineering of tobacco trees to improve non-food oil content and biomass extractability. Gene stacking approaches were used for the generation of ketocarotenoid producing transgenic plants. Due to the lack of available plant sources of ketocarotenoids, this approach provides a means of renewable production of these molecules.
- *N. glauca* transgenic lines producing high levels of phytosterols were also successfully generated. High value compounds of the isoprenoid pathway showed increased levels in these transformants. Furthermore, these high sterols producing lines could be used as a source to feed *Mycobacterium* spp. strains in the bioconversion of natural sterols to the economically important steroids.
- Modern molecular cloning platforms such as GoldenBraid were used to build multigene constructs that would ultimately be used for the engineering of other high value isoprenoids in *N. glauca.* Unfortunately, and despite numerous attempts, the constructs were not efficient and therefore plant transformations were unsuccessful.
- Environmental assessments carried out on WT *N. glauca* and ketocarotenoid transgenic plants showed that *N. glauca* is intrinsically resistant to such environmental conditions/challenges and therefore has a lot of potential for bioremediation treatments.
- Field-grown genetically modified poplar trees altered in either Cinamoyl CoA reductase (CCR) or Cinnamyl alcohol dehydrogenase (CAD) were demonstrated to display similar phenotypes to greenhouse grown Arabidopsis and poplar bearing the same mutation.

Work package 4. Technology platforms for the assessment of biomass extractability and chemical composition.

WP4 objectives.

O4.1. Establish experimental protocols for baseline metabolomics of poplar and *N. glauca*.

O4.2. Primary metabolite and saccharification evaluation of transgenic plants and for Arabidopsis mutants.

O4.3. Perform crude cell wall analysis on poplar transgenics.

O4.4. Perform lignin, cell wall and suberin analyses on specifically engineered *N. glauca* and poplar transgenics.

The main activities in WP4 employed state-of-the-art metabolite profiling platforms to assess i) the chemical composition of *Nicotiana glauca*, *Nicotiana tabacum*, poplar and *Arabidopsis* samples, and ii) derived waste material and fractionated extracts. The following profiling/evaluation platforms were optimized for this purpose a) metabolite profiling procedures for polar derived extracts prepared from poplar and *Nicotiana glauca* material, b) metabolite profiling procedures for non-polar derived extracts prepared from poplar and *Nicotiana glauca* material, c) assessment of macromolecules and d) saccharification of biomass. For these purposes a battery of techniques including GC (gas chromatography)-MS (mass spectrometry), DiMS (direct infusion mass spectrometry), Raman spectrometry, FTIR (Fourier transform infra-red), X-ray, TEM (transmission electron microscopy) and LC (liquid chromatography)-MS were employed. First, we established a platform to analyse high/semi-polar and non-polar compounds for *Nicotiana glauca* and poplar samples (MultiBioPro Consortium, 2014; JoVe. 87: doi:10.3791/51393). This platform includes metabolite profiling for primary metabolites by GC-MS, secondary metabolites by LC-MS, hydrocarbon measurement by HPLC and isoprenoids/phytosterol profiling by GC-MS. None of the extraction and analytical methods in this platform was largely altered from methods typically used for plant materials, but due to the many species/tissue specific metabolites including a large number of unknown compounds were detected in each plant species, data processing was optimized de novo.

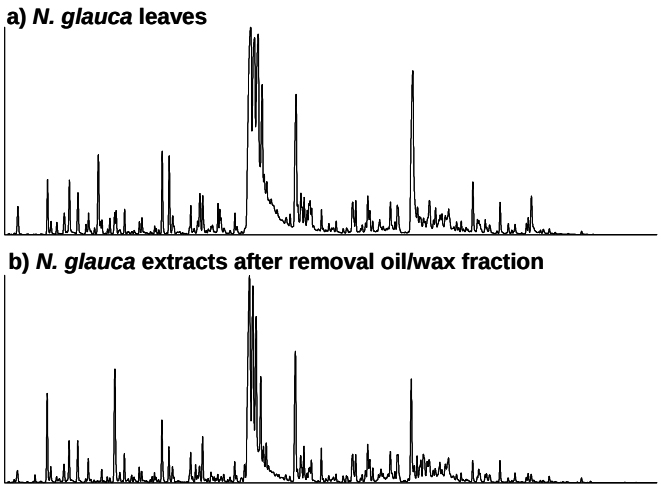

Figure 1: Mass chromatogram of a) *N. glauca* leaves used for extraction and b) *N. glauca* extracts after removal oil/wax fraction.

Within the above-defined work, a different combination of platforms were employed dependent on the sub-project`s aims. Profiling of chemical composition was mainly used for functional evaluation of candidate genes from the network analysis of WP2, evaluation of GM

plants in WP3 and WP5 and derived waste material and fractionated extracts for WP7. The GM and non-GM outputs from WP2 and 3 were also subjected to comprehensive metabolic analysis in order to i) ascertain and evaluate putative molecular breeding strategies and ii) identify potential high-value renewable bioproducts of both macromolecule and small molecule origin. Identification of these commodities acted as an initial basis for the selection of biomass for industrial testing and at the same time adds potential value to the biorefining feedstocks of the multipurpose crops targeted in MultiBioPro. High/semi-polar compounds including primary and secondary metabolites and lipids whilst suberins and phytosterols, and cell wall properties were profiled and focused on, respectively. Plant and plant derived material was shared between partners according to the focus of respective projects. In order to characterize the function of candidate genes involved in lipid and/or hydrocarbon formation identified in the co-expression networks of Arabidopsis in WP2, knockout mutants and transgenic lines of candidate genes were evaluated with regard to their content of non-polar compounds. Additionally, partners in WP4 profiled poplar transgenic including material deficient in CCR (Van Acker et al., 2014; PNAS. 111:845-850) grown both in greenhouse and field trial conditions.

As a second model species, MultiBioPro utilized *Nicotiana tabacum*. Using LC-MS, a total of 105 secondary metabolite peaks were found in *N. tabacum* tissues. Around 30 peaks were identified or annotated based on the co-elution profile of standard chemicals and common plant extracts of well-known model plants such as Arabidopsis and tomato, alongside an extensive literature survey of metabolites characterized in *N. tabacum* tissues. This metabolite profiling established that secondary metabolism, namely chlorogenic acids (CGAs), anthocyanin, flavonol and nicotianoside biosynthesis are largely regulated in a tissue specific manner. Based on this finding, the obtained data were integrated with transcriptional co-expression network analysis in order to elucidate metabolic regulatory mechanisms and identify novel genes responsible (Ruprecht C. *et al.* (2016) FamNet: A framework to identify multiplied modules driving pathway diversification in plants. *Plant Physiol,* 170, 1878-1894).

The results of metabolite profiling of *N. glauca* leaf developmental series were provided to WP2 for integrative analysis. We performed profiling of primary and secondary metabolites, phytosterols, isoprenoids and lipids. The results of metabolite profiling were combined and evaluated by simple clustering and correlation analysis in order to categorize compounds in groups showing similar metabolic changes during leaf development. Clustering analysis resulted in five major clusters comprising i) TCA intermediates and some amino acids, ii) remaining amino acids, iii) unknown compounds, iv) sucrose and chlorogenic acids, and v) glutamates and flavonol glycosides. Within the MultiBioPro project, phenolic profiling of *N. glauca* leaves was performed using reversed phase ultrahigh performance liquid

chromatography (RP-UHPLC) coupled to a Fourier transform Ion Cyclotron Resonance mass spectrometer (FT-ICR-MS) resulting in the detection of a large number of *N. glauca* secondary metabolites. Structural characterizations were based on the recorded accurate mass, MS^n fragmentation spectra, compound database survey, MS databases and MS spectral elucidation software (such as MetFrag). Using these approaches, peaks annotated as "unknown compounds" which decreased during leaf development were annotated as those phenolamides which are commonly detected in other Solanaceae plants such as tomato and *N. tabacum* (their respective biosynthetic genes are equally well-conserved in Solanaceae species). Since phenolamides have been discussed in the context of bioactivity against insects, phenolamides can be considered as one of the beneficial compounds found in *N. glauca*. Phenolamide metabolism was thus further focused on in the network analysis of WP2 in order to elucidate regulatory mechanisms for enhancement of production of phenolamides in Solanaceae species including *N. glauca*.

Biorefinery feedstocks and their corresponding fractions (provided by WP7) were evaluated by the aforementioned platforms in order to facilitate a comprehensive survey of beneficial compounds in fractions of *N. glauca.* Pilot scale extracts and waste fractions which could be obtained after oil extraction were screened using mass spectral analysis. Detected compounds are generally similar to the compounds found in dried leaf materials (fructose, glucose, myo-inositol, sucrose, glycerol, malate, succinate, glycerate, pyroglutamate, oils, fatty acids and lipids). Estimation of absolute concentrations revealed that five sugars constitute a high percentage of biomass in *N. glauca* (~15% in *N. glauca* dry material) extracts. Peaks of putative protocatechuic acid (PCA), which have been reported as an anti-inflammatory antioxidant inducing reported activity against apoptosis of human leukemia cells, were annotated in the water steam extracts of *N. glauca*. We additionally surveyed beneficial metabolites in poplar branches and bark which could be obtained after wood pulp production. For this purpose we screened metabolite profiles using GC/MS analysis for young/old poplar bark and their extracts. As for *N. glauca* materials, both poplar bark feedstocks and their fractions contain extremely higher level of sugars, such as fructose, sucrose and maltose and thus could be considered as candidate targets for bio-ethanol and bio-plastic production. Estimation of absolute concentrations revealed that five sugars constitute a high percentage of biomass in poplar (~18% in poplar dry material) extracts. In spite of high detection of catechins, which are highly beneficial compounds having anti-oxidant and antifungal activities, in the original feed stock poplar bark materials, catechin contents were significantly reduced in almost all fractions obtained from pilot scale extraction. However, we found high content of catechins in the washing waste water in the first step before the poplar bark extraction process. Since catechins can be expected for usage with

their antifungal activity, for example for house blocks, plastic tableware and cat toilet sand etc., they represent a potential valuable product of the refinery.
In summary the objectives of WP4 were fully realised. Most importantly in addition to oil the biorefinery was able to yield considerable amounts of other valuable products in both *N. glauca* and poplar.

Conclusions

- Procedures to analyse primary metabolites, hydocarbons, non-polar components, isoprenoids and phenolic secondary metabolites have been established.
- Metabolite profiling platform established for *N. glauca* and poplar materials in WP4 has been efficiently used in projects in WP2, WP5 and WP7. This platform was also employed for the projects of functional characterization of candidate Arabidopsis genes and networks analysis of *N. tabacum*.

Work Package 5. Assessing poplar biomass for the production of bio-materials with high performance and functionality

WP5 objectives

O5.1. Determine structural, surface and mechanical properties of fibers from transgenic poplar.
O5.2. Production of polyesters for biomaterial design and fabrication using long hydrocarbon chains and alkanes from *N. glauca* and refined suberin from poplar bark.
O5.3. Design, fabrication and assessment of the properties of cellulose-based composites.

The main objective of WP5 was to isolate and determine the properties of individual compounds produced by plant varieties generated in WP2 and WP3, and to exploit the isolated (macro)molecules for the generation of novel high-performance green materials. Of particular interest were carbohydrate-based fiber structures from wild-type and transgenic poplar and long hydrocarbon chains from bark suberin and *Nicotiana glauca* leaves. Lignocellulosic fibers from poplar were used for the preparation of biocomposite materials. Long hydrocarbon chains, on the other hand, were used for the production of polyesters that can subsequently be incorporated into cellulosic materials.
Protocols based on saponification were established to fractionate the suberin components from poplar bark as well as fatty molecules and long hydrocarbon chains from the leaves of the tobacco tree *N. glauca*. Yields were determined for each type of extract and compositional analyses were performed on the fractions from poplar that contained the highest suberin amount and on the fraction from tobacco leaves that was richest in fatty

acids and derivatives. The data revealed the presence of valuable bioactive triterpenoids in the bark of poplar trees. In addition, fatty acids, fatty acid alcohols and diacids were identified in the fractionated samples. These can potentially be used for the development of 'green' esters and polyesters. Their exploitation for materials design and fabrication has, however, not been possible due to the low yield obtained for each component at the scales tested in the project. The tobacco tree sample was composed essentially of linoleic acid, either as a free fatty acid or released from esters upon saponification. These samples are thus ideal for biofuel production.

Another approach, aimed at altering the cell wall ultrastructure of tobacco cells, was the overexpression of a carbohydrate-binding protein named CBM3 in cell culture. This protein binds to cellulose and was thus expected to alter cellulose crystallinity in mature cells. Surprisingly, the structure of the cellulose was altered, but without any apparent decrease in crystallinity. Cellulose nanocrystals prepared by acid hydrolysis of cellulose isolated from the CBM3 cell line were significantly longer than those from the non-modified line (Figure 2). Interestingly, the use of the intact nanofbrils from the CBM3 cells allowed the fabrication of nanopapers that exhibited enhanced mechanical performance. This demonstration suggests that *in planta* expression of CBM3 would lead to similar modifications of cell walls as observed in cell suspension cultures and allow the large scale production of strong nanopaper derived from cellulose nanofibrils from, e.g., secondary cell walls from modified poplar trees.

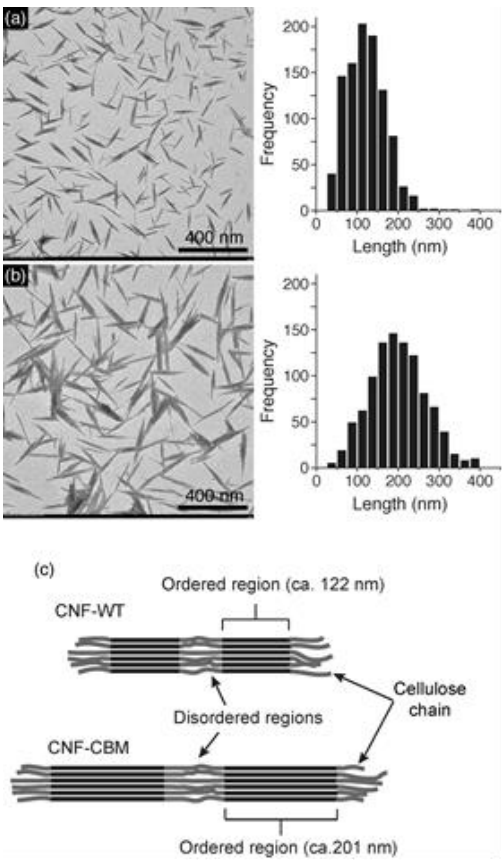

Figure 2: Micrographs (electron microscopy) and corresponding histograms of length distribution are shown for cellulose nanocrystals hydrolyzed from the wild-type (a) and CBM3 (b) lines. (c) Schematic illustration of a possible distribution of ordered regions along the corresponding cellulose nanofibrils.

In addition, by studying the effect of the different genetic engineering approaches on mechanical properties, structure, and chemistry of the poplar wood, we were able to gain new insights into structure-function relationships of plant cell wall mechanics.

Mechanical properties were investigated by micromechanical tensile tests of xylem tissue strips and the density of these strips was determined. Specific elastic moduli and specific strength values (density normalized) were compared to eliminate the influence of changes in density on the mechanical properties. X-ray diffraction was employed to elucidate the orientation distribution of the cellulose microfibrils in the cell walls and has a dominant influence on the mechanical behavior. Finally, changes in the chemistry of the cell wall due to the modification were resolved with Raman and Infrared spectroscopy. For the latter, spectroscopic analysis was performed on the mechanically tested samples to directly correlate to changes in mechanics with the changes in density, structure and chemistry.

The downregulation of the specific enzymes of lignin biosynthesis led to different effects regarding mechanical properties and chemistry of the plants. Whereas in one case, no difference in mechanical behavior was visible, two other treatments led to significant decreases in mechanical properties. For one genetic modification, the decrease in elastic modulus could be correlated with the decrease in lignin content as no other significant

changes in tissue density, cell wall structure and chemistry occurred. Thus, for some treatments it was possible to demonstrate the feasibility of genetic engineering for gaining new insights into structure-function relationships of plant cell walls. However, both wild type and transgenic poplar plants showed a pronounced intra-sample and intra-line variability in mechanical properties, which can be assigned to the natural variability in structure and chemistry of the cells and cell walls independent of the genetic treatment of the plants. This variability makes it difficult to resolve the underlying structure-function relationships as it may partly obscure the actual effects of a genetic treatment, especially in cases of more subtle changes and probably explains the difference between these results and those previously reported.

The effect of the genetic engineering on the mechanical properties can also be taken as a valuable input for estimating the (mechanical) robustness of the plants in field plantations. However, this effect may vary with the age of plants and also with the impact of environmental factors such as wind or precipitation on morphology and mechanical properties of the plants. Thus, there might be a difference between greenhouse plants and field grown plants. Such differences need to be taken into account when arguing about tree robustness.

The concept of hybrid materials often found in nature, combining constituents of different material classes in highly ordered arrangements of organic and inorganic phases, yields outstanding multifunctional properties. Bone and nacre are probably the most prominent and best studied biological materials, wherein a combination of different material classes arranged in a highly sophisticated manner comprises material functionality i.e. an outstanding mechanical performance and adaption to additional functions. Another example for a highly sophisticated hierarchical structure fulfilling various functions and yielding high mechanical integrity is wood. The hierarchical structure of wood renders it highly suitable to be used as a carrier scaffold for the formation of light-weight, multifunctional materials (Figure 3).

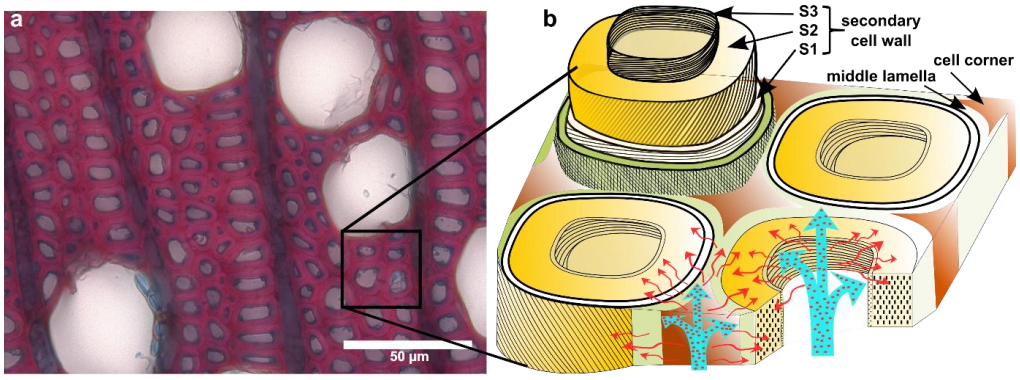

Figure 3: a) Optical micrograph of cross-section of poplar wood. b) Cell wall model with potential diffusion pathways of aqueous media. The secondary cell wall with distinct layers, S1, S2, S3, differentiated through chemical composition and microfibril angle.

The design, fabrication and assessment of the properties of cellulose based composites were followed for spruce, used as a model system, and for the genetically modified poplar wood. Former studies have shown that the integration of additional functionality into wood based cellulose scaffolds for new material profiles can be facilitated by increasing the porosity of the cell walls. Therefore, one can profit from genetic treatments which reduce lignin contents in the plant or make the lignin more soluble in chemical processes.

A main criterion for the feasibility of pretreatment processes in this work was the preservation of the hierarchical arrangement of the cellulose present in the native wood cell wall, combined with a profound removal of the lignin. Transgenic plants were used to study whether the genetic treatment can facilitate further chemical processing to achieve cellulosic scaffolds. Standard procedures from the pulp and paper industry for a removal of the lignin were applied to increase the porosity in the cell wall and allow for an easier accessibility of the remaining scaffolds. FT-IR spectroscopy and X-ray diffraction techniques, combined with spatially resolved Raman spectroscopy were conducted in order to monitor lignin removal and possible changes in the cellulose structure. The latter, especially, allowed for a comprehensive analysis of the chemical changes introduced through delignification processes, with clear distinction to the different anatomical regions present in the wood cell wall.

Depending on the pretreatment of the wood, the porosity of the modified cell wall and the free volume available for newly added material could be influenced. *In-situ* mineralization and metallization treatments as well as infiltration of pre-synthesized nanoparticles were conducted to functionalize the wood and the cellulosic scaffolds. For the post-functionalization of the wood-based scaffolds through *in-situ* formation of nanoparticles, we focused on classical water-based and microwave-assisted synthesis routes. Electron and optical microscopy of the scaffold backfilled with *in-situ* formed nanoparticles indicated the structural differences of the obtained cellulose scaffolds in dependence of the applied pretreatment and confirmed the preservation of the integrity of the hierarchical, anisotropic structure. This general materials approach may lead to the implementation of a multitude of sustainable material combinations resulting in manifold potential applications.

In summary the objectives of WP5 were fully realized. Importantly, the suberin for wild type and transgenic plants was characterized and genetic modification of cell wall properties was demonstrated to exhibit beneficial characteristics. Moreover, it was possible to demonstrate modified functionality of poplar wood which offers legion possibilities of commercial application.

Conclusions

- Protocols based on saponification were established to fractionate the suberin components from poplar bark and long chain hydrocarbons from *N. glauca* leaves and suitable substrates for the development of “green polyesters” were isolated. However, their abundance was too low to allow facile exploitation for materials design and fabrication.
- Genetically modified poplar plants were demonstrated to have considerably different mechanical properties and expression of a carbohydrate binding protein in tobacco cell lines also resulted in an unexpected enhanced mechanical performance with both phenomena having potential industrial applications.
- New fundamental insights into structure-function relationships of plant cell wall mechanics were obtained.
- Modification of cell wall functionality via pre-treatments of wood including in situ mineralisation and metalisation demonstrated that modified functionality of poplar wood offers a massive scope for commercial application.

Work package 6: Evaluating biorefinery measures for the selected multipurpose crops

WP6 Objectives

O6.1. Utilisation of *N. glauca* and poplar as new feedstocks for biorefining and novel biomaterials.

O6.2. Production of lead (bio-)chemicals and fibers for materials development.

O6.3. Recovery of compounds other than hydrocarbons and fibers as additional added-value products.

The main activity in WP6 was the design of advanced biorefinery processes wherein specific classes of molecules with the highest potential for biomaterial engineering and non-food oil production were isolated in a pilot plant scale. The work was performed for *N. glauca* and poplar bark. The final processes developed for both feedstocks are described below. Importantly, both were optimized for high fuel production and in maximising recovery of value added compounds – in the case of *N. glauca* solanesol and in the case of poplar suberin.

Biorefinery for *N. glauca* leaves

The biorefinery process began with the harvesting and cleaning of the *Nicotiana glauca* farm site. The leaves were dried to prevent them from rotting. After shipment to the production site, the biorefinery process steps took place. These consisted of several extraction steps and the biodiesel production step (see Figure 4). The first extraction used 96 % ethanol as

solvent. This separated waxes, oily compounds and sugars. The resultant ethanol-miscella was cooled to crystalize and separate waxes as valuable product. The remaining miscella was evaporated to separate the ethanol and then extracted with hexane to gain the oily compounds. The second extraction step for the leaves after ethanolic extraction used hexane as solvent to gain a second oily fraction from the leaves that was enriched with additional valuable compounds like sterols and carotenoids. The last extraction step was the hexane extraction of extract 2 to separate oily compounds as valuable products. This oil fraction was used to produce the fatty acid methyl esters (FAME). After the FAME was produced from the oil, it then went through two separations, where the mixture was cooled, causing certain fractions to crystallise. The focus was on converting all of the FAME into biojet fuel through hydro-deoxygenation and hydrocracking.

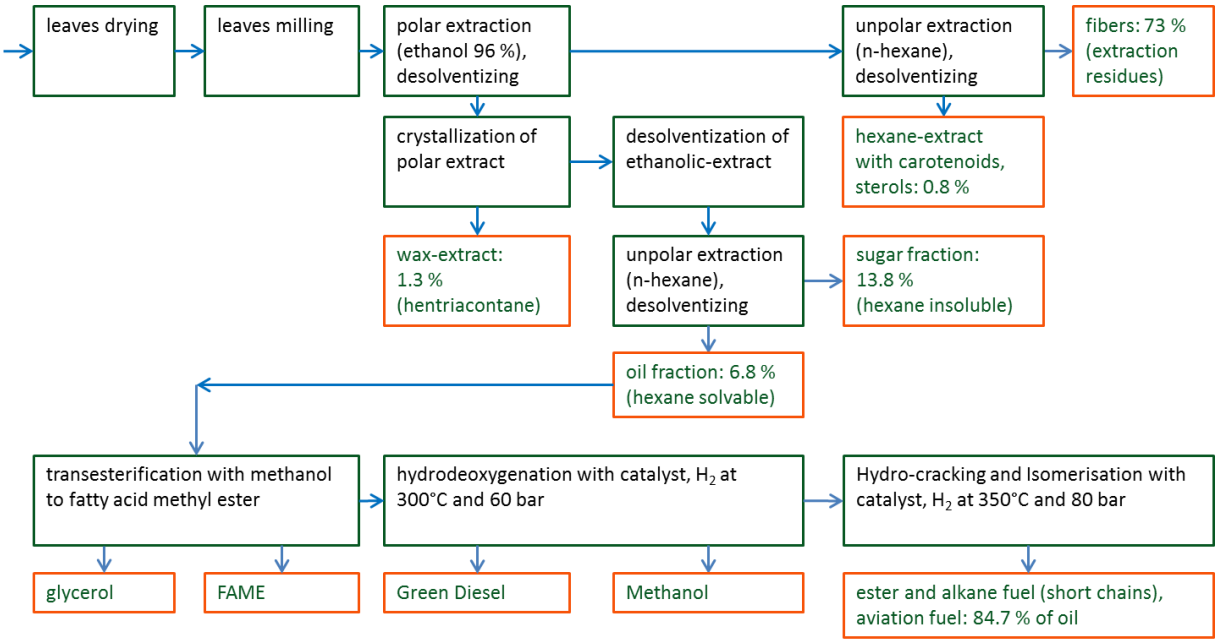

Figure 4: Biorefinery process for *N. glauca*

Biorefinery for poplar bark

In the biorefinery of poplar, the wood was processed to produce cellulosic materials. Like for *N. glauca* the process began at the farm site. The poplars needed to be harvested and then debarked. In the MultiBioPro project, further process steps of the bark as by-product were considered, but not the well-established processing of the wood. The bark was broken and cleaned of stones. It was then dried to prevent the bark from rotting. After shipment to the production site, the biorefinery process took place (see Figure 5). For the first extraction step, process water was used as solvent. Sugars were removed at this stage. The second step was saponification with KOH dissolved in methanol. After acidification the last extraction step was the extraction of the suberin-product using supercritical carbon dioxide.

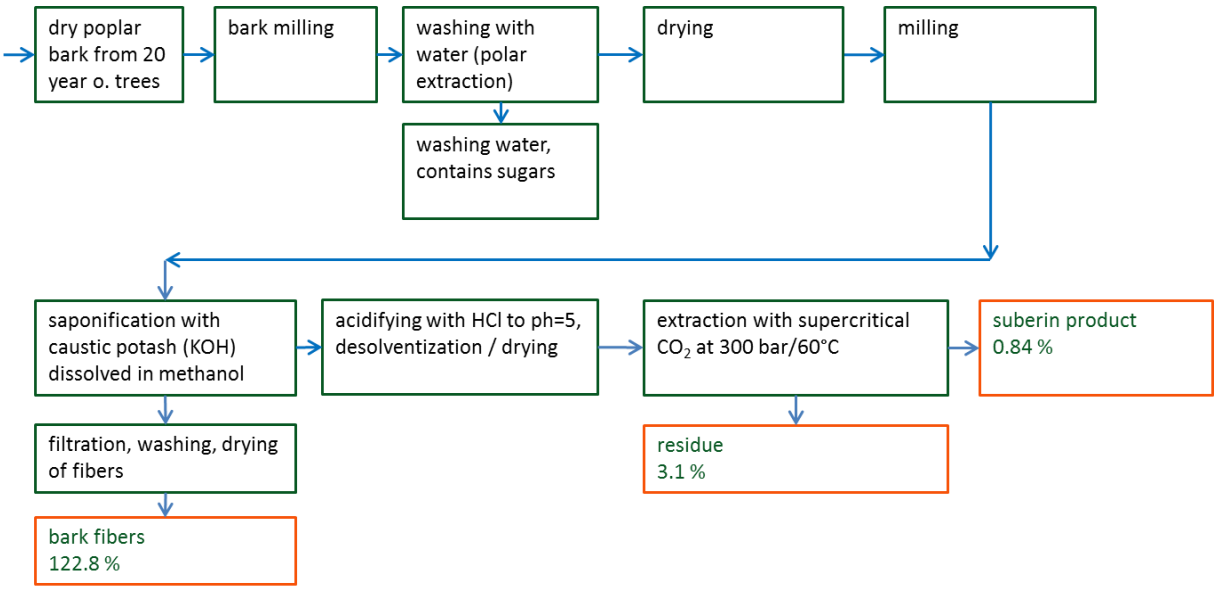

Figure 5: Biorefinery process for poplar bark

Conclusions

- Both laboratory and pilot plant scale extractions and processing of *N. glauca* leaves and poplar bark were established which provide both good fuel yields and considerable production of the identified exploitable by-products namely solanesol and suberin.
- Planning for full scale biorefineries have been carried out including equipment lists and mass balances models. In addition, a commercial scale plant has been designed and modified for jet fuel and a respective market analysis and business model has been designed for this application.

Work Package 7. Agronomic and environmental assessment of MultiBioPro varieties

WP7 objectives

O7.1. Determine the agronomic properties and perform environmental assessment of poplar and *N. glauca* under field trial conditions performed in multiple geographical locations.

O7.2. To determine agronomic parameters and environmental effects of second generation MultiBioPro polar and tobacco genotypes with enhanced biomass extractability and oil content.

O7.3. Generate the experimental datasets to enable the assessment of the marketability and implementation of business strategies suitable to deliver MultiBioPro technologies into commercial practice.

The main objective of WP7 was to conduct growth trials of the selected poplar and tree tobacco varieties. Field trials are essential in assessing the agronomic characteristics of

experimental crops under field conditions. Within the MultiBioPro project, the field trials were carried out for the following key objectives:

i. to facilitate the comparison of agronomic and environmental performance among different geographical locations and under different environmental conditions
ii. to assess the potential of modern breeding technologies
iii. to generate data for developing a market assessment of the best strategies to deliver the varieties and processes to be developed into commercial practice

A series of growth trials were conducted in Belgium for Short Rotation Coppice (SRC) Poplar and in the UAE and Madagascar for wild-type *Nicotiana glauca*. *N. glauca* was also observed growing under wild conditions on Ascension Island, and biomass harvested there was sent to the UK for extraction and analysis.

The trials assessed a number of traits including resistance to pest infestation, germination rates, ease of propagation methods, growth rates under different treatments, biomass production, irrigation requirements and fertilizer application. Due to regulatory restrictions, transgenic *N. glauca* was not included in the field trials.

Belgium

A new poplar trial in Wetteren, Belgium was established with transgenic poplars downregulated for cinnamyl alcohol dehydrogenase (CAD), which was documented to have improved saccharification yields in greenhouse trials. 240 copies of three transgenic lines and the wild type control were planted in 2014 and harvested in 2015 with material currently under analysis for cell wall composition and saccharification potential. The field grown lines had higher saccharification potential. A field trial with cinammyl CoA reductase (CCR) downregulated poplars in Zwijnaarde, Belgium was completed after seven years of growth. The entire site was excavated to destroy the root systems and monitoring for additional root suckers is planned for the next year. Results from these field trials confirmed higher saccharification yield but only in a single line and only after acid or strong alkali treatments.

Ascension Island

Ascension is a small volcanic island located in the Atlantic Ocean. Its flora consists entirely of species introduced through human intervention. *N. glauca* grows prolifically all over the island. Climatic conditions on Ascension are tropical and semi-arid. The MultiBioPro Consortium obtained 60kg of *N. glauca* biomass that was harvested from the island for extraction and analysis of compounds in WP4.

UAE trial

Generally, the climate in the UAE, including at Sharjah where the trial took place, is hot and dry, with extreme temperatures in excess of 40^{o}C reached during the months of July and August. Precipitation rates are very low, with no rain recorded for 5 -6 months of the year.

The trial site was cleared and demarcated into 2 zones; one half with net shading, the other without. An irrigation system consisting of tank, pipes and watering loops was installed. This enabled relatively accurate metering of irrigation water. The plants survived for the 10 months from September to June under open field conditions.

The plants thrived in the UAE, with rapid and robust growth observed throughout the growing season. The unshaded plants outperformed their shaded counterparts on stalk girth measurements, with the reverse observed for average height.

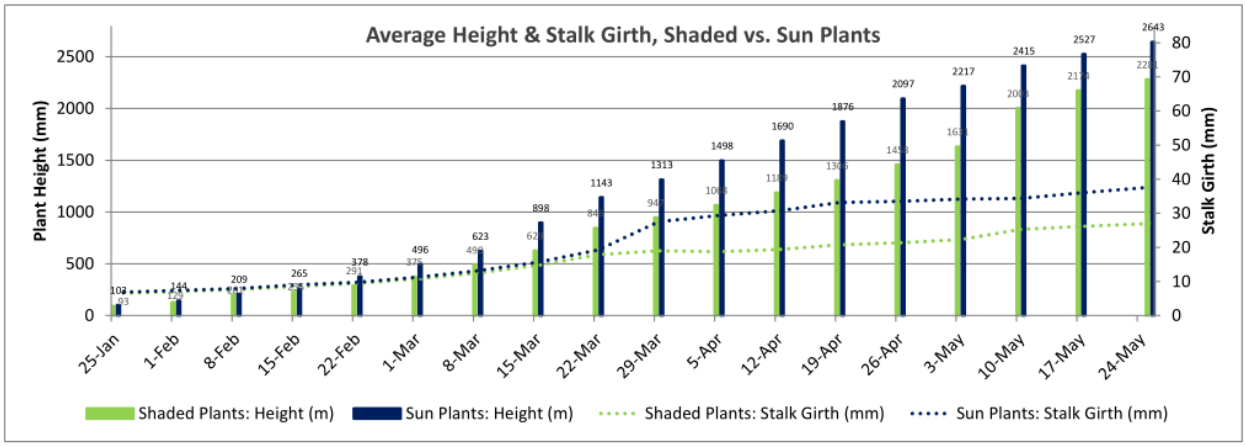

Figure 6: Physical measurements of UAE trial plants

The plants were irrigated intensively throughout the duration of the trial, with four distinct regimes established (as shown in Table 1). The plants responded most favorably to the high intensity irrigation regime in terms of observable vigor and biomass produced.

Table 1: Irrigation regimes

Intensity	**Rate (litres/m²/day)**
Low	3.5
Medium	7
High	10.5
Very high	13

Mealybug and aphid colonisation occurred in the plants, with the highest levels observed in the young specimens. Ant population was also high, perhaps because of their attraction to the sticky exudates produced by the mealybugs. Treatment was achieved by means of *Mospilan SL* applied for 5 days. This was effective at eliminating the pests, but caused some leaf yellowing. Fungal infestation level was minimal.

Table 2: Physical characteristics of wild type *N. glauca* from UAE trial

Av. Height (cm)	Av.Leaf Width (cm)	Av. Leaf Length (cm)	No.of plants harvested	Leaf mass weight (kg)	Leaf to Stalk ratio
146	7	12	17	8.5	0.8
149	7.5	12.1	19	9	0.8
162	8.3	12.5	16	6.1	0.81
133	7.1	11.6	13	3	1.54
160	8.8	12.7	15	9.3	0.81
136	7.7	11.9	15	6.5	0.69
171	6	10.3	12	5.6	0.79
131	7.9	12	11	5.7	0.82

Madagascar trial

Madagascar's climate is generally defined into two distinct seasons; a hot and humid rainy season which lasts from November to April, and a cooler dry one from May to October. Due to its topography and relative position to trade winds, there is significant variation in weather over the island, with rainfall ranging from 1.5 meters to as much as 3.5 meters annually.

861 seedlings were planted out in total during the first phase. There was a high incidence of fungal infestation among the young plants, which, coupled with seasonally high solar irradiance, caused a high mortality rate among the young plants in the field. The fungal infestation was controlled with Azadiracta applied on a twice-weekly basis.

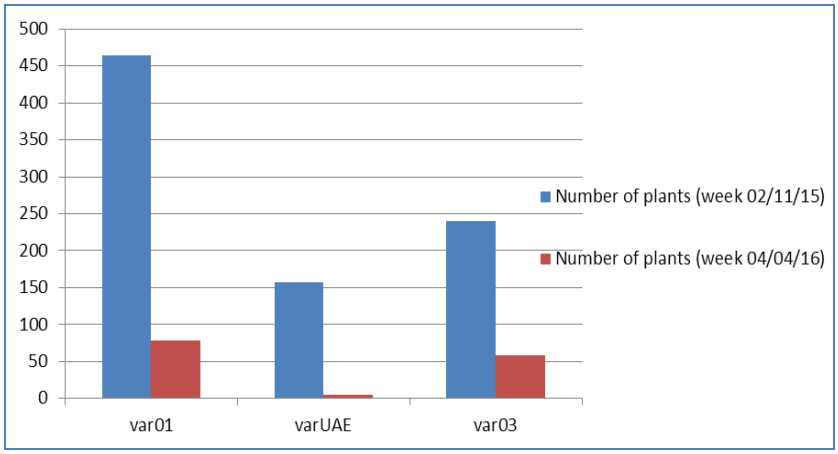

Figure 7: Number of live specimens of each variety in week 2/11/2015 compared to week 4/4/2016

Across harvests the average weight of dried leaves per plant per harvest was 22g. This equates to 880kg/ha, assuming 40,000 plants/ha. For comparison, the tobacco variety Burley in Madagascar had an average yield of around 25g/plant, equivalent to 1t/ha also at 40,000 plants/ha.

N. glauca is susceptible to fungal and aphid infestation when grown in subtropical locations, although these can be successfully treated with fungicides and pesticides. The metabolite content in harvested leaves diminishes with drying and time, meaning that the distance from plantation to processing facility must be minimal.

In summary, both crops could be grown and produce substantial yields in marginal growth conditions as well as to provide suitable feedstock for the project. For *N. glauca,* an essential first characterization of the crop for this purpose was achieved whereas for poplar comparative work on transgenic lines was possible.

Conclusions

- Whilst the agronomic performance of poplar is well established, the MultiBioPro project confirmed greenhouse results suggesting that gene modified poplar exhibited properties rendering them superior in terms of biofuel production.
- The *N. glauca* field trial evaluations carried out in the MultiBioPro project demonstrate that this crop can generate considerable biomass in both arid and even extreme desert conditions
- Data obtained from these field trials will inform future best strategies for exploitation of these crops for production of biofuels and value added bioproducts on marginal land.

Final considerations

The MultiBioPro project delivered considerable knowledge concerning multiple facets of the production of multi-use biofuels on marginal land. As such it is difficult to address the entire scope of the project, however, salient points which summarize the major outputs are provided below.

MultiBioPro used state-of-the-art approaches and technologies to advance our fundamental understanding of key genetic and metabolic processes underpinning the performance of multipurpose crops. Prime amongst these were the adoption and further development of methods to translate information gleaned from Arabidopsis co-expression studies to other species and the clarification of the biosynthetic pathways of lignin, flavonoids and glycoalkaloids and phytosterols in the model species Arabidopsis and tomato. In addition, considerable datasets have been collected which encompass the transcriptional and metabolic complements of multiple tissue types of *N. glauca* in order to complement work by the *N. glauca* genome sequencing initiative which includes several partners of the MultiBioPro project.

Transfer of the abovementioned knowledge by the MultiBioPro consortium from Arabidopsis to *N. glauca* and poplar proved highly efficient with *N. glauca* generated that contained increased phytosterols thus providing a valuable renewable resource of these compounds.

Current studies are evaluating the possibility of biotransforming these sterols into commercially important steroids. Evaluation of transgenic poplar plants downregulated in *CCR* and *CAD* within the MultiBioPro project revealed that these exhibited significant alteration in lignin structure and a corresponding enhancement in saccharification yields. Thus for both crops the focus on translation of knowledge gleaned from Arabidopsis was highly successful.

Detailed chemical compositional analysis of *N. glauca* and poplar samples including characterization of both small molecule metabolites and polymers as well as analysing the saccharification potential allowed a comprehensive survey for potential added value compounds which highlighted suberin, and solanesol as exceptional target compounds for this purpose. Discussions between the consortium and the chemical industry are already in progress.

Genetic modification of lignin content and consequently of cell wall properties was demonstrated to exhibit beneficial characteristics with regard to biofuel production. Additionally, the novel biopolymers produced by these lines have interesting properties and as such are likely to have a range of potential commercial applications.

Biorefinery of both *N. glauca* and poplar was performed and they were proven to be effective feedstocks not only in terms of fuel output but also in terms of value added compounds being recovered at appreciable amounts from the fractions obtained during the refinery process.

Equally as important as the demonstrated efficacy of the feedstocks was the fact that we were able to obtain enough biomass both for poplar (which was expected) and for *N. glauca* on marginal land. This was certainly achieved with *N. glauca* field sites in the United Arab Emirates and Madagascar and the poplar sites in Belgium. The field data being similar for both crops to that previously obtained in greenhouse conditions.

4. Description of the potential impact

The outputs from the MultiBioPro project will impact directly on several key strategic areas including improved sustainable bio-production and processing, enhanced quality of life, important socio-economic factors, the knowledge based economy, increased competitiveness and prosperity, international development and advancing fundamental science and the exploitation of knowledge. Examples are provided below.

1. Scientific advances and knowledge exploitation. Important underlying fundamental knowledge can be acquired from the project`s output. This information will likely benefit the broader scientific community and substantially contribute to the progression of science by opening up new research avenues in addition to furthering prospective societal and

commercial exploitation. When taken together, the research carried out in the MultiBioPro project improves upon the state-of-the-art in the following ways:

(i) MultiBioPro improved the fundamental knowledge of plant primary and secondary metabolic pathways of plants and their regulation for the production of plant derived commodities. The isolation, identification and characterization of genes encoding newly discovered biosynthetic enzymes involved in the biosynthesis of lignin and or phenylacetylated flavonoids represent prime examples of how the utilization of systems level profiling of transcripts and metabolites can be employed as important discovery tools. As such, MultiBioPro advanced fundamental knowledge of plant metabolism in general, in fact it did so to such an extent that the textbook pathway of the biosynthesis of the essential biopolymer lignin will need to be modified. The application of the same tools in tomato also suggested that there were several errors in the previously suggested pathways of glycoalkaloid metabolism that have now been corrected and incremental advances in our knowledge of flavonoid and lipid metabolism in N. glauca have also been made.

(ii) Development and utilization of innovative tools which facilitate the leverage of knowledge from model to crop species formed an important cornerstone of the MultiBioPro project. Their utility for understanding metabolism in our models was demonstrated by the successful identification of genes in Arabidopsis and tobacco which qualify as central enzymes of lignin and phytosterol metabolism in poplar and N. glauca, respectively.

(iii) The MultiBioPro project collated detailed transcriptomic and metabolomic datasets associated with biosynthesis and deposition of lipids in *N. glauca* and via outreach to other scientists in both the academic and industrial sectors working with *N. glauca* to integrate this with the genome sequence. In doing so, a highly complete molecular characterization of this species was achieved which will prove essential for further work with both fundamental and applied perspectives.

(iv) The project delivered the tools and expertise necessary to develop a network-based approach as a discovery pipeline for optimization of lipid production/secretion and improving chemical composition for biomass extractability and biomaterial characteristics. Several suitable genes were characterized; however, vectors for combinatorial transformation did not prove effective. This fact notwithstanding, clear effective metabolic engineering strategies have been developed and validated for the improvement of biofuel recovery.

(v) Poplar and tobacco tree genotypes with combined biomass, oil and organic properties were generated. The MultiBioPro project did not assess fuel production in isolation but rather in the context of overall biomass production and in the case of *N. glauca* also in terms of its properties for bioremediation. The purpose of these integrative studies was to identify interventions which maximize oil yields in a manner which is neutral or even beneficial to the environment and which allows biofuel production on land that is regarded marginal for food

production. Greenhouse trials suggested that this was achievable despite the fact that the poplar interventions resulted in reduced biomass production (due to the high importance of lignin for plant growth), whilst *N. glauca* was demonstrated to be able to tolerate growth in water impoverished and heavy metal contaminated soil with little effect on biomass production. The lignin compromised poplars were additionally demonstrated to have considerably higher saccharification yields.

(vi) New transfection technologies/procedures were developed. Through activities in the MultiBioPro project, *N. glauca* transformation is now a routine laboratory practice.

(vii) The application of transcripts and metabolomics was improved and extended. Over the course of the MultiBioPro project, the very latest developments in transcriptomics and metabolomics were implemented into the work program. The utility of RNA-Seq was demonstrated through the identification of transcripts associated with the lead compounds targeted and was also extensively used in annotation of the *N. glauca* genome. The availability of these RNA-Seq datasets to the community will likely be vast since they are of high quality and therefore can directly be used in the co-expression pipeline generated during the MultiBioPro project. Furthermore, comparison of RNA-Seq datasets between *N. glauca* and tobacco will likely be highly informative in understanding the woody stem trait and as such the regulation of bark deposition. Similarly, accurate mass metabolomics approaches have been developed and used to identify and assign putative formula to unknowns especially in poplar. Furthermore, a battery of metabolomics approaches was utilized to provide gold standard structural identifications of the phenylacetylated Arabidopisis flavonols, named saiginols. It is likely, that these could readily be introduced into several crop species to confer added stress tolerance. Thus, the MultiBioPro project has further developed both metabolomics and transcriptomic technologies.

(viii) A repository of meta and experimental datasets relating both to the metabolite content of poplar and *N. glauca* and to the optimisation or perturbation of metabolic pathways for the production of biorefinable compounds from these species has been produced. Large datasets of high quality were generated in the MultiBioPro project. These can be utilized by the scientific community in the future. To facilitate community use, the data will be made fully available via appropriate databases.

(ix) Detailed lignin, cell wall and suberin analysis on specifically engineered *N. glauca* and transgenic poplar was performed. The structural, surface and mechanical properties of wood/fibers from transgenic poplar were assessed, whilst the production of modified polyesters based on long chain hydrocarbons of *N. glauca* and refined poplar bark suberin was evaluated. The resultant information will further allow the design and fabrication of cellulose based polymers of novel functionality and a myriad of opportunities for the commercialization of these novel polymers will be followed.

(x) A catalogue of biorefinery strategies of both MultiBioPro crops was compiled which details both the fuel and value added components at each process at a qualitative level with additional quantitative data available for the most important compounds. These studies demonstrated that both poplar and *N. glauca* performed similar to the industry standard maize within the biorefinery process. They additionally identified solanesol in *N. glauca* and suberin from poplar bark as highly valuable by-products of the biorefinery. First contact has been made with the chemical industry regarding these compounds. With regard to biofuel, a commercial scale biorefinery has been established which is intended to use *N. glauca* seeds for the production of aviation fuel. Biorefinery of transgenic *N. glauca* revealed that recovery of the enhanced phytosterols following processing was possible.

(xi) Within MultiBioPro extensive field and wild growth trials were carried out. In the case of *N. glauca* these were restricted to wild type material. The material was collected from the wild growth trials in Ascension Island, a small volcanic island in the Atlantic Ocean, and from field trial sites in the extreme desert of the United Arab Emirates and the highly arid Madagascar. A full agronomic assessment of the yield was carried out in controlled field experiments. The data will shortly be made available in conjunction with publication of the genome. Given the success of the biorefinery trials it is likely to be of high use for informing policy regarding biofuel production as well as for agronomists and academics interested in plant growth in extreme environments. The poplar trials, carried out in two different locations in Belgium, included transgenic material which had been demonstrated in greenhouse trials to have improved extractability and therefore enhanced saccharification yields. These trials demonstrated that the results could by-and-large be reproduced in the field – thus demonstrating the efficacy and applicability of the strategies adopted. Assessment of the marketability of both *N. glauca* and poplar within commercial strategies should be completed within the next two years.

(xii) A supportive environment for sharing knowledge (know-how), information, expertise and innovation (knowledge transfer) was created. MultiBioPro guaranteed knowledge exchange not just through project specific activities but also in association with the concurrent COST ACTIONs and projects e.g. MultiHemp, Quality Fruit and Fruit Engine and through outreach to Multi-national companies such as Philip Morris International and Pierre Fabre which strengthen both academic and industrial horizons of the project. This aspect was of particular benefit to the SMEs giving them direct access to developments in fundamental research for example benefiting greatly from the exchange of chemical analytical expertise. The MultiBioPro project will clearly advance our fundamental scientific knowledge and highlight potential exploitation of multipurpose crops grown on marginal land.

2. Improved sustainable bioproduction. The MultiBioPro project has demonstrated the success of an approach where fuel and other value added compounds are produced by two crops *N glauca* and poplar that grow on marginal land and as such do not compete with food and feed crops. Moreover, *N. glauca* has been demonstrated to be able to produce considerable biomass even when grown in the desert. Therefore they represent high-yielding renewable bioresources at competitive levels. Furthermore, the field trials with transgenic poplar demonstrate how effective containment of the transgenes is ensured via excavation of the field sites and subsequent monitoring. Whilst a considerable amount of field trials have been carried out, feasibility studies are now required to demonstrate this on a large production scale. The technologies and resources generated by the proposed project are generic and can potentially be exploited to impact directly on creating sustainable renewable biofuels and for most high-value plant natural products. The new genotypes created during the MultiBioPro project can now provide us with this opportunity to show the feasibility of even improving on the potential of these crops. In MultiBioPro, we have identified solanesol and suberin as high value compounds that can be generated as biorefinery by-products to create added-value co-products. To recapitulate, the outputs of the MultiBioPro project offer a means of producing biofuel products using cheap renewable bio-sources that have a dramatically reduced impact on the environment and do not compete with the production of food and feed. These are important factors that will impact directly on society in the coming years, particularly given the twin problems of the exploding world population and deteriorating environment.

3. Increasing competitiveness and prosperity. The tools, resources, plant varieties and training generated and provided by the MultiBioPro project will contribute to ensuring European enterprises working in plant agricultural biotechnology and related fields are competitive. In this way the potential has been developed for European industry to compete with their US and Asian counterparts or envisaged to license technologies to these established and emerging global partners. It is hoped, that the renewable and cheaper means of production developed in the MultiBioPro project will facilitate expansion of existing, and the creation of new, markets and companies, increasing European prosperity. For example, the high-value demonstration molecules targeted in the project offer an opportunity to upgrade plant based raw materials to high value co-products. Such approaches offer an important alternative market for European agriculture at a time where subsidized production of bulk products is being considerably reduced. Given increased energy costs, environmental implications as well as the fossil fuel based production systems, the market for biofuel production coupled to high-value natural products like those targeted in the present project are likely to expand rapidly. In addition such products have a high export value and given the

choice of vehicles of the project can be grown on marginal land in a non-competitive manner to food and feed.

4. Socio-economic impact. In 2050 it is estimated that the human population will exceed 10 billion people. Meeting the nutritional and energy demandwill have important economic implications across Europe and the world. Producing renewable energy in a manner that is not competitive to food and feed production such as growth of crops capable of growing on marginal land, like *N. glauca* and poplar targeted in the MultiBioPro project, and in a way that additionally optimizes the by-production of value added compounds will thus represent a mechanism by which to reduce costs. The utilisation of renewable bio-resources as source for chemicals will also enable the consumer to have a natural choice instead of a reliance on the synthetic market. In Europe, the present chemical synthesis mode of production is often linked to the petro-chemical industry for the supply of precursor molecules and thus subject to intense price fluctuations. The use of plant based production platforms would improve self-sustainability and promote independence of volatile markets. Thus, through its dual objectives of identifying genotypes capable of providing fuel and value added compounds the outputs from MultiBioPro will have direct economic implications in the coming years. The project has demonstrated how crops grown on marginal land can replace fossil fuels and the chemical industry with regard to meeting society's energy and chemical needs**.**

5. Training. One main objective of the MultiBioPro project was to equip the young scientists recruited in the project with a comprehensive expertise at the interface of the academic and industrial sectors. Through WP9, the trainees have acquired state-of-the-art knowledge in their specific areas of research as well as in other disciplines through exchanges between the different academic laboratories. Exchanges between academic and industrial partners also took place (in both directions) in several instances to ensure that the trainees gain inter-sectorial experience and awareness of industrial needs in a general context of plant biotechnology research and development. Access to data and sharing of protocols between the partners and their trainees was provided during the entire duration of the project through a specifically dedicated intranet platform, accessible via the website (www.multibiopro.eu). The platform was also used for project reporting, to capture specific dissemination tools, background information and templates. A calendar function was used for the coordination of meetings and archiving of the presentations held during the workshops. A series of workshops were organized in conjunction with the annual meetings where selected partner members and external speakers delivered lectures in aspects that were most relevant to the project, including, for example, plant biology, genetics and breeding, plant biochemistry and metabolic engineering, bioinformatics and systems biology, biorefinery processes and

chemical analyses, agronomy and field trials, and biomaterials design, structure analysis and property assessment as well as life cycle analysis. Expert skills in specific research activities and technologies were provided to the trainees through a 'training-on-the-job' approach. This comprised in particular the acquisition of top-level research strategies at each partner laboratory, approaches and specialised technologies, as well as troubleshooting approaches, data handling and storage, ethical issues, and other activities relevant to research in a laboratory setting (work in a group, coordination of activity, shared instruments, etc.). In order to facilitate the implementation of the training-on-the-job approaches, each trainee was assisted by his/her mentors at each partner laboratory to develop a Personal Career Developmental Plan at the beginning of the project. Mentorship was maintained throughout the entire training period at each partner laboratory to ensure that the training objectives would be successfully reached.

As a result, a new generation of scientists has been trained and many of the trainees remain employed at the host laboratories whereas PhD students are in a phase of finalising their theses at the time of writing this report. It is expected that the high-level of training provided in discipline-specific activities through MultiBioPro will greatly facilitate their future career development in the public or private sector. There is indeed an increased demand for R&D experts in 'green' chemistry in the broad sense to overcome our unsustainable dependence on fossil fuels as a source of chemicals, materials and biofuels with the early career scientists employed by the MultiBioPro project being ideally poised to fill, at least in part, this demand.

6. Rural development. The MultiBioPro project has provided opportunities for rural development within and beyond Europe (e.g. Belgium, United Arab Emirates, and Madagascar). Expertise developed has in each case remained local. This approach is in line with the policy of involving rural areas into the globalization process by ensuring an active economic role.

7. Improved consumer quality and safety. The MultiBioPro project has delivered the tools and resource necessary to provide the consumer with the choice of a non-food competitive biofuel over fossil fuel and of natural (by-)product(s) over synthetic products derived from the chemical industry. There is also potential for a wider consumer choice, greater availability and cheaper costs. The utilisation of bio-resources instead of chemical synthesis means safety aspects are generally improved. There is less chance of chemical contamination so the work setting is improved for employees and the process platform will have less impact on the environment. In addition, preliminary experiments suggest that *N. glauca* is extremely tolerant to growth on heavy metal contaminated soils potentially hinting to its ability to aid in

phytoremediation. The fact that the MultiBioPro project already carried out field and biorefinery based studies means that we have already been able to evaluate these aspects in some detail; however, ultimately further feasibility studies will be required to fully complete such assessments.

8. Industrial policy and employment. The scientific and technological advances will potentially impact on European competitiveness creating new and increased markets leading to increased economic growth and job creation. Industrial partners including SMEs are actively involved in the program fostering industrial cooperation and capacity building at a European level. Due to their participation in the MultiBioPro project, partners were able to create and/or safeguard employment opportunities.

9. European cohesion. The MultiBioPro project has through the transfer of technology and joint efforts of selected experts with complementary skills and industrial cooperation created European cohesion and capacity building at the European level. In addition, further links have been established which has resulted in MultiBioPro activities contributing to capability building in developing countries.

10. Contribution to policy developments. MultiBioPro partners have presented and contributed data to both National agencies and the European Commission. These events contribute to policy especially in the case of reducing reliance on fossil fuels and the promotion of biorefining procedures.

11. Ethical issues. Participants in MultiBioPro have at all times complied with current legislation and regulations in the countries where the research has been carried out. We have performed ethical reviews on the use of genetic manipulation periodically throughout the project.

12. Contributions to standards. MultiBioPro implemented procedures to ensure accuracy, traceability and reliability of the data and procedures, to health and safety standards, and to staff development and training. The consortium worked to community standards relating to the reporting of “omic” datasets. The contextual information or documentation (“metadata”) is essential for the project partners and future secondary users and thus the consortium agreement will ensure access to all datasets and accompanying metadata for the project participants. After publication and intellectual property rights having been secured, which we predict to be within three years after the completion of the project, the metadata and datasets

will be made publicly available through third party databases. All outputs from the scientific activities will adhere to the principles of good laboratory practice.

13. Dissemination. The MultiBioPro project has disseminated its findings and activities through a number of verifiable formats for example, scientific papers submitted to journals and presented at conferences, through press releases, through meetings organized and/or attended by partners, public engagement including local meetings, and open days, project movies, and partner websites, patents, PhD theses and through the project website.

Thus, the dissemination activities have been extensive. Three different examples will be highlighted;

(i) Dissemination to the scientific community: MultiBioPro has already generated 21 peer reviewed scientific publications in high impact journals such as PNAS, Nature Communications, and Science. We anticipate more publications of this quality to follow. Presentations have also been given at many high profile conferences such as the international Solanaceae conference in Bordeaux 2015 and the 20th International Conference for Renewable Resources and Plant Biotechnology 2016.

(ii) Interaction with industry: MultiBioPro outputs have been disseminated to industry through individual seminars and at large conferences. It is through the latter activity that a potential route to market was forged and that the *N. glauca* genome consortium was formed.

(iii) Dissemination to the general public: The website has acted as a focal point for interested parties from the general public. It informed visitors about project contents and participants, main objectives, as well as project results. Through news releases and linked publications, project advancements were displayed. Information about the project was largely accessible to a broad public and since its launch, the MultiBioPro domain was accessed about 103.000 times. The project website will be maintained also after the project end to display further developments as applicable.

The MultiBioPro Consortium:

Max Planck Institute of Molecular Plant Physiology

Royal Holloway, University of London

KTH Royal Institute of Technology

Eidgenössische Technische Hochschule Zürich

Vlaams Instituut voor Biotechnologie

European Research and Project Office GmbH

Pilot Pflanzenöltechnologie Magdeburg e.V.

Green Fuels

Neutral Supply Chain

Capax Environmental Services