Bacterial Hosts for Production of
Bioactive Phenolics from Berry Fruits

BACHBERRY Project Grant Agreement n°613793

Final Report

25.12.2016

About BACHBERRY project

BacHBerry is a 3-year project funded by the EC-FP7 on "BACterial Hosts for production of Bioactive phenolics from bERRY fruits". The BacHBerry project aims to develop a portfolio of sustainable methodologies to mine the potential of the untapped biodiversity of the bioactive phenolic compounds in an extensive collection of berry species. The consortium comprises a full chain of research and innovation, with 18 partners from research groups, SMEs and a large enterprise, representing 11 countries.

www.bachberry.eu.

Project coordinator

Project partners

About this document

This report corresponds to the Final Report of the BACHBERRY project. It has been prepared by:

Organization name: DTU-Biosustain, Technical University of Denmark
Organization address: Kemitorvet, Building 220, 2800, Kongens Lyngby, Denmark
Person's name: Alexey Dudnik
E-mail: adud@biosustain.dtu.dk
Phone: +45 93 511 101

This project is co-funded by the European Union within the 7th Frame Programme. Grant Agreement n°613793. The sole responsibility of this publication lies with the author. The European Union is not responsible for any use that may be made of the information contained therein.

Table of contents

1 Final publishable summary

1.1 Executive summary

BACterial Hosts for production of Bioactive phenolics from bERRY fruits (BacHBerry) was a project aiming at establishing a sustainable and economically-feasible pipeline for the microbial production of novel added-value polyphenols isolated from berry fruits.

As the first step, berries from a total of 112 species/varieties from various genera (*Aristotelia, Berberis, Lonicera, Lycium, Ribes, Rubus, Ugni* and *Vaccicium*) were collected from different locations in Chile, China, Portugal, Russia and UK. These samples were extracted and subsequently analyzed using LC-MS-based metabolomics tools. In order to store and share the data, a new database was developed, allowing the user to access the germplasm and metabolite information.

The 34 most phytochemically-diverse extracts were selected for bioactivity screening using the yeast-based SMART platforms, and the 3 most promising extracts were further fractionated. Each generated fractions was tested for bioactivity, and the positive ones were chemically characterized, resulting in a large library of potentially-bioactive compounds (~189). Out of those, four compounds were identified as the candidate effectors associated with bioactivities for Amyotrophic lateral sclerosis, Huntington's disease, inflammation, or Parkinson's disease. Additionally, colorogenic properties of selected berry extracts were evaluated in food products.

For isolation of candidate genes responsible for production of the identified bioactive compounds, transcriptome of 13 berry species spanning eight genera from seven families were sequenced and annotated. More than 6,000 candidate genes encoding enzymes for the phenolic biosynthetic pathway and its regulation have been identified. Enzyme functionality has been confirmed for a variety of berry genes and genetic markers for anthocyanin production have been identified.

Corynebacterium glutamicum and *Lactococcus lactis* were selected as host organisms due to their long history of industrial use and the safety status. Here, *C.*

glutamicum has been engineered for producing a variety of polyphenols, such as stilbenes (resveratrol, pinosylvin, pterostilbene) and flavonoids (naringenin, eriodictyol, kaempferol, quercetin). The key to this success was the identification of the hitherto unknown pathway for phenylpropanoid degradation and its subsequent deletion, along with 17 other genes. Furthermore, production pathways for resveratrol and anthocyanins were successfully introduced into *L. lactis*. The developed strains were further used for optimization of culture conditions and bioprocess scale-up in fermenters. Resveratrol production using *C. glutamicum* was chosen as the model process to scale-up using 250L fermenters. These tests demonstrated the technical viability of a potential industrial process.

In order to assist with strain engineering and improve polyphenol production, a number of original mathematical models and computational methods were developed, dealing with either a single strain or microbial consortia. These have been applied to data provided by project partners, leading to some interesting outcomes, some of which were experimentally validated.

Downstream processing was optimized using both liquid-liquid and solid-liquid extraction-based processes. This was done using detailed thermodynamic models that were developed to describe/predict partitioning of polyphenols based on their structure. Furthermore, techno-economic evaluation of potential process for the microbial production of polyphenols was performed.

Moreover, economic, regulatory, and sustainability aspects of the project were also investigated, and the best practices for Access and Benefit Sharing (ABS) were outlined. Lastly, stakeholder workshops were organized, interviews with consortium scientists and other interested parties were conducted, as well as a documentary film and an online science game were made.

1.2 Summary description of the project context and the main objectives

BacHBerry (BACterial Hosts for production of Bioactive phenolics from bERRY fruits) has received funding from the European Commission's the Seventh Framework Programme (FP7) under the theme KBBE.2013.3.1-01: "*Plant High Value Products - from discovery to final product*". The project team consisted of experts from the fields of plant biology, industrial biotechnology, analytical chemistry, and social sciences in order to assemble and validate a complete process for discovery and bacterial production of novel phenolic compounds originating from berry species. The consortium included twelve research groups, five small and medium sized enterprises and one large enterprise from eleven different countries, namely Austria, Chile, China, Denmark, France, Germany, Netherlands Portugal, Russia, Switzerland, and United Kingdom. Having partners from a multitude of countries and geographic regions allowed getting access to a variety of endogenous plants that might not yet been extensively characterized.

Historically, plants have not only been important food and nutritional source in a diet of every human, but also the foundation and the most important resource for traditional and modern medicine. The metabolite diversity in the Plantae kingdom has been estimated to be vast, and include up to 200.000 different chemical compounds (Weckwerth 2003), with many of those compounds having proven or potential medical applications. This corresponds to the high number of pharmaceutical products that are either based on, or derived from, plant natural products, such as morphine, quinine, paclitaxel and artemisinin (Cragg and Newman 2013) . Among the different classes of plant metabolites, polyphenols stand out due to their large diversity and ubiquity in the plant kingdom. Over the last two decades, plant polyphenols have been gathering more and more attention due to their powerful antioxidant properties (Stevenson and Hurst 2007). Furthermore, polyphenols were demonstrated to have anti-inflammatory and anti-cancer properties, and are increasingly being associated with putative bioactivities for several cardiovascular

and neurological diseases (Santangelo et al. 2007; Kim et al. 2010; Kishimoto et al. 2013; Goszcz et al. 2015).

Polyphenols are a lot more prevalent and diverse in berry species than in other common fruits and vegetables. For example, cranberry, lingonberry and blackcurrant were demonstrated to have a high flavonoid content (Häkkinen et al. 1999), and anthocyanins, one of the most abundant groups of plant pigments, are also enriched in berries (2-5 $g.kg^{-1}$ fresh weight) (Määttä et al. 2001; Grotewold 2006). Additionally, berries constitute one of the most important dietary sources of ellagitannins (Landete 2011) and proanthocyanidins (Hellström et al. 2009). Thus, it becomes obvious that berry species contain a broad diversity of different polyphenols and their derivatives, and tapping into this diversity should allow discovery of novel health-beneficial compounds.

Moreover, the polyphenol market has been steadily going up over the past few years, and is expected to exceed 850 million USD by 2018 worldwide (Aranaz et al. 2016). This has been mainly due to accumulating scientific evidence regarding benefits of polyphenol consumption that resulted in increased consumer awareness on this matter. Additionally, there has been a boost in the usage of polyphenol-containing extracts in beverages, food, and cosmetics products, in particular in the Asian region (Aranaz et al. 2016; Grand View Research Inc. 2016).

It is obvious that in order for the increasing demand to be met, there is a need for more sustainable and eco-friendly manufacturing approaches, as current polyphenol production mostly relies on extraction from various plant sources (e.g. roots, leaves, or fruits) via complex downstream processing (Wang et al. 2016). However, the utilized extraction procedures have numerous shortcomings that limit the potential exploitation of polyphenols (Kolewe et al. 2008; Wang et al. 2016). Hence the consortium has decided to address these challenges by setting the following objectives:

(i) to systematically analyze the phenolic composition of the extensive germplasm and berry collection available to the consortium members,

including species of the following genera: *Rubus*, *Ribes*, *Vaccinnium*, *Lycium*, *Lonicera*, *Berberis*, and *Ugni*.

(ii) to develop a publicly-available resource of genomic and metabolic data obtained from berry bioprospecting within, as well as outside of the project, in order to have a centralized resource highlighting the chemodiversity potential of active compounds in berries

(iii) to discover novel bioactivities in berry extracts and to identify the corresponding functional biomolecules, conferring protection against a multitude of human pathologies, such as Alzheimer's disease and Amyotrophic lateral sclerosis, by high-throughput screening using the SMART platform

(iv) To perform functional characterization of the genes and the corresponding gene products involved in the biosynthesis of the high-value phenolic compounds identified through bioprospecting (core biosynthetic enzymes, decorating enzymes, such as glycosyltransferases, acetyltransferases and P450s, as well as transcriptional regulators)

(v) to evaluate a selection of the identified biosynthetic genes for functionality in Gram-positive bacterial hosts and subsequently use those to construct bacterial cell factories for the production of phenolic compounds

(vi) to further improve the production efficiency in the engineered bacterial cell factories by functional integration of heterologous pathways, modifying the host metabolic networks and performing flux balancing using predictions obtained via rational design or computational tools developed within the frame of the project

(vii) to design and optimize cost-effective methods for extraction of phenolic compounds from bacterial fermentation broth, achieving purities fulfilling the requirements of the European regulation 231/2012 for food and directive 2001/20/EC for pharmaceuticals

(viii) to optimize fermentation conditions and subsequently upscale the production to the pilot plant levels (250L). The functionality of phenolics produced by the designed bacteria through the optimized fermentation

process is to be validated in the SMART yeast platforms and further verified in mammalian cell lines

References

Aranaz I, Acosta N, Mengíbar MN, et al (2016) InFiQuS: Making the Best of Leftovers. In: Domínguez de María P (ed) Industrial biorenewables : a practical viewpoint. John Wiley & Sons, Inc., Hoboken, US, pp 363–364

Cragg GM, Newman DJ (2013) Natural products: A continuing source of novel drug leads. Biochim Biophys Acta - Gen Subj 1830:3670–3695. doi: 10.1016/j.bbagen.2013.02.008

Goszcz K, Deakin SJ, Duthie GG, et al (2015) Antioxidants in Cardiovascular Therapy: Panacea or False Hope? Front Cardiovasc Med 2:29. doi: 10.3389/fcvm.2015.00029

Grand View Research Inc. (2016) Polyphenols Market Analysis By Product (Grape Seed, Green Tea, Apple), By Application (Functional Food, Functional Beverages, Dietary Supplements) And Segment Forecasts To 2024.

Grotewold E (2006) The genetics and biochemistry of floral pigments. Annu Rev Plant Biol 57:761–780. doi: 10.1146/annurev.arplant.57.032905.105248

Hellström JK, Törrönen AR, Mattila PH (2009) Proanthocyanidins in Common Food Products of Plant Origin. J Agric Food Chem 57:7899–7906. doi: 10.1021/jf901434d

Häkkinen SH, Kärenlampi SO, Heinonen IM, et al (1999) Content of the Flavonols Quercetin, Myricetin, and Kaempferol in 25 Edible Berries. J Agric Food Chem 47:2274–2279. doi: 10.1021/JF9811065

Kim J, Lee HJ, Lee KW (2010) Naturally occurring phytochemicals for the prevention of Alzheimer's disease. J Neurochem 112:1415–1430. doi: 10.1111/j.1471-4159.2009.06562.x

Kishimoto Y, Tani M, Kondo K (2013) Pleiotropic preventive effects of dietary

polyphenols in cardiovascular diseases. Eur J Clin Nutr 67:532–535. doi: 10.1038/ejcn.2013.29

Kolewe ME, Gaurav V, Roberts SC (2008) Pharmaceutically Active Natural Product Synthesis and Supply via Plant Cell Culture Technology. Mol Pharm 5:243–256.

Landete JM (2011) Ellagitannins, ellagic acid and their derived metabolites: A review about source, metabolism, functions and health. Food Res Int 44:1150–1160. doi: 10.1016/j.foodres.2011.04.027

Määttä KR, Kamal-Eldin A, Törrönen R (2001) Phenolic Compounds in Berries of Black, Red, Green, and White Currants (Ribes sp.). Antioxid Redox Signal 3:981–993. doi: 10.1089/152308601317203521

Santangelo C, Varì R, Scazzocchio B, et al (2007) Polyphenols, intracellular signalling and inflammation. Ann dell'Istituto Super di sanità 43:394–405.

Stevenson DE, Hurst RD (2007) Polyphenolic phytochemicals – just antioxidants or much more? Cell Mol Life Sci 64:2900–2916. doi: 10.1007/s00018-007-7237-1

Wang J, Guleria S, Koffas MA, Yan Y (2016) Microbial production of value-added nutraceuticals. Curr Opin Biotechnol 37:97–104. doi: 10.1016/j.copbio.2015.11.003

Weckwerth W (2003) Metabolomics in systems biology. Annu Rev Plant Biol 54:669–689. doi: 10.1146/annurev.arplant.54.031902.135014

1.3 The main S & T results/foregrounds

1.3.1 WP3: Phytochemical mining and databasing of diverse germplasm

Objectives and tasks

The main goal of WP3 was to prepare and collect all the chemical information necessary to WP4 (see below). More specifically, the objectives include the collation of global data sources to create a database of phenolic diversity, as well as organization and overseeing of berry sampling, Standard Operating Procedures (SOP)-based extraction and shipping for bioprospecting within WP4, and chemical characterization, preparation, and fractionation of the obtained extracts.

Task 3.1: *Construction of a knowledge database for structured and germplasm targeted bioprospecting*

Global scientific data sources were mined for traditional-to-modern knowledge on functionality and bioactivity in order to expand the existing databases of berry phytochemistry and bioactivity. These were structured to facilitate searches for specific bioactivities and chemical classes in species within the following genera; *Rubus, Ribes, Vaccinnium, Lycium, Lonicera, Berberis,* and *Ugni*. The database was constructed to take the data generated from the collation, fractionation and bioactivity characterization (WP4) data. Following a period of closed access for addressing IP issues, an open access policy will be applied, at least for the data attributed to germplasm and chemical characterization.

Task 3.2: *Access to, and preparation of, biodiverse global germplasm and the generation of material for screening*

The established germplasm collections of IBET, JHI, PUC, IBCAS & VIR was utilized for phytochemical diversity assessment and bioactivity screening. These germplasm collections contain varieties, landraces and wild accessions derived from the UK, mainland Europe, Russia, Chile, New Zealand, North America and China, and berries of the following genera, *Rubus, Ribes, Vaccinnium, Lycium, Lonicera, Berberis,* and

Ugni. Material Transfer Agreements was completed and submitted for collation, to ensure that all activities are performed in accordance to Nagoya protocol.

Task 3.3: *The application of state-of-the-art purification and metabolomics approaches to bioactive and functional molecule characterization.*

A standard operating procedure (SOP) was developed for preparing berry extracts, and circulated to all the partners that will be providing germplasm for screening. The SOP, based on freeze-dried berries, generated polar and nonpolar extracts that were prepared according to standardized protocols developed in other EU projects. Validated un-targeted LC-MS profiling was employed to characterize the extracts and fractions. For each chemical class (anthocyanin, ellagitannins procyanidin, flavonol, etc.) quantification was done using the available commercial compounds, *eg.* cyanidin-3-*O*-glucoside, ellagic acid etc. All data was deposited in the project database.

Task 3.4: *Sub-fractionation and compound purification*

Following the results of the bioprospecting and color functionality screening procedures (WP4), further fractionation of the extracts was undertaken using semi-preparative HPLC. Definitive compound characterization was assured using accurate mass (Orbitrap) MS and MS/MS fragmentation. Confirmation of structure was done using analogues or scaffolds of authentic compounds. Compounds were purified to pharma grade purity for further biotesting, as well as for identification of the corresponding biosynthetic pathways (WP5) to be used for microbial production.

Summary of results

D3.1 - Material transfer agreements for the transfer of germplasm extracts and, where necessary, biological material

A material transfer agreement has been drafted in accordance with the Nagoya protocol and reviewed by the legal departments of both JHI and DTU. This has been subsequently approved and signed by the consortium partners.

D3.2 - Standard operating procedures (SOPs) for the extraction of the target germplasm

In order to ensure that extraction procedures were standardized across all of the different germplasm collection sites and that sufficient quality control methods were in place, a new method was successfully developed providing non-living material in the form of extracts, which ensured that this project adheres fully to the Nagoya protocol and protects IP associated with the shared germplasm. Furthermore, the extracted material was found to contain adequate quantities of the types of metabolites being studied (*i.e.* polyphenols), with verified stability of more than 3 weeks at -20 ˚C storage temperature. Furthermore, the method was improved by using an internal standard, and the amount added optimized to allow adequate quality control of the extraction procedure. This method was subsequently adapted for the generation of bioactivity screening extracts by omitting the internal standard from the procedure. The developed chromatography and mass spectrometry methods were found to be adequate for the analysis of targeted classes of compounds (*i.e.* anthocyanins, flavonols, etc.).

D3.3 – Report on the current knowledge associated with phenolics bioactives from the target species

A deliverable report has been submitted that summarizes a body of literature regarding the phenolic diversity present in the diverse studied genera, the contribution of these to food quality, and the effects of polyphenols found in these genera on human health. It has been noted that the genera studied within BacHBerry displayed a great diversity in terms of the classes of polyphenols present, the aglycone forms present within each polyphenolic class of interested, and the decoration diversity of individual aglycones, suggesting a great potential for bioprospection in the assembled germplasm collection. In addition to contributing to some organoleptic properties of food, several studies have linked the consumption of

polyphenol-rich foods with beneficial health effects. In particular, the report focused on reviewing the evidence of the effects of polyphenols on several diseases, such as cancer, diabetes, and cardiovascular and neurodegenerative diseases, citing studies ranging from *in vitro* and animal work to human intervention studies.

D3.4 – Germplasm extracts for preliminary screening

The developed SOP were used for the analysis of the germplasm collection of berry species from diverse genera originated from Chile, China, Portugal, Russia and Scotland. The total number of species/lines analyzed comprised 113 samples and included samples from the following genera: *Aristotelia* (1), *Berberis* (1), *Lycium* (1), *Lonicera* (10), *Ribes* (57), *Rubus* (35), *Vaccinium* (6) and *Ugni* (1). The complete germplasm set was extracted and analyzed in triplicate utilizing the LC-ToF-MS-based untargeted approach. An automated MS mining feature in the LC-MS software was used to generate 1506 and 384 features for the dataset in positive and negative modes, respectively. This dataset was then subjected to multivariate statistical analysis and the species/accessions that displayed the largest diversity were selected for the preliminary screening. As a result, a total of 34 species/accessions were selected for bioactivity screening and re-extracted using the SOP developed for this purpose. These extracts were subsequently sent for bioactivity screening to WP4. In addition, a color strength screening was performed for the entire dataset in order to select the berry species with the highest potential for food application purposes as colorants.

D3.5 - Fractionated extracts for further screening

This deliverable was directly aligned with the outputs of WP4. The workflow leading to this deliverable can be summarized as follows: first, an initial phytochemical analysis of the wide germplasm collection and the selection of the most chemical diverse species/lines were conducted. Subsequently, extracts of these species/lines were tested for bioactivity and antimicrobial activity, and the number of species was further narrowed to the species that demonstrated the highest potential. These species were then subjected to fractionation, the fractions were tested for their bioactivity and anti-microbial activity, and lastly, chemical analysis of the promising

fractions was performed. The initial results of the bioactivity screening have highlighted that the extracts of two *Rubus sp.* berries contained high bioactivity potential against a variety of diseases (Huntington's disease (HD), Amyotrophic lateral sclerosis (ALS), Parkinson's disease (PD), and inflammation). Fractionation methodologies were developed for each of the berry extracts and included a chromatographic separation using a C18 silica resin. Additionally, (and although not berries) leaf extracts of *Corema sp. 1* were included in the fractionation process due to its bioactivity for Parkinson's disease. This resulted in the generation of 13, 28 and 16 fractions for *Rubus sp.* 1, *Rubus sp.* 2 and *Corema sp. 1*, respectively, which were then passed on to WP4 for further bioactivity testing.

D3.6 - Report a complete project database housing all project data; germplasm, chemical and bioactivities

The data generated within the BacHBerry project consists of numerous variables, whose nature often differs significantly. For example, throughout the project, species information and site of collection were recorded in parallel to bioassay results, as well as metabolite characteristics and content. These corresponded to descriptive (species information), discrete data (bioassay results), and continuous data (metabolite content and characteristics such as $(M-H)^+$, $(M+Na)^+$, $(M-H_2O)^+$, etc.). Although, compiling these different types of data into a single dataset posed no significant challenge, providing a database that would allow the user to explore the different types of data according to their goals provided a significant challenge. In order to overcome this challenge, a novel database (Berrybase) was created based on the Germinate 3 platform, and custom-made to host the germplasm and metabolite data. Germinate 3 is a generic plant genetic resources database, created by JHI, and offers facilities to store both standard collection information and passport data, along with more advanced data types such as phenotypic, genotypic and field trial data. The Berrybase installation implements novel tools that allow the user to explore metabolomics-type datasets. These features focused on implementing a parallel metabolite database that includes metabolite information (name, molecular formula, monoisotopic mass, etc.), images and links to other relevant databases, while additionally implementing visualization tools that allow the user to explore the

diversity in compound abundance across several species. BerryBase is the first database of its kind and is already proving its worth in-house at JHI as the central repository to input all data for the JHI *Ribes, Rubus* and *Vaccinium* breeding programs, allowing a much faster and a more detailed control of the programs' outputs, whilst keeping record of the breeding rejections for later analysis. This completely supersedes the traditional Excel-based systems ubiquitous in almost all plant/crop breeding programs.

1.3.2 WP4: Bioprospecting and biodiscovery of novel berry high value phenolics

Objectives and tasks

The bioprospection of berry extracts from the samples harvested in Portugal, Scotland, Russia, Chile and China for the discovery of new added-value phenolic compounds with bioactivities for chronic human diseases (neurodegenerative and haematological diseases, cancer, inflammation, and type II diabetes), novel natural colorants, and compounds with anti-microbial activity, were the main objectives of WP4.

In **Task 4.1**, the bioprospection for health-beneficial effects of the most chemically-diverse extracts were successfully carried out using the SMART screening platform. Among several species conferring bioactivities for human diseases, the two *Rubus* species and *Corema sp. 1* that demonstrated bioactivities for pathological processes of ALS, HD, inflammation and PD, were the most promising samples that were selected for a bio-guided fractionation. After detailed chemical characterization of the fractions that retained bioactivity, four candidate compounds with activities for ALS, HD, inflammation or PD were identified and validated. Furthermore, additional evaluation of selected model compounds, representing different classes of polyphenols that were expected to be present in berries, allowed identification of a novel bioactivity for the flavonol fisetin. This will result in at least four publications that are currently in preparation.

In **Task 4.2**, functional screening for antimicrobial activities and colorogenic properties of selected extracts were performed using the antimicrobial activity platform and a color evaluation panel. *Corema sp. 1* and *Rubus sp.* 1 were identified as the samples with the highest antimicrobial activity, supporting the decision to select these samples for fractionation and further identification of bioactive compound. The antimicrobial activity of *Corema sp. 1* for *Staphylococcus aureus* was lost after the fractionation, suggesting a putative synergistic effect among the compounds present in the original extract. Regarding the color functional screening, *Rubus occidentalis* was pointed as the one presenting the most interesting color

characteristics. However, methanolic extracts used in this work were shown to be unsuitable for application in food products due to formation of insoluble particles that were visible in yoghurt.

In **Task 4.3**, resveratrol produced within the frame of BacHBerry was used as a proof of concept in order to evaluate the efficacy of polyphenols produced in food-grade microorganisms. For that, different sources of resveratrol (native from plant extraction - commercially available, produced by *Saccharomyces cerevisiae* (Evolva), or produced by *Corynebacterium glutamicum* (Juelich/Biotempo)) were tested in the inflammation and PD models of the yeast SMART screening platform. Even though resveratrol from *C. glutamicum* was less pure than the other two samples (~50%), it was still able to reduce activation of the reporter construct in the anti-inflammatory screening system as efficiently as the plant and the yeast resveratrol. Moreover, given the attractive features of anthocyanins as natural colorants, these compounds were produced in *Escherichia coli* and *Lactococcus lactis* and color properties of the corresponding extracts were determined and compared with elderberry and blackcurrant standards using the color evaluation panel. Both *E. coli* and *L. lactis* extracts had attractive red-orange and orange shades, respectively. However, the turbidity caused by the presence of water-insoluble components strongly limits their application, suggesting a need to develop better extraction procedures.

Summary of results

Task 4.1: *Bioactivity screening: Human pathologies*

One of the goals of BacHBerry was the identification of novel health-promoting bioactivities among the samples provided by iBET, JHI, PUC, and IBCAS. This was achieved by using the yeast SMART platform to identify polyphenol-enriched extracts with bioactive properties for several neurodegenerative disorders, inflammation and cancer, among the mostly chemically-diverse extracts indicated by chemical analysis done by WP3. A summary of the identified bioactivities is given in **Table 1**. *Rubus sp. 1, Rubus sp. 2 and Corema sp. 1* were subsequently chosen for the bio-guided

Table 1: Summary of the bioactivities identified using the “humanized” yeast SMART platform

Bioactivities are highlighted as pink squares.

Species	Disease models						Source of Extracts
	Parkinson's disease	Alzheimer's disease	Huntington's disease	Amyotrophic lateral sclerosis	Cancer	Inflammation	
Rubus vagabundus	–	–	–	–	–	–	Portugal (8)
Rubus brigantinus	–	–	–	–	–	+	
Rubus sampoianus	–	–	–	–	–	+	
Rubus henriquesii	–	–	+	–	–	–	
Rubus sp. 2	–	–	+	+	–	–	
Rubus hochstetterorum	–	+/–	–	–	–	–	
Corema sp. 1	+	–	–	–	+	–	
Corema sp. 2		–	+	–	–	–	
Rubus loganobaccus 1	–	–	–	–	–	–	UK (19)
Rubus sp. 13	–	–	+	–	–	+	
Rubus loganobaccus 2	–	–	+	–	+	+	
Ribes holosericeum	–	–	–	–	+	+	
Rubus sp. 1	–	–	+	+	+	+	
Rubus sp. 14	–	–	–	–	+	–	
Rubus fruticosus 1	–	–	–	–	–	+	
Rubus armeniacus	–	–	–	–	+	+	
Rubus occidentalis	–	–	–	–	+	+	
Rubus fruticosus 2	–	–	–	–	–	–	
Rubus ideaus	+	–	–	+	–	+	
Ribes nevadense	–	–	+	–	–	+	
Ribes grossularia	–	–	–	–	+	+	
Ribes rubrum	–	–	+	–	–	–	
Ribes sp. 644217	–	–	–	–	–	+	
Ribes petraeum	–	–	+	–	–	–	
Ribes sp. 2292-1	–	+	+	–	+	+	
Ribes sp. 1126	–	–	+	–	–	+	
Ribes bethmontii	–	–	–	–	+	+	
Berberies buxifolia	–	–	–	–	+	+	Chile (4)
Aristotelia chilensis	–	+	–	–	+	-	
Wild ribes	–	-	–	–	-	-	
Ugni molinae	–	+	+/–	–	+	+	
Vaccinium vitis-idaea	–	–	–	–	–	–	China (3)
Vaccinium uliginosum	–	–	–	–	+	–	
Lycium chinense	+	+	+	–	+	+	

fractionation procedure. After the fractionation, the bioactivity re-testing of the *Rubus sp.* 2 fractions in the ALS model allowed separation of the protective fractions. Comparative chemical analysis of the bioactive and non-protective adjacent fractions allowed the identification of several compounds with potential bioactivity for ALS. The most likely candidate compound was tested and its bioactivity was confirmed. Similarly, bioactivity screening of the fractionated *Rubus sp.* 1 extract revealed independent protective fractions for HD and for inflammation, and the respective candidate compounds for both diseases were identified. Moreover, their bioactivity was also validated in mammalian cell models. Given the novelty of the bioactive compound for HD, it has been decided to try to produce this compound in the engineered food-grade bacteria. Concerning the cancer model, bioactivity screening of *Rubus sp.* 1 fractions revealed a loss of the protective activity, suggesting that a synergistic effect of multiple *Rubus sp.* polyphenols may be responsible for the

bioactivity observed in the total extract. *Corema sp. 1* fractionation and bioactivity re-testing allowed identification of the positive fractions for PD and cancer. Chemical analysis of the bioactive fractions for PD resulted in a total of 97 annotated mass spectral features, and from those, a single candidate compounds was confirmed as the bioactivity determinant for this disease. Using statistical analysis tools, the entire *Rubus* dataset was used to integrate metabolomics and bioactivity data to identify additional putative bioactive compounds. This resulted in the generation of a list of potential hits for HD (3), inflammation (1), ALS (1) and cancer (8).

Task 4.2: *Functionality screening: color and antimicrobial activity*

The obtained results reinforce the idea that berry phenolics are potent antimicrobial agents. Nevertheless, identification of specific antimicrobial compounds was shown to be difficult due to synergistic effects among the compounds present in the original extracts. Therefore, even though we successfully identified berry extracts with antimicrobial properties, isolation of individual antimicrobial compounds was not possible. Berry species, including the *Rubus, Ribes, Vaccinium, Fragaria and Corema* genus were tested for their antibacterial properties against Gram-positive and Gram-negative pathogens. The berries extracts with the highest antimicrobial activity were the two tested *Rubus idaeus* cultivars, two *Ribes* species, the *V. vitis-idaea* extract and *Corema sp. 1* samples. Together with the results of the screening for health promoting bioactivities, these results supported the choice to select *Rubus sp.* 1*, Rubus sp.* 2 and *Corema sp. 1* for the bio-guided fractionation procedure.

Regarding the functional screening for colorogenic properties, evaluation of the suitability of the berry extracts revealed that assessment of color characteristics and stability in beverages could be easily performed on methanolic extracts, whereas use of acetronitrile resulted in presence of insoluble particles and turbidity once the extract is dissolved in a buffer. Among the different tested berry extract, the sample of *Rubus occidentalis* has been identified as the one presenting the most interesting color characteristics due to its high color strength and a bright orange shade. The samples of *Rubus henriquesis*, *Rubus loganobaccus*, *Rubus occidentalis*, *Rubus armeniacus* and *Ribes holosericeum* have been selected for stability test in beverage

application, presenting very close shades when applied in beverage. The extracts of *Rubus henriquesis*, *Rubus occidentalis* and *Rubus armeniacus* presented better stability than elderberry standard to light, but not to heat. In addition, *Rubus occidentalis* was shown to be the most stable sample, and for that reason it was tested as an ingredient for yogurt, revealing that methanolic extracts are not suitable for such tests due to the presence of spots of insoluble particles when added to yogurt. Lastly, evaluation of the juices from the engineered purple tomatoes showed that the indigo cultivar presents an interesting bright and violet shade. Unfortunately, the stability of anthocyanins from this source to light and heat is limited.

Task 4.3: *Bioactivity validation of bacterial produced phenolic compounds*

The testing of resveratrol from different sources (plant, yeast, or bacteria from BacHBerry) in the inflammation and PD models of the SMART platform revealed that all of the resveratrol samples were capable of attenuating Crz1 activation, therefore confirming that the bioactive compounds produced by microbial cell factories retain native bioactivity of the original plant compounds for attenuating inflammatory pathways. Furthermore, the resveratrol produced in *Saccharomyces cerevisiae* was shown to exert protective activities against inflammation at lower concentrations, as compared to the one commercially available (coming from native plant extraction). Resveratrol produced in bacteria was shown to be as efficient as the commercial ones in activating anti-inflammatory responses, proving the sustainability of production of plant phenolic compounds in food-grade bacteria. Concerning the ability to reduce α-Syn toxicity in the yeast PD model, both of the commercially-available resveratrol samples (purified from plants and produced *S. cerevisiae*) partially restored the growth defects associated with its toxicity. However the resveratrol sample produced in bacteria within BacHBerry did not overcome the α-Syn cytotoxicity, possibly due to low purity of the compound in the extract.

Concerning the anthocyanins produced in bacteria (see WP6 for details) and their use as food colorants, it can be concluded that in its current state, these extracts present low color strength and high color indexes. *E. coli* extracts contained cyanidin-3-glucoside and cyanidin as the major pigments, while in *L. lactis* extracts cyanidin

and another unknown pigment were identified as the major components. When applied to a soft drink, both bacterial extracts exhibited a high level of turbidity, which makes the extracts unsuitable for use as food ingredients due to the presence of water-insoluble compounds. In term of shade, *E. coli* extracts exhibited a bright red-orange shade, whereas *L. lactis* extracts had an orange shade. The color of the two bacterial extracts might be interesting, if there could be found a way to get rid of the insoluble particles. In conclusion, in their current state, bacterial extracts are not usable as natural food color for several reasons: (i) use of methanol for extraction (extraction with acidified water is recommended); (ii) low color strength and low anthocyanin content; (iii) presence of water insoluble pigments, which leads to high turbidity when applied to beverages.

1.3.3 WP5: Identification of metabolic pathways and regulators in phenolic production

Objectives and tasks

The Work Package 5 (WP5) represents the combined effort of three research groups located at the University of Copenhagen (UCPH), Evolva and the John Innes Centre (JIC). The main objective of WP5 was to provide the knowledge required about the biosynthetic genes and genetic regulation of phenolics production in plants to assist WP6 with bioengineering the production of bioactive phenolic compounds in bacteria. Specific objectives included the provision of experimentally validated plant gene sequences needed for pathway engineering in bacteria, as well the large-scale mining of genes involved in the phenolic biosynthetic pathway, its regulation and the decoration of its products in target berry species. Finally, WP5 aimed to develop genetic markers to assist breeders to produce berry species with enhanced bioactive levels.

Six main tasks were set-up to achieve these objectives:

Task 5.1: identify and disseminate gene sequences of key enzymes from the core phenolic biosynthetic pathways in plants to enable BacHBerry partners to *de novo* bioengineer the core phenolic pathways in bacteria.

Task 5.2: this task involved sequencing the transcriptome of a wide range of berry species, including those identified by WP3&4 as containing bioactive compounds. Identify, clone and test the activity of genes encoding decorating enzymes of phenolic compounds identified as bioactive by WP3&4.

Task 5.3: identify genes encoding regulators of phenolic biosynthesis in berry species identified by WP3&4 as containing bioactive compounds.

Task 5.4: clone and test the activity of genes encoding regulators of phenolic synthesis in berry species identified by WP3&4 as containing bioactive compounds.

Task 5.5: develop marker for different alleles of genes encoding regulators of phenolic synthesis in berry species identified by WP3&4 as containing bioactive compounds.

Most of the activities of WP5 were conditional to the availability of plant berry material or on the identification of bioactive compounds in berry species by other WPs. The WP5 activity, therefore, took place as these underpinning resources or information were made available during the course of the BacHBerry project. In sum, all deliverables and milestones contributing to the tasks and objectives of WP5 have been successfully achieved and completed.

Summary of results

Task 5.1: *Provision of genes encoding enzymes for the different phenolic biosynthetic pathways.*

13 key plant enzymes from the core phenolic biosynthetic pathway were identified (black boxes in **Figure 1**) and the sequences of experimentally validated genes from different species corresponding to these 13 enzymes were disseminated to other WPs.

Task 5.2: *Identification of genes encoding decorating enzymes*.
As berry plant material became available during the course of the project, the transcriptome sequences of 13 berry species spanning eight genera, seven families and seven orders were produced, annotated and disseminated to partners *via* a dedicated site developed at JIC. This represents the largest portfolio of berry transcriptome sequences available worldwide. In addition, the transcriptomes of berries at different stages of maturation (green, red and ripe) were sequenced for the two target *Rubus* species identified by WP3&4 as containing bioactive molecules (see above).

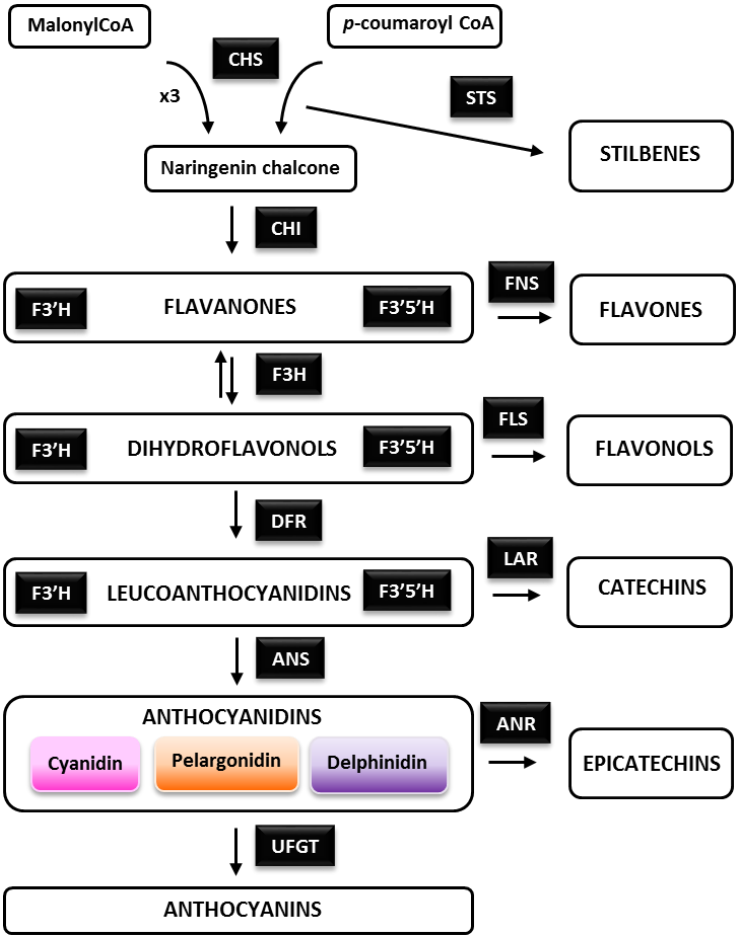

Figure 1: Polyphenol biosynthetic pathway in plants (modified from Falcone Ferreyra *et al.* 2012)

PAL, phenylalanine ammonia-lyase; C4H, cinnamic acid 4-hydroxylase; 4CL, 4-coumaric acid: CoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3-hydroxylase; F3'5'H, flavonoid 3'-5'-hydroxylase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase; ANR, anthocyanidin reductase; UFGT, flavonoid 3-O-glucosyltransferase; LAR, leucoanthocyanidin reductase; FNS, flavone synthase; FLS, flavonol synthase; STS, Stilbene synthase.

In plants, acyltransferases, glucosyltransferases, *O*-methyltransferases, hydroxylases, reductases, aurone synthases, dehydrogenases, dehydratases and dirigent proteins are the main types of decorating enzymes of polyphenolic compounds. 120 decorating enzymes from a range of plant species were identified by UCPH and JIC. These 120 sequences were used in BLAST searches against the transcriptome sequences of 13 berry species that we had previously generated. A total of 4,475 sequences homologous to decorating enzymes were identified. Multiple candidates for each type of decorating enzyme could be identified in each

transcriptome. On average, 19 acyltransferases, 96 glucosyltransferases (UGTs), 39 *O*-methyltransferases, 91 hydroxylases, 55 reductases, 6 aurone synthases, 16 dehydrogenases, 17 dehydratases and 2 dirigent proteins candidates were identified per species. This provided the project with a broad portfolio of decorating enzymes that could be mobilized for the bacterial production of target phenolic compounds by other WPs.

Table 2: Transcriptome sequence resources for berries developed on BacHBerry

Species	Plant material	Total number of raw reads
Aristotelia sp.	ripe berries	397,707,372
Berberis sp.	ripe berries	444,362,698
Corema sp. 1	leaf	353,604,932
Lonicera sp.	ripe berries	397,214,254
Ribes sp. 1	ripe berries	336,479,242
Ribes sp. 2	ripe berries	393,665,630
Rubus sp. 1	berries (3 stages)	1,040,224,680
Rubus sp. 2	berries (3 stages)	1,064,858,518
Rubus sp. 3	ripe berries	505,754,030
Rubus sp. 4	ripe berries	390,608,452
Ugni sp.	ripe berries	405,024,920
Vaccinium sp. 1	ripe berries	373,159,882
Vaccinium sp. 2	ripe berries	375,778,718

The testing of enzyme activity of candidate decorating genes followed up on the identification of bioactive compounds in berry species by other WPs which occurred during the second half of the BacHBerry project. Bioactive compounds for which a complete biosynthetic pathway has not yet been elucidated in plants, such as the compound RC-1, were not fully investigated, as the range of putative decorating enzymes was too large and too speculative. We, therefore, focused our work on enzymes associated with known bioactive phenolic compounds, such as, for example, UGTs expected to act on well-known model flavonoids (quercetin, pelargonidin, pterostilbene, fisetin). Candidate UGT genes were identified, cloned

and showed activity towards a range of target substrates (fisetin, pelargonidin, delphinidin and cyanidin).

Task 5.3: *Identification of genes encoding regulators of phenolic bioactive synthesis in berries***.**

As for **Task 5.2**, this activity followed up on the identification of bioactive compounds in berry species by other WPs, and focused on biosynthetic pathways already elucidated in plants. MYB, and basic-Helix-loop-Helix (bHLH) transcription factors, as well as WD-repeat (WDR) proteins can form a complex (MBW) and activate polyphenol / anthocyanin biosynthesis in plants. 68 regulators of the polyphenol biosynthetic pathway have been identified from a range of plant species by JIC. These 68 sequences were used in BLAST searches against the transcriptome sequences of 13 berry species that we had previously generated (**Task 5.2**). A total of 1,232 sequences homologous to the know transcriptional regulators were retrieved. On average, 85 MYB, 5 bHLH and 4 WDR candidates were identified per species. This work produced a large portfolio of sequences of regulatory genes that could be used to improve the production of target phenolic compounds and the development of markers for berry breeders.

Task 5.4: *Functional characterization of regulatory genes controlling bioactive accumulation.*

This activity followed up from task 5.3. RT-PCR products of a range of MYB and bHLH transcription factors (TFs) from *Rubus sp. 1* and *Rubus sp. 2* (identified by other WPs as target species containing bioactive compounds) were cloned into agrobacterial binary vectors for overexpression studies. Transient expression assays in leaves of tobacco plant *Nicotiana benthamiana* of MYB10- and AN1-type TFs led to the production of significant levels of anthocyanin (*i.e.* red coloration in infiltrated leaf patches; see **Figure 2**). This demonstrated that functional proteins (here, regulators of the polyphenol biosynthetic pathway) could be identified using the sequencing resources developed by WP5.

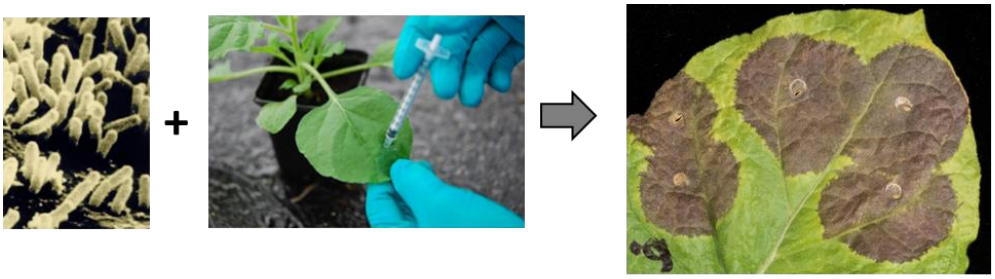

Figure 2: Overexpression of R2R3-MYB + bHLH transcriptional factors in *Nicotiana benthamiana* (tobacco) leaves using the *Agrobacterium tumefaciens*-based transient expression system

Task 5.5: *Development of markers for regulatory genes for selection for berry improvement*

During the activity tests of regulatory genes (**Task 5.4**), different isoforms of transcription factor-encoding genes exhibited different activity levels in terms of trigger the overexpression of anthocyanin in tobacco leaves following patch infiltration of overexpression vectors. Three PhAN1-like bHLH homologues cloned from *Rubus sp. 2* (RgAN1-1, RgAN1-2 & RgAN1-3) via RT-PCR showed different anthocyanin overexpression levels in transient studies. Nucleotide and amino sequence alignments of RgAN1-1, RgAN1-2 and RgAN1-3 revealed several variations that could be exploited for the development of molecular makers (*e.g.*, SNPs) to assist in breeding this trait in varieties to improve anthocyanin production.

1.3.4 WP6: Gram-positive bacteria as cell factories for the efficient production of berry high value phenolics

Objectives and tasks

Goal of this WP was the development of Gram-positive microbial cell factories for the efficient production of high-value phenolic compounds originating from berries. *Corynebacterium glutamicum* and *Lactococcus lactis* were made the organisms of choice for this task due to their Generally-regarded as safe (GRAS) status, high robustness, ease of genetic manipulation, and a long history of use in industrial processes. Production of polyphenols was achieved through the following tasks:

Task 6.1: *Assembly and grafting of pathways for synthesis of berry high-value phenolics*

This task aimed at constructing platform strains for production of core polyphenol molecules in the above-mentioned bacteria. This was done using gene sequences provided by WP5 and/or selected from the literature. It was split into two sub-tasks:

Task 6.1.1 Design of synthetic genes and operons

Translation efficiency of heterologous genes is affected by various parameters, including secondary structure of the transcript, codon usage, presence of slow/rare codons, etc. The target gene sequences were analyzed using various bioinformatics tools, and the identified features aiming at providing optimized expression in *L. lactis* and *C. glutamicum* were incorporated into the synthetic genes that were then obtained commercially. Additional collaboration with DNA2.0 was established in order to develop and test an improved algorithm for codon optimization for *L. lactis* and *C. glutamicum*. Furthermore, monocistronic and operon-like expression modules were designed using standardized parts (*eq.* promoter elements, ribosome binding site, terminators).

Task 6.1.2 Functional integration of heterologous pathways in bacterial hosts

Firstly, the effect of target phenolics on growth and global gene expression was evaluated and modifications were made based on the outcomes in order to make the strains suitable for production of phenolics. Next, biosynthetic pathways were reconstructed by assembling the corresponding genes via plasmid-based expression. Evaluation of biosynthetic genes from different sources, as well as of the gene order and the structural organization of the expression modules (monocistronic versus pseudo-operon versus operon) on the overall productivity was performed. Gene functionality was accessed by analyzing product formation using state-of-the-art analytical methods.

Task 6.2: *Host engineering for enhanced phenolics production*

Rational metabolic engineering approaches (*eq*. inactivation of competing pathways and (over)-expression of genes leading to accumulation of desired metabolites) was combined with model-guided approaches (developed by WP8) in order to re-wire the bacterial metabolic networks for enhancing phenolics production. Additional efforts were done for increasing the precursor availability of: i) the aromatic amino acid $_L$-tyrosine (*C. glutamicum*, Juelich), and ii) malonyl-CoA, which has been pinpointed as a major bottleneck in bioproduction of phenylpropanoid derivatives (DTU, Juelich). High-throughput screening methods (HTS, using colorimetric metabolite biosensors coupled to FACS analysis, T6.4) were deployed for selection of the best strains obtained by random mutagenesis or overexpression.

Task 6.3: *Balancing pathway expression for optimal yield and productivity*

Expression of the biosynthetic genes was modulated by varying the promoter strength and vector copy number. Expression of bottlenecks reaction steps was assessed by monitoring mRNA levels. Biosynthetic intermediate profiling experiments were also performed.

Task 6.4: *Development of expression platform for membrane-bound proteins*

Attempts to establish a platform for optimizing the expression of membrane bound enzymes of plant origin were performed in *L. lactis*. The system was based on fusion

of target cytochrome P450 genes with a folding reporter GFP gene. Bacterial surface display experiments were also done using an anthocyanin synthase gene.

Task 6.5: *Metabolite sensor development for intracellular detection of phenolics and their precursors*

Metabolite sensors for the intracellular detection of precursors of polyphenol biosynthesis were developed. Target metabolite-responsive transcriptional regulator/promoter combinations were identified from literature search or based on results from previous experiments, and the promoter was positioned in front of a fluorescent protein-coding gene (eq. GFP). Elevated intracellular concentrations of the target metabolite will trigger synthesis of the protein, enabling fluorescence intensity-based FACS screening. The developed metabolite sensors were evaluated for their sensitivity, specificity towards the target metabolite, and dynamic range. These were then used for strain development with improved precursor supply.

Task 6.6: *Synthetic genetic circuits to control the temporal profile of gene expression*

Genetic circuits based on well-established regulatory networks were used for controlling expression of the polyphenol biosynthetic genes. For example, bacteriocin nisin-inducible promoter P_{nisA} was used for controlling resveratrol production in *L. lactis* and IPTG-inducible T7-polymerase-based expression system was used in *C. glutamicum*.

Summary of results

Corynebacterium glutamicum

Polyphenols are produced from phenylpropanoid starter units, such as *p*-coumaric or cinnamic acids, which in turn are synthesized from amino acids $_L$-tyrosine and $_L$-phenylalanine (Marienhagen and Bott 2012). First expression cassettes for *C. glutamicum* contained genes encoding for tyrosine ammonia lyases (TAL) and 4-coumarate:CoA ligases (4CL) as artificial bicistronic operons under control of the

strong T7 promoter enabling synthesis of such phenylpropanoids from L-tyrosine **(D6.1)**. For all engineering tasks eventually leading to *C. glutamicum* cell factories, *C. glutamicum* MB001(DE3), a prophage-free variant of the wild-type strain *C. glutamicum* ATCC 13032, was used.

During strain construction and testing we found that *C. glutamicum* is able to utilize phenylpropanoids, including *p*-coumaric acid, ferulic acid and caffeic acid, as sole carbon and energy sources. Global gene expression analyses identified a gene cluster, which showed increased transcription levels in response to phenylpropanoids (Kallscheuer et al. 2016a). Cultivation experiments conducted with *C. glutamicum* strains carrying single gene deletions showed that loss of *phdA*, *phdB*, *phdC* or *phdE* abolished growth of *C. glutamicum* with all tested phenylpropanoid substrates. These results, together with the intracellular accumulation of pathway intermediates determined via LC-ESI-MS/MS, showed that the *phd* gene cluster encodes for a CoA-dependent, β-oxidative deacetylation pathway **(Figure 3)**, which is essential for the utilization of phenylpropanoids in *C. glutamicum* (Kallscheuer et al. 2016a). Subsequent deletion of key genes of this pathway, along with 17 other genes involved in the catabolism of aromatic compounds, yielded the strain *C. glutamicum* DelAro, which is unable to catabolize any phenylpropanoids previously produced from endogenous L-tyrosine via the TAL and 4CL activity. This strain provided sufficient precursors for the synthesis of phenolic compounds **(D6.2)**.

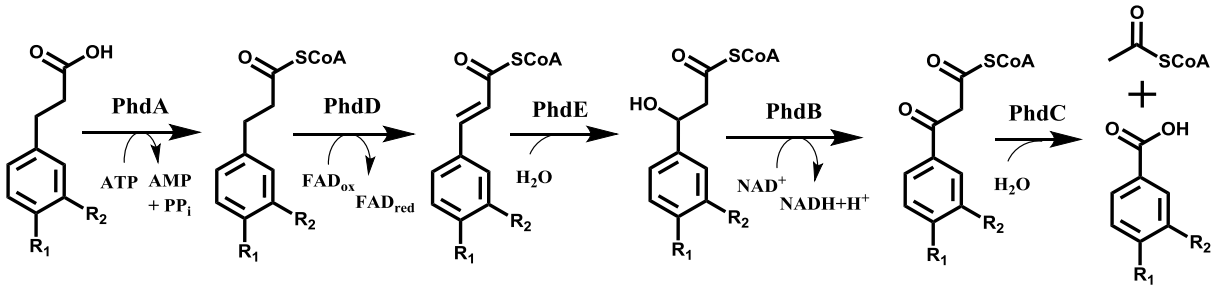

Figure 3: Catabolic route for the degradation of phenylpropanoids in *C. glutamicum*
Phenylpropanoids are degraded by a β-oxidative side-chain shortening pathway yielding benzoic acid derivates. Subsequent degradation leads to TCA cycle intermediates by ring cleavage of 3,4-dihydroxybenzoic acid = protocatechuic acid, a central aromatic substrate of C. glutamicum.

Two recombinant *C. glutamicum* strains, engineered for production of polyphenols, were sent to project partner Biotempo in June 2015 for upscaling **(D6.4)**. *C. glutamicum* DelAro pMKEx2_STSAh$_{Cg}$-4CLPc$_{Cg}$ is able to convert *p*-coumaric acid into resveratrol, and product titers of up to 158 mg/L could be achieved with this strain. Strain *C. glutamicum* DelAro pMK2_CHSPh$_{Cg}$-CHIPh$_{Cg}$ accumulates up to 35 mg/L naringenin and 37 mg/L eriodictyol from *p*-coumaric acid and caffeic acid, respectively**.** Additional engineering of the amino acid metabolism for an optimal connection to the synthetic plant polyphenol pathways enabled resveratrol production directly from glucose, which allows avoiding supplementation with expensive phenylpropanoid precursors (Kallscheuer et al. 2016b). This strain generation contained the final design of the desired expression cassettes, which allow for the temporal expression of heterologous, plant-derived genes (**Figure 4**)**, (D6.7, D6.8)**. Furthermore, modularity of the expression cassettes enabled the rapid engineering of *C. glutamicum* strains capable of producing *O*-methylated stilbene derivatives, such as pinostilbene and pterostilbene, as well as of the flavonols kaempferol and quercetin (Kallscheuer et al. 2017a *in preparation*) **(D6.9)**. These strains were also made available to project partner Biotempo for upscaling.

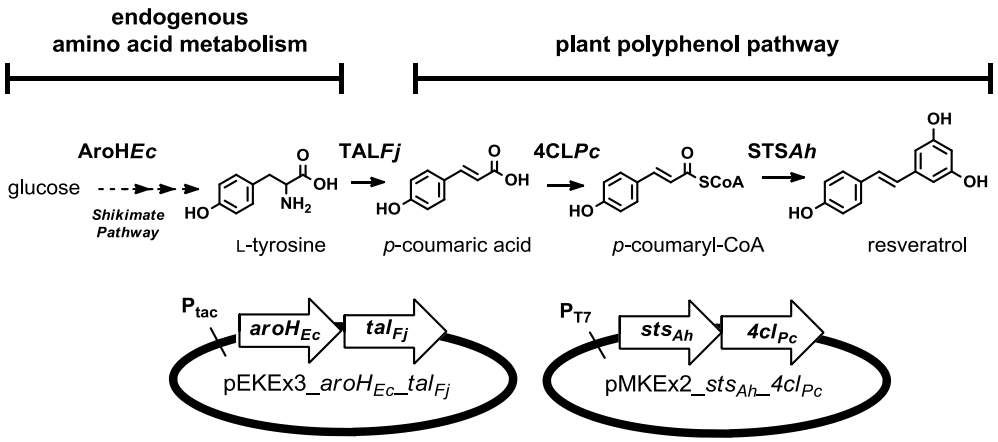

Figure 4: Schematic representation of the overall pathway for resveratrol production in *C. glutamicum*
The pathway from glucose and two plasmids harboring relevant gene expression cassettes for resveratrol production are shown. The genes aroH$_{Ec}$ and tal$_{Fj}$ were expressed as synthetic operon under temporal control of the P$_{tac}$ promoter in pEKEx3, 4clPc and stsAh were cloned as synthetic operon into pMKEx2 under temporal control of the T7 promoter.

In the course of the project, two metabolite sensors for the intracellular detection of phenylpropanoid CoA-thioesters and malonyl-CoA, both being precursors for the synthesis of plant polyphenols, were constructed and characterized **(D6.3)**. The sensor pSenCg0343 for phenylpropanoid CoA-thioesters was used for a protein engineering campaign to increase the heterologous activity of a TAL in *C. glutamicum*, which is originally from *Flavobacterium johnsoniae*. Here, the experiments are not finished yet, and will be continued after BacHBerry is over. The malonyl-CoA sensor pSenFapR was used to screen chemically-mutagenized *C. glutamicum* DelAro cells for variants with an increased intracellular pool of malonyl-CoA. Screening of more than 1 x 10^6 cells yielded several variants with an increased malonyl-CoA availability, which was also reflected by very high titers of plant polyphenolics. In the conducted experiments, the best strain K1A1 accumulated up to 115 mg/L resveratrol without any cerulenin-mediated manipulation of the malonyl-CoA pool **(D6.7)**. Currently, the genome of this strain is sequenced for identifying the underlying genomic mutations.

Lactococcus lactis

L. lactis is a Gram-positive bacterium that has a long history of use in the production of cheese and buttermilk. The majority of work within this project was done with *L. lactis* NZ9000, a derivative of MG1363 with the *nisRK* genes integrated into chromosome (Kuipers et al 1998), and its lactate dehydrogenase (*ldh*)-negative mutant.

L. lactis strains with plant biosynthetic pathways for anthocyanin production were constructed. Pelargondin-3-*O*-glucoside is produced from naringenin in four steps, catalyzed by F3H (flavanone 3-hydroxylase), DFR (dihydroflavonol reductase), ANS (anthocyanidin synthase), and 3GT (anthocyanidin 3-*O*-glucosyltransferase) from different plant sources were tested for expression and functionality in *L. lactis*. Expression of each protein was verified using Western Blot and the best-expressed genes were assembled into operons. However, DFR activity was determined to be the rate-limiting step. Another bottleneck was identified at the level of 3GT, where over-expression from a plasmid resulted in pronounced toxicity. Therefore, the gene

was integrated into the chromosome of *L. lactis*, which allowed maintaining sufficient activity and avoiding the toxicity.

Another anthocyanin, cyanidin-3-*O*-gucoside can be produced from catechin without the need for DFR. In order to express this pathway in *L. lactis,* the *ans* gene were placed on a plasmid under control of P_{nisA}, while the *3gt* was integrated into the chromosome under P_{usp45}. This strain was able to produce various red-yellow cyanidin derivatives in mg/L range. Some of the compounds had unusual and, presumably, novel structures. In addition, production of cyanidin from tea extracts was achieved.

The entry of the precursors catechin and naringenin into lactococcal cells was identified as the main bottleneck for anthocyanin production. In order to enhance their uptake, various cell wall permeabilization strategies were employed, such as pre-treatment with lysozyme or a surfactant Tween-20. Furthermore, in order to understand how the cell responds to naringenin, gene expression of the production strain was analyzed in the presence of naringenin and deletion of a multidrug-resistance transporter, which was upregulated in its presence, made this strain more sensitive to naringenin, thus allowing better uptake.

Moreover, the resveratrol biosynthesis pathway, consisting of three enzymes: TAL, 4CL, and STS (stilbene synthase), has been introduced into *L. lactis*. The biosynthetic genes were placed on a high copy number vector under the control of P_{nisA}. Two configurations were tested: operon and pseudo-operon, where each gene is preceded by an individual promoter, but only a single terminator is located at the 3' end of the expression cassette. Upon cultivation of the designed strain in a chemically defined medium, the strains were producing resveratrol on mg/L scale (**D6.1**). When the *tal* gene was replaced with a promiscuous PAL/TAL enzyme, production of pinosylvin (*trans*-3,5-dihydroxystilbene) from $_L$-phenylalanine was also detected. Moreover, production of methylated variants of *trans*-resveratrol was also achieved via co-expression of resveratrol-*O*-methltransferases from various plant species.

Similarly to *C. glutamicum, m*etabolic sensors for the limiting precursors, malonyl-CoA and *p*-coumaroyl-CoA, were constructed in *L. lactis* **(D6.3)**. The former biosensor was successfully used for selection of a strain with enhanced resveratrol production via overexpression of acetyl-CoA carboxylase enzymes from different sources. The best performer had the production increased by more than 2-fold.

We have also tried to develop a platform for production of plant-derived cytochrome P450s in *L. lactis* that are involved some steps of polyphenol biosynthesis. In plants, these enzymes are normally associated with the endoplasmic reticulum membrane, hence their production in bacterial hosts is often challenging (Effendi and Koffas 2007). A model cytochrome was fused with a GFP gene, so that a fluorescent signal could be observed if the enzyme was expressed. It was therefore possible to detect expression of a model P450 CYP79A1 involved in the dhurrin pathway, but not of CYP73A from the phenylpropanoid pathway. Furthermore, no activity of CYP79A1 could be detected under any of the condition tested. Hence, in the remaining work, alternative enzymes that bypass the P450-catalyzed steps, such as TAL, were utilized (**D6.5**).

In the earlier stages of the project, we used *E. coli* as a platform for heterologous assembly of polyphenol biosynthetic pathways in bacteria. We aimed at reconstruction of a previously unknown biosynthetic pathway for the flavonol fisetin, a potent anti-inflammatory, antioxidant, and neuroprotective compound that was investigated within WP4. A pathway consisting of 7 steps and 9 enzymes was assembled and the final yield of fisetin was 0.3 mg/L directly from $_{L}$-tyrosine (Stahlhut et al 2015). Moreover, in order to stabilize fisetin, several glycosyltransferases, both previously known and newly-discovered within BacHBerry, were assayed for functionality and integrated into the pathway.

Other activities related to this WP included analysis and development of molecular tools to assist strain engineering. An extensive set of tools is available for both *C. glutamicum* and *L. lactis* (For reviews, see Pátek and Nešvera J (2013), Kortmann *et al.,* (2015), and Gaspar *et al*. (2013)), many of which were used for strain engineering within the project. Moreover, several origins of replication with different copy-number

and different selection marker genes were evaluated for functionality in *L. lactis*. These plasmids are to be combined with a set of designed synthetic double- and quadruple- inducible promoter systems, which should allow simultaneous expression of long biosynthetic pathways. Lastly, multiple attempts to establish CRISPR/Cas-based genome engineering in *L. lactis* were made using *cas9* genes from different sources.

1.3.5 WP7: Extraction and bioseparation of phenolics

Objectives and tasks

The ultimate goal of WP7 was to design and optimize methods for extraction of bioactive compounds for testing in WP3&4, as well as for extraction of phenolic compounds produced by bacteria after fermentation. The specific objectives were: design and optimization of methods to extract bioactive compounds from berry biomass, as well as design and optimization of methods to extract bioactive compounds from fermentation broth.

Task 7.1: *Design & optimization of berry biomass phenolics extraction methods*.

WP7 assisted WP3 in the development of a SOP for phenolics extraction and identification from berry biomass, by elaborating upon previous work and fine tuning existing extraction methods and techniques.

Task 7.2: *Design & optimization of fermentation phenolics extraction methods*.

Task 7.2.1: *Liquid-Liquid (L-L) Extraction*. Small scale extraction experiments to identify optimal L-L extraction possibilities (e.g. ethanol extraction). Optimization of solvent composition and separation conditions for optimal capacity, selectivity, and recovery yield.

Task 7.2.2: *Adsorption-based extraction*. Small scale extraction experiments to identify optimal solid phase extraction (adsorption) possibilities. Perform screening of various resins in order to identify optimal matrix type, ligand type, manufacturer, size, etc. in terms of capacity, selectivity, and recovery yield. Select the most suitable operational mode (e.g. Fixed Bed, Simulated Moving Bed, or Expanded Bed Adsorption).

Once the most optimal extraction procedures were selected and the parameters were optimized, scaled-up of the designed process were perform up to pilot levels. Larger-scale processes (with chosen production capacity) were designed based on modelling simulations and experimental results and tested in fermenters of various sizes in close consultation with other WPs (e.g. for optimal fermentation broth

composition or with regard to optimization of the extraction and purification approaches). Polyphenols extracted from the designed fermentation processes were provided for confirmation of bioactivity to WP4. Lastly, techno-economic evaluation of the designed industrial fermentation-based production and extraction process was performed.

Task 7.3: *Semi-preparative fractionation and purification of phenolics from plants for testing*

This task is in-line with **Task 3.4** and aimed at facilitation and performing of larger scale semi-preparative chromatographic fractionation of the most interesting berry extracts, following the results of the bioprospecting and color functionality screening procedures.

Summary of results

Liquid-Liquid (L-L) extraction has been investigated as a potential method for purification of polyphenols from fermentation broth. Similarly, Solid-Liquid (S-L) extraction, or adsorption/chromatography, has been experimentally and computationally investigated. Based on the obtained result, different process concepts have been developed and compared via detailed conceptual process design calculations. The most promising one has been experimentally investigated in collaboration with WP9, and successfully operated. Additionally, semi-preparative fractionation protocols have been developed, and applied to obtain berry fractions for bioprospecting in WP3 and WP4. A brief overview of the main results is given below.

Semi-preparative fractionation and purification of phenolics from plants for testing

Fractionation of bioactive fruit extracts was developed using chromatography. The freeze-dried extract of selected berries (eq. *Rubus sp.*) were re-suspended in Milli-Q water, filtered and adsorbed onto a semi-preparation column packed with a C18 resin (PREP C18 55-105 µm 125Å, Waters Co. USA) using an AKTA Explorer (GE USA). 0.1% Formic Acid (Sigma, USA) in Milli-Q water was chosen as mobile phase A and

Methanol (Chromasolv purity≥99.9%, Sigma, USA) as mobile phase B. The approach was modified for the different berry extracts in order to achieve proper fractionation, by tuning the hydrophobicity of the mobile phase (water-methanol). Modifying loading and elution mobile phase composition and gradient profile served to generate a maximum number of well identifiable peak/fractions for further bio-activity testing (ranging from 12 to 28 fractions – **see Figure 5**). Fractions were brought under vacuum using a rapid-vac to remove methanol for further bioactivity testing.

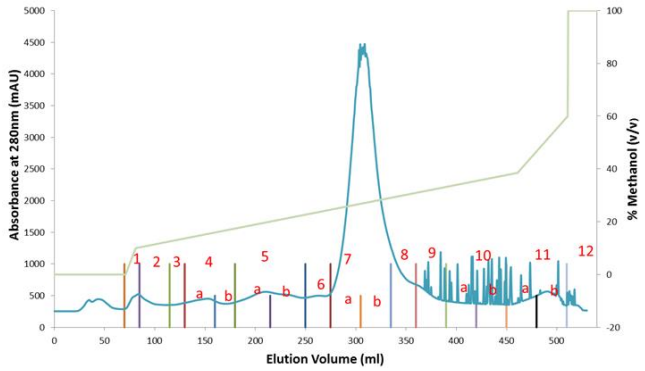

Figure 5: Semi preparative berry fractionation profile of *Rubus sp. 1*.

Thermodynamic models for prediction of partition coefficients in liquid-liquid extraction

Next to the MPP-UNIFAC model developed by our group (Méndez Sevillano et al, 2014), an approach has been developed without the need for previously-regressed interaction parameters, namely COSMO-RS. This model can be used next to the more general mod-UNIFAC in situations when no reliable data is expected from this mod-UNIFAC. A publication is in preparation describing both thermodynamic modelling approaches combined with extensive solvents testing.

Adsorption for polyphenol capture and purification

Next to liquid-liquid extraction, adsorption using a solid auxiliary phase can be used, leading to very high purities. Ten different hydrophobic resins were tested and isotherms were determined for model polyphenols resveratrol, quercetin and fisetin. Different ethanol percentages were used as modifier as desorbent (**Figure 6**). The separation behavior on chromatographic columns was simulated accurately with

Aspen Chromatography®, allowing using this tool in total process development and design. A publication is in preparation describing resin screening and chromatographic separations of key model polyphenols as well as chromatographic separation simulations.

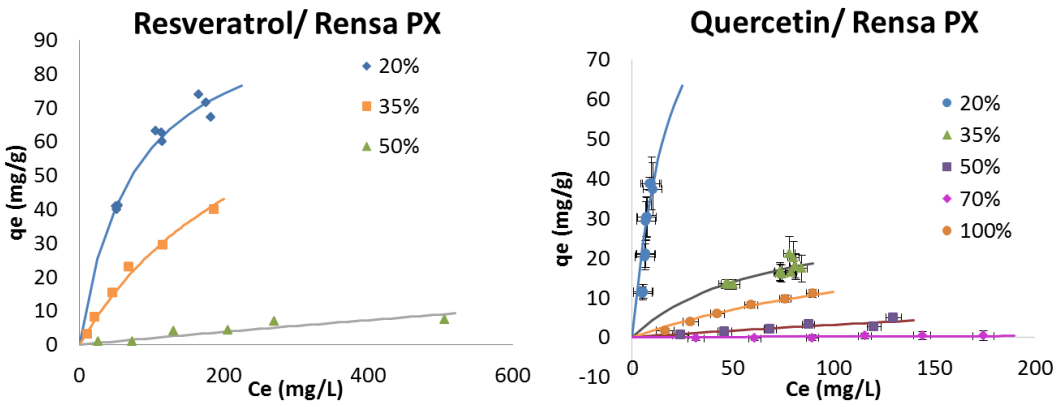

Figure 6: Adsorption isotherms on Rensa PX at different ethanol liquid fractions. Left: resveratrol; Right: quercetin.

Polyphenol Diffusion

Diffusion influences mass transfer and is of key importance in designing an absorber, a crystallizer, or an extraction protocol, next to the thermodynamic equilibrium (translated to solubility, partition coefficients and adsorption isotherms). Using a microfluidic diffusion cell, accurate results of polyphenol diffusion were obtained. With that, an adaptation has been developed for the well-known Wilke-Chang diffusion model, based on the number of hydroxyl group of the polyphenol to accurately predict diffusion coefficients of a representative selection of polyphenols (Méndez Sevillano et al, *in preparation*)

Preferential crystallization of similar polyphenols

Polyphenols normally have low solubility in aqueous solutions. An advantage of this property can be made by crystallizing the polyphenols, followed by filtering and washing the obtained crystals, thereby simplifying the downstream processing significantly, reducing CAPEX and OPEX. An elegant method was developed and

successfully tested to preferentially crystallize two model polyphenols with similar solubility, namely resveratrol and naringenin, in two different and connected crystallizers by controlling the flow between the crystallizers and the temperature profiles. A publication is in preparation describing preferential crystallization of similar polyphenols.

Conceptual process design of the entire industrial production process

Based on the obtained experimental and modelling results, combined with the data on the microbial production efficiency achieved by other partners, a conceptual process design was made to assess economy feasibility of the pipeline. Several processing concepts were evaluated and the most optimal process scheme was designed in detail. Results have been reported in the deliverable report D7.3 (Techno-economic aspects of optimal fermentation based on large-scale process for phenolics production) as well as a Group Design Report (da Silva, Ferreira et al, Techno Economic Evaluation of Polyphenol Production and Purification on Industrial Scale, Report GDP 2015/2016 BioProcess Design, Group Design Project, Delft University of Technology, 29 February, 2016). A part of the total process is shown in **Figure 7**.

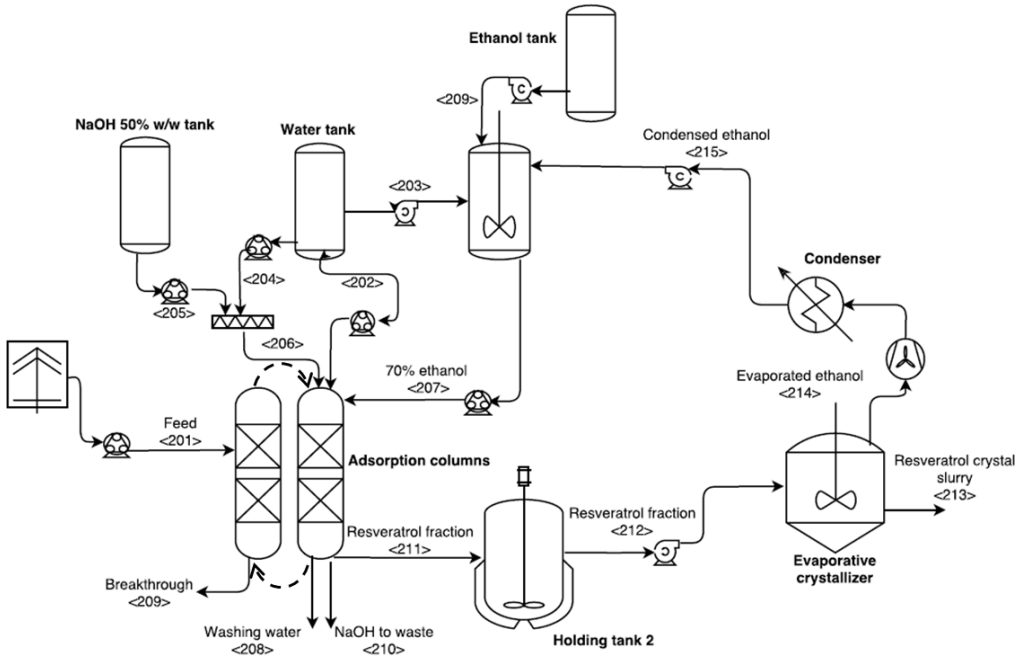

Figure 7: Process Flow Diagram over the Adsorption section of the overall process

Large scale proof of concept trials: adsorption for polyphenol capture and purification

In a close collaboration with WP9 (Biotempo), large scale fermentations and purification have been carried out using one of the downstream processing methods developed in WP7, namely adsorption. A process concept using In Situ Product Recovery was successfully applied. A joint article describing this ISPR method is in preparation.

1.3.6 WP8: Bioinformatics and modelling methods

Objectives and tasks

The main goal of WP8 was to explore the existing mathematical and computational methods to engineer an optimal production of phenolics using *Lactococcus lactis* and *Corynebacterium glutamicum*. The specific objectives were to evaluate the effect of the shortcomings presented by each method and design new ones that make a full use of the different types of the available omics data. The models and algorithms to be developed were divided into three main tasks corresponding to the different steps or approaches to the issue of plant bioactive molecule production using microorganisms. These included: (**Task 8.1**) Sequence analysis and plant functional genomics aimed at developing algorithms for analysis of berry transcriptome data generated in WP5; (**Task 8.2**) Sequence design for heterologous expression of plant genes in bacteria; and (**Task 8.3**) Metabolic modelling for optimal production of molecules. The latter task was subsequently subdivided into: (**Task 8.3.1**) Genome-wide reconstruction of the embedded metabolic pathways; (**Task 8.3.2**) Network optimization for polyphenol production in a single uniform population of bacteria and then (**Task 8.3.3**) for optimizing polyphenol production in a community of bacteria.

Summary of results

Task 8.1: *Sequence analysis and plant functional genomics*

MASSBLAST - A workflow to accelerate RNA-seq and DNA database analysis

Two pieces of software were developed for analyzing the transcriptome (RNA-seq) data, namely GENEEXTRACTOR and MASSBLAST. GENEEXTRACTOR searches KEGG and NCBI databases for different compounds and downloads the related genes. MASSBLAST behaves as an intermediate step to perform BLAST queries. MASSBLAST queries multiple files from GENEEXTRACTOR against existing RNA-seq data of the berries from the project and summarizes the result in a single file.

MASSBLAST was applied to the project data, in close collaboration with WP5. The most recent results on the available blackberry transcriptome data show a significant improvement in the analysis, with a processing speed of >1000 times faster than manual curation, with excellent accuracy.

MASSBLAST can be applied beyond the BacHBerry project virtually on any transcriptome data. It is also freely available online for academic users at https://github.com/averissimo/mass-blast with its accompanying software ORFFINDER (https://github.com/averissimo/orf_finder).

Task 8.2: *Sequence design for heterologous expression*

Machine learning models capable of predicting protein expression levels based on their codon encoding were investigated. In particular, support vector regression and partial least squares were used to build models using two available benchmark datasets of three proteins from *E. coli*. The results indicated that support vector regression results in more accurate gene expression predictions, as compared to partial least squares. Moreover, it is possible to improve gene expression predictions by using two more input features, codon identification number and codon count, besides the already used codon bias and minimum free energy. Finally, we showed that ensemble averaging allows improving gene expression prediction efficiency even further.

Task 8.3: *Metabolic modelling for optimal production of molecules in L. lactis and C. glutamicum*

Task 8.3.1: Genome-wide reconstruction of the embedded pathway

Genome-scale models of *L. lactis* and *C. glutamicum* were expanded to accommodate the introduced heterologous pathways, derived from the plant analyses and phenolic pathways identification previously performed in WP6. They are available in SBML format for MATLAB and COBRA and can be further used in other applications.

Task 8.3.2: Network optimization for a single population

OptPipe – a pipeline for optimizing metabolic engineering target

Recently, several computational methods for *in silico* metabolic network optimization have emerged and are currently being applied in metabolic engineering. They are based on different approaches and rationales, thereby leading to distinct solutions. One key issue is the multitude of obtained results, known to be dependent on the specific algorithms and the solvers used. In order to address this problem, a consensus-based approach was proposed and developed. This method is based on running several optimization procedures and analyzing *a posteriori* the solutions, looking for a consensus. The rationale is to have rankings of hypotheses that may provide confidence in particular sets of proposed genetic alterations from various aspects.

In the case of a single population of microbes *(i.e.* a monoculture), the genome-scale models of *L. lactis* and *C. glutamicum* were extended with the biosynthetic pathways of the four above-mentioned model polyphenols, during *Task 8.3.1. Genome-wide reconstruction of the embedded pathway*. Subsequently, five different optimization methods were applied to the genome-scale models and the consensus ranking was obtained for each strain and compound. The hypotheses were generated with the rank product test, and the outputs were lists of deletions that achieve the best ranks using defined criteria, namely maximum predicted target compound production, maximizing the minimum predicted target compound production and distance from the wild-type flux distribution. The obtained results were then further filtered and experimentally validated, resulting in several promising mutant strains where production was increased by up to three-fold. The software package called OPTPIPE is available at: https://github.com/AndrasHartmann/OptPipe.

Task 8.3.3: Network optimization for a community of bacteria

MultiPus – MULTIple species for the synthetic Production of Useful biochemical Substances

For a long time, single microbial strains (monoculture) have been used for the production of high-value compounds of interest, as exemplified by the use of *Saccharomyces cerevisiae* for production of artemisinic acid, which is important for obtaining an effective anti-malaria drug, artemisinin. However, there has been an increasing body of literature over the last decade exploring the use of not only a single microbial strain, but of a community of microbial species for such tasks. There are various reasons for this. One is related to a potentially increased efficiency of the production system. Another important motivation is to augment the chances of avoiding toxic effects for each of the recruited microorganisms. Microbial consortia are thus believed to be able to perform more complex functions, and to be more robust as a group to environmental fluctuations. The challenge remains, however, in precisely establishing which consortia are the best for the production of a given compound or of set thereof. We introduced an initial model, topological for now, and a combinatorial algorithm, called MULTIPUS, that enable proposing optimal consortia to synthetically produce compounds that are either exogenous to it, or are endogenous, but where interaction among the species in the consortium could improve the production line. MULTIPUS was then applied to several case-studies.

The software MULTIPUS is freely available for academic users at http://multipus.gforge.inria.fr/.

MultiOpt - Multi-objective mixed integer optimization

A crucial part of our future work will be to put together the model and method established in **Task 8.3.3** with the multi-objective optimization that we also are now developing. We indeed explored the concept of multi-objective optimization in the field of metabolic engineering, when both continuous and integer decision variables are involved in the model. In particular, we proposed multi-objective models to suggest reaction deletion strategies and to deal with problems where several functions must be optimized simultaneously, such as maximization of bioproducts while minimizing toxicity. We thus introduced MULTIOPT (Multi-Objective Mixed integer Optimization for metabolic engineering: applications to strain optimization for improving bioproducts), a computational framework that aims at modelling and

solving optimization problems, designed for predicting reaction knockout strategies, by means of multi-objective programming. We compared the results from MULTIOPT with those obtained by using the well-known bi-level optimization model (OPTKNOCK). Furthermore, we studied two multi-objective optimization problems arising from the metabolic engineering of microorganisms. Preliminary results illustrate that a multi-objective approach can provide more hypotheses regarding candidate deletions for improving target compound production.

A prototype of MULTIOPT is available at: http://web.ist.utl.pt/~susanavinga/MultiOpt/

1.3.7 WP9: Optimization of phenolics production by fermentation and scale-up of bioprocesses

Objectives and tasks

The main goal of WP9 was the development of an optimized production process of phenolic compounds in microbial cell factories using lab-scale automatically-controlled fermenters. A strong emphasis was given to the economic feasibility of a possible industrial process using such technologies. In order to achieve this goal, the WP objectives were split into: **Task 9.1:** *Fermentation Optimization* and **Task 9.2:** *Bioprocess scale-up*. The former task was further split into sub-tasks: optimization the environmental conditions, medium composition and operation strategy (batch vs fed-batch), and evaluation of cheese whey as an alternative carbon source for fermentation. The latter task aimed at using the obtained data and the optimized parameters for preparation of the scale-up to be executed in WP10.

Summary of results

Task 9.1: *Fermentation Optimization*

The following bioreactor operational parameters were addressed within this WP:

Task 9.1.1: Environmental conditions

The environmental conditions were optimized for *C. glutamicum* and *L. lactis* strains. The most relevant results were related to the optimization of the dissolved oxygen concentration in the system, as well as the supplementation of the media with limiting salts when the fermentations are run in fed-batch mode. The main issue with the former parameter was that on one hand, oxidation of dissolved phenolic compounds was observed during the fermenter operation, whereas on the other hand, the operation in aerobic conditions enhanced the cellular growth for both species. Hence, equilibrium had to be found between the product degradation and the growth rate in order to maximize the productivity.

Furthermore, *C. glutamicum* proved to be strongly dependent on good aeration rates in order not to produce organic acids. On the other hand, the high levels of aeration and agitations required by these strains, made it mandatory to use high levels of anti-foaming agents. The choice of the anti-foaming agent and the determination of its feeding policy was also one of the key parameters that needed to be optimized.

Task 9.1.2: Medium composition

The first set of optimization efforts was centered on the use of hemin to reinforce the growth rate of *L. lactis* under aerobic conditions. However, the results would vary according to the strains under study.

The second important consideration was the cost of the fermentation media in a potential future industrial operation scenario. In order to address these issues, market studies were conducted to evaluate the cost of each component in a large-scale production scenario. Several media, which allow avoiding the use of expensive reactants, were proposed in collaboration with WP6 and further tested using bioreactors. The most cost-affecting omissions were the supplementations with the BHI fermentation medium and the polyphenol production enhancer cerulenin. In the case of the anthocyanin production by *L. lactis*, an industrial-scale lysozyme supplier was found and its product was tested. All of this allowed achieving significant reduction of the media prices.

Lastly, in order to evaluate the possibility of using cheese whey, a waste product of dairy industry that contains lactose, as an alternative nutrient source for the fermentation, WP6 designed a *L. lactis* strain capable of utilizing this sugar. The strain was then cultivated in the presence of lactose as a carbon source. The results obtained with this strain demonstrate that it is possible to produce phenolic compounds from cheese whey. However, the productivity is about half of what was obtained with glucose.

Task 9.1.3: Bioreactor operation

In silico optimization of a fed-batch operation strategy was performed, based on phenomenological models developed for the wild type strains of *L. lactis* and *C.*

glutamicum. The achieved results clearly pointed out the advantages of a fed-batch operation strategy. These strategies would be later on applied in real fermenter-based cultivation experiments with the engineered resveratrol-producing *L. lactis* and *C. glutamicum*.

The use of fed-batch operation mode was demonstrated to be the most relevant tool for improving the production efficiency for both resveratrol and the compound RC-1. In general, it was possible to double the amount of biomass produced per batch and to increase the amount of produced phenolic compounds by 1.5 fold. Both the titre and the productivity were improved with this strategy. The best performing resveratrol producer strain of *C. glutamicum* could reach the titers of about 100 mg/L in this mode of operation, which is more than 20 times the titre obtained in shake flask operation.

Following the findings that resveratrol is degraded through oxidation during the operation of the fermenter and that a large amount of resveratrol was found inside the cells, two strategies were put in place in order to stabilize the produced resveratrol and to dislocate a possible equilibrium in the production reaction towards resveratrol. Both strategies consisted of promoting the adsorption of resveratrol from the medium into a chromatographic resin during the fermentation process itself. These strategies were designed together with TUD and subsequently integrated into the downstream processing. Although the total amount of resveratrol produced per batch did not increase, the simplification of the downstream process and the stabilization of a significant amount of resveratrol in the fermenter could justify by itself this operation mode. The reason why the productivity did not increase is related to the observation that not only resveratrol was adsorbed but also its precursor – *p*-coumaric acid. Future research should focus on alternative, more selective, adsorption techniques.

Task 9.2: *Scale-up*

Within WP9, the scale-up was performed in shake flasks, 2L, 3.5L and 5L fermenters. The use of fermenters with varied configurations allowed a better understanding of the systems behavior and prepared the ground for further scale-up in WP10.

The most important of the addressed bottlenecks was the foaming control in fermentation systems with *C. glutamicum*. On one hand, foaming was responsible for a significant loss of productivity due to biomass getting trapped in the foam. On the other hand, the addition of anti-foaming agents to the system reduced its productivity in terms of both biomass and final product titers. The strategies, defined to deal with this issue, were dependent on various vessel characteristics, such as aspect ratio, propeller configuration, aeration capacity, existence or absence of baffles, etc. There was no clear recipe identified to minimize foaming and the addition of the anti-foaming agent was carried out “as needed” and was always dependent on the opinion of the operator.

1.3.8 WP10: Bioprocess scale-up: fermentations at 250-L scale and downstream processing

Objectives and tasks

WP10 had a single task (**10.1**) that aimed at demonstrating the feasibility of a possible industrial-scale process for microbial production of phenolic compounds using strains developed within BacHBerry. In order to achieve this goal, 10L, 50L and 250L (pilot-scale) fermentations were carried out using the best-producer strain, available at that point of time, for the production of the selected model phenolic compound, namely resveratrol. One of the designed downstream processing systems for continuous product removal was tested simultaneously.

Summary of results

Task 10.1: *Bioprocess scale-up: fermentations at 250-L scale*

The production of the model polyphenol resveratrol using *C. glutamicum* strain DelAro^4, developed by Juelich, has been tested in 10L-, 50L- and 250L-scale fermenters. The fed-batch operation with integrated continuous product removal strategy was put in place. The results confirmed the scalability of the developed technologies and confirmed the need to fine-tune the strategy to the characteristics of specific fermenter to be used for production. Obtained titers were comparable to those obtained in 5L fermenters. The most relevant parameters were the initial glucose concentration, the oxygen transfer capacity of the vessel, and the feeding strategy during the fed-batch phase.

1.3.9 WP2: Dissemination, training, societal and ethical issues

Objectives and tasks

The objectives of WP2 were to train our researchers, to disseminate the aims and results of the project to a general public, to identify and characterize markets for novel phenolic products, and perform an analysis for the most promising market. To fulfill these objectives, seven tasks have been assigned to this WP. They are tasks to study the economic issues and regulatory frameworks (**Task 2.1**); to follow up the new development of implementations of access and benefit sharing regulations (**Task 2.2**); to conduct sustainability assessment of the developed pipeline and processes (**Task 2.3**); to organize open dialogues with potential stakeholders and other interested parties (**Task 2.4**); to disseminate the project and its outcomes to general public through multiple channels of media (film and game) (**Task 2.5**), as well as to scientists other EU-funded consortia (**Task 2.6**); to organize special training sessions, such as career workshops and summer schools (**Task 2.7**); and to draft an exploitation plan for future product development based on the outcomes of the project (**Task 2.8**).

Summary of results

A report was delivered, focusing on the economic, regulatory, and sustainability impacts of producing phenolic compounds, derived from the mining of berry genomic resources, using microorganisms for pharmaceutical and nutraceutical uses, and as novel food additives. The assessment of economic issues of the project has been focused on the market potential of natural colorants identified during the screening of berry germplasm. An overview was written, analyzing the relevant regulations and procedures for approval of novel food additives, including colorants, in the EU, the United States (US), and China. Regulations within both the EU and China confer a strict set of requirements for bringing a novel food colorant to the market, even a natural one. These rules consider not only the composition and purity of the substance, but also the process through which the compounds have been produced.

In contrast, in the US, a novel nutraceutical can gain access to the market by simply passing the Generally Recognized as Safe (GRAS) assessment. Hence, if a novel polyphenolic colorant is to be produced using engineered microorganisms, the US market should probably be the first one to be targeted due to less strict regulatory framework for approval. To assess the possible contributions of BacHBerry's research activities towards building a more sustainable society, sustainability of the project has been assessed based on environmental, economical, and social impacts (Pei and Schmidt 2016).

Another report was produced to follow up on the updates of the implementation of *the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization (NP-ABS)* (CBD 2012) in the EU, as well as preparing recommendations for best practices in Access and Benefit Sharing for the consortium. The Nagoya Protocol came into force in the EU as of 12.10.2014. As most of the BacHBerry project partners were from EU Member States, the best practices suggested from both the Convention on Biological Diversity (CBD) and the EU were reviewed, and the relevant guidelines and updates on implementing ABS were provided to the partners. As of October 2016, from the countries participating in the project, Denmark, Germany, Switzerland, and UK ratified the NP-ABS and began implementing the aforementioned regulations.

We have also used multiple channels of communication to engage with the stakeholders in order to broaden the societal impact of project. This was done through one student workshop (Vienna, November 2014), three stakeholder workshops in Austria (in Vienna, March 2015), Denmark (in Copenhagen, February 2015), China (Beijing, December 2014) and post-workshop interviews, a documentary film about the project and its activities, and lastly, an online science game has been made as a more interactive way to engage the young audience. Moreover, a career development workshop for young researchers (Lyon, France, November 2015) and a summer school on computational and modeling tools (Lisbon, Portugal, November 2014) were organized. Lastly, a satellite workshop "Bacterial Hosts for Production of Bioactive Phenolics from Berry Fruits" was conducted at the 28th International Conference on Polyphenols (Vienna, Austria, July 2016).

A report on the open dialogue activities was prepared, analyzing the key issues brought up in discussion during the interactive workshops and the stakeholder interviews, as well as the responses from the scientists in the consortium. When the results from the internal brainstorming are compared with those of the workshops, it is clear that economic aspects, health and the issue of acceptance are important to both groups. The association with GMO and the related debate is only prominent in the scientists' group from the consortium, whereas it only played a minor role in the stakeholder workshops. The crucial issue regarding the societal impact of the project is, therefore, connected with the relevance of the health claims for the polyphenolic compounds to be produced. Another important point raised during the discussions was access benefit sharing. While the scientists were already aware of the Nagoya protocol and its intentions, the second group suggested a need for a framework for regulating genetic resources and indigenous knowledge.

On the topic of whether potential products might be rejected in context of the GMO debate, an event much feared by the consortium scientists, the analysis clearly shows that food issues and medical issues are of higher importance with respect to public opinion. Use of genetically-modified organisms is only one point of concern among the multitude of others, such as trust, naturalness, food safety vs. clinical trials, synergistic effects, *cui bono,* etc. While presenting new projects, new technologies and new potential products to the general public, scientists, industry and policy makers are and will be challenged with questions on risk versus worth (benefit): what are the possible health and environmental risks, and potential social impacts? And, in return, what are the possible and potential gains and benefits? In the case of BacHBerry, the scales were heavily weighted towards the latter, especially with its aim towards production of health-beneficial compounds, while carrying minimal risks, and it was therefore perceived as not overly controversial.

Eventually, an exploitation plan has been drafted aiming to bring the results of the project to the next Technology Readiness Levels. It consists of a summary of all partners and their exploitable results produced within the BachBerry project. There is a description of the agreements reached between the partners for the exploitation of the results, including the Material Transfer Agreements signed is presented.

Considering that an applied R&D project, such as BacHBerry, produces more than scientific reports, other exploitable project outcomes have been compiled in the “Exploitation Plan” (EP). This document starts with the characterisation of the different phenols market: food colourants, functional food ingredients and pharmacological drugs. The BacHBerry consortium partners are then presented, together with their specific expertise. All exploitable BacHBerry outcomes are presented and each specific exploitation strategy is described. Finally, an economic analysis of a possible industrial process to produce phenolic compounds with the developed technology is presented, together with a SWOT analysis if such a process. The Intellectual Property and contracts put in place during the project are presented to serve as guidelines for a future implementation of the technologies developed.

1.4 Potential impact (including the socio-economic impact and the wider societal implications of the project so far), the main dissemination activities and the exploitation of results

The health benefits of the ingestion of certain phenolic compounds present in berry fruits is, nowadays, unquestionable. As this project has demonstrated, we have only started to grasp the immense variety of such compounds and their effect to the human organism. With an increasing world population and the need to provide healthy food to a larger number of human beings, humanity has to deal with the scarcity of farmland to produce nutritionally-rich food. It is in this context that the alternative of producing the most healthy food ingredients at an industrial scale reveals itself as a viable alternative to improve the wellbeing of an ever growing percentage of the world population. The production of functionally active phenolic compounds one of the vectors that such a strategy should address.

The BacHBerry partners have successfully established a complete pipeline from screening (WP3), bioprospecting (WP4), identification (WP3 and 4), microbial production of plant phenolic compounds in Gram-positive bacteria (WP5, 6 and 8), to development of an industrial processes to produce such compounds in large scale (WP7 and 9). Overall, this project provided the scientific community with enabling tools and resources to further advancing research on berries, phenolic compounds, and metabolic engineering of microorganisms. A knowledge database for berries has made available by WP3. Several software tools for bioinformatics and modelling methods have also been available to future exploitation on not only our target bacteria (*Corynebacterium glutamicum* and *Lactococcus lactis*), but also other microbes for the scientific community (see below).

The BacHBerry project developed a set of tools that, in the future, will facilitate the continued research in this area of science. RC-1 was one of the compounds with health-beneficial properties identified in the project. A fermentation-based process has been developed for the industrial productions of RC-1. This is a definitive proof-of-concept for all the strategies put in place by the BacHBerry project. Hence, the

prepared Exploitation Plan presents all these outputs and the strategies adopted to exploit them.

A dedicated work package (WP2) was assigned to provide a comprehensive assessment on the regulatory, economic, environmental and social issue related to the discovery and production of high added value products from the bioprospecting on plant genetic resources, in combination of subsequence development of biological synthesis process for scale-up production. The **Figure 9** shows the general dissemination strategy done by the project.

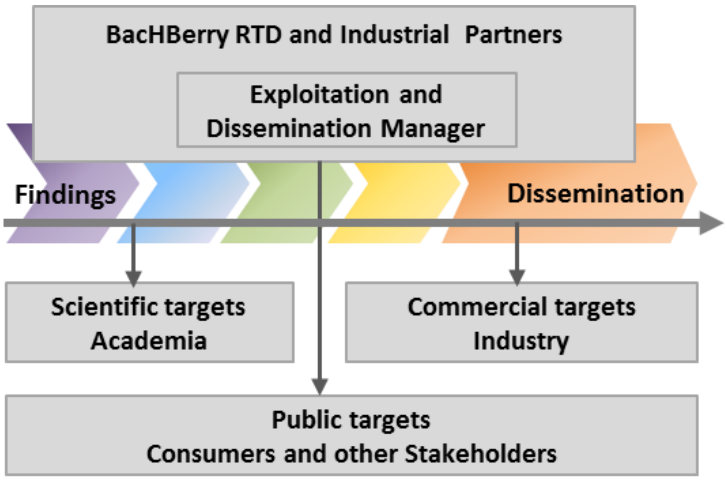

Figure 9: Dissemination strategies used throughout the project

Based on this strategy we have aimed to:

- Contribute to the knowledge for bio-economy via scientific dissemination activities (peer-reviewed articles, academic conferences, dedicated workshops, open resources, etc.)

- Attract the interest for further commercialization of the findings from the project via exploitation plan for potential industrial partners

- Enhance project's visibility at local, national and international levels by public dissemination activities (press releases, project website, dedicated partner Facebook, documentary film, online science game, open dialogue events, etc.)

Our project aimed at contributing to the development of bio-based products and economy by accelerating the turnaround time for going from the discovery to the

most promising and ready products in basic R&D to arrival of end-products to the phenolics market. Thus, the analysis of economic impacts has focused on the market potential of polyphenols as colorants (food additives) in our study. The sustainable assessment has been carried out based on the concept of BacHBerry, covering not only the environmental sustainability, but also the economic and societal sustainability.

A number of stakeholders who have some sort of interest with respect to the outcome of the project were identified and invited to participate in one-day workshops that included elements of brainstorming, as well as group discussions (Steyaert 2005/2006, Goldschmidt 2014). The results of the brainstorming session were categorized using qualitative inductive content analysis strategies.

Based on the findings from the latest Eurobarometer on biotechnology, a large majority of Europeans is unfamiliar with synthetic biology, its risks and benefits, regulation and ethical issues. To tackle this issue, our approach included: open dialogue sessions featuring interactive workshops, documentary film, and a project-inspired science game with free access. Furthermore, dissemination to scientific audiences was done through seminars, invited lectures, conference talks and poster presentations. A summer school (ADDICTION) and a satellite workshop at a polyphenol research-dedicated conference were organized, and a team preparing for the iGEM synthetic biology competition was supported. Lastly, a joint review manuscript covering the entire scope and the progress of BacHBerry has been prepared by the consortium and sent to the Phytochemistry Reviews journal (Dudnik, *et al*, *submitted*)

The dissemination of knowledge obtained within WP3 has been well underway with the delivery of a report on current knowledge regarding the phytochemical composition of different species and their associated bioactivities, which is available to all project partners. It is expected that this could be translated into a review manuscript for submission in a peer-reviewed journal. In addition, the project progress and aims has been disseminated to relevant stakeholders in several occasions: meeting with the Scottish Government Food, Drink and Rural

Communities Policy group (13.03.2015); two independent meetings with members of the Scottish Parliament (12.12.2015 and 20.02.2015); discussions with LR Suntory (a global beverage producer) (04.12.2015); participation in a cross Scottish agri-food research institute engagement (27.02.2015); participating in a meeting with senior policy-makers, business leaders and key scientists entitled Food security: mapping risks, building resilience (1.-2.12.2015).

Moreover, Biofaction produced a documentary film of about 15 minutes to show the aim, efforts and potential results of the BacHBerry project to a wider audience of research-interested non-scientists.

Figure 10: A screenshot from the documentary film

The film was produced in 2015 and 2016 by travelling and filming in St. Petersburg, Russia (VIR), Dundee, Scotland (James Hutton Institute), Montpellier, France (Chr. Hansen) and Lisbon and Braga in Portugal (partners IBET and Biotempo). During the project meeting in Montpellier, France, we carried out additional interviews with key partners and the coordinator from DTU (Denmark). The film is available at https://vimeo.com/193467652 and can be shared and inserted into any website or social media channel.

Biofaction also produced an online/app game called “Berry Maker”, which has been made available at http://berrymaker.modalog.at/index.html, and soon will be available through iTunes and Google Play for mobile devices, once iTunes and Google have completed categorization and age rating for the game. The game mirrors the efforts of the real world to identify genes for useful health-promoting substances in berries

and to transfer them to microorganisms so they will produce these compounds in fermenters. As the CEO, the aim is to maximize the quarterly profits of your company. That way the player can stay ahead of the competition and become the strongest competitor on the market. After every 90 in-game days the player can upload their quarterly profit figures to the high-score board.

Figure 11: The logo of the Berry Maker game

At first, the player needs to buy equipment for the FACTORY, identify and then extract useful genes in berries (collected from all over the world) in the LABORATORY, and finally transfer these genes to microorganisms and start to produce a soft drink in the FERMENTER. In the fermenter the microorganisms will then produce healthy ingredients, which, back in the factory, the player can fill into bottles and sell to consumers. Different combinations of berries, genes, microorganisms and ingredients, as well as costly resources, such as feed for the microorganisms, heat and service personnel, need to be optimized to make the startup company an economic success.

As the company grows from a start-up to a large corporation, the player experiences different challenges. At first, the player has limited financial resources to buy equipment and needs to establish a positive cash flow. Later on, the focus will be on increasing the efficiency of the production process by constantly searching for the right selection of berries, genes, and microorganisms.

Berry Maker favors the strategic and forward-looking player who keeps the laboratory, fermenters and factory working well. The player has to establish and build up the company step by step to create the most effective biotech start-up on the market. The player also needs to understand and react to outside market events that are beyond their control, but that influence the business and require finely-calibrated adjustments. See **Figures 12** and **13** for a brief description of the game elements.

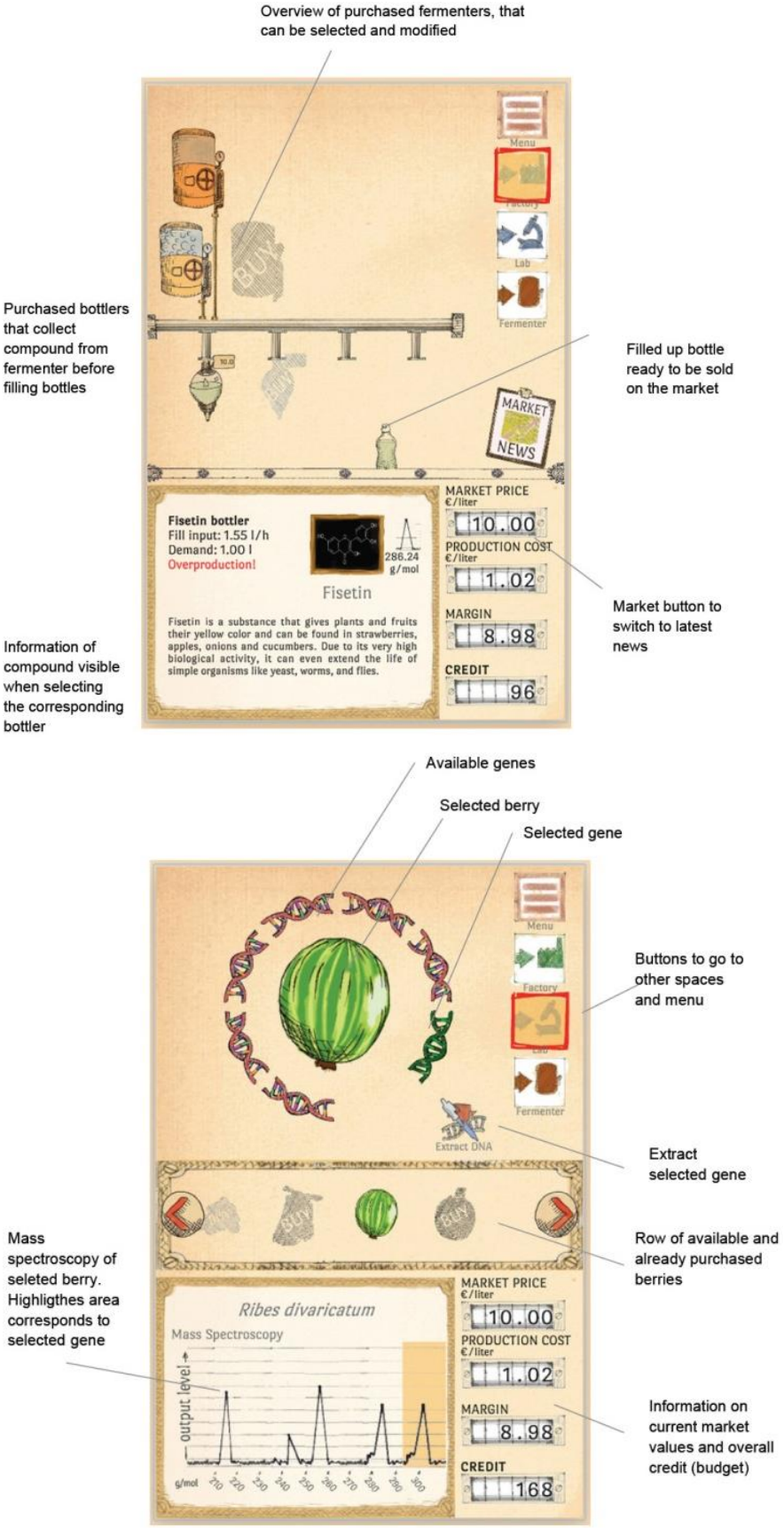

Figure 12: Factory (top) and laboratory (bottom) environments of the Berry Maker game

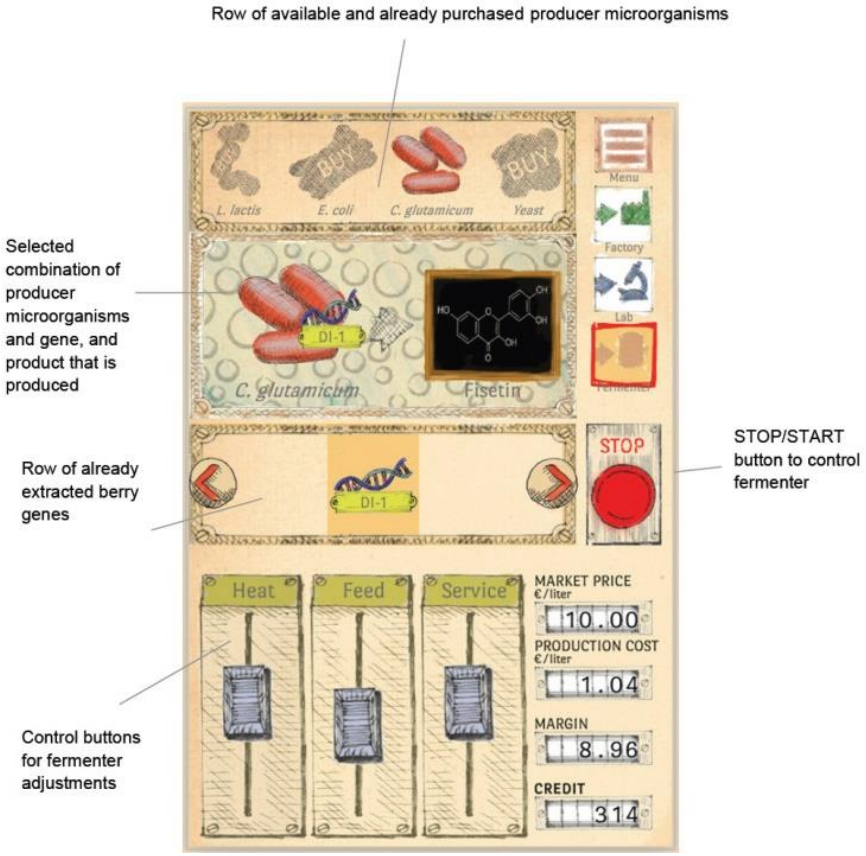

Figure 13: Fermenter settings of the game

Furthermore, the BacHBerry project has produced the largest portfolio of berry transcriptome sequences available worldwide. This community resource represents a valuable contribution to future scientific research on berry species and particularly on studies on polyphenols. Moreover, the project partner Juelich developed a microbial platform strain on the basis of the Gram-positive soil bacterium *Corynebacterium glutamicum*. This strain designated *C. glutamicum* DelAro^4 represents a promising strain for the synthesis of a broad range of small aromatic compounds and is not limited to the production of polyphenols alone. More information on this strain can be found in the publication “Construction of a *Corynebacterium glutamicum* platform strain for the production of stilbenes and (2S)-flavanones” published in the Journal “Metabolic Engineering” (http://dx.doi.org/10.1016/j.ymben.2016.06.003).

During the project, several computational tools where developed and made freely available to the scientific community. These tools include knowledge made

accessible through the Internet, such as a database, and also software for the analysis of metabolic networks. Individual software packages could be found at:

KiMoSys (http://www.kimosys.org/),
OrfFinder (http://github.com/averissimo/orf_finder),
MassBlast (http://github.com/averissimo/mass-blast/),
GeneExtractor (https://github.com/averissimo/gene-extractor),
SON-EM (http://www.mathworks.com/matlabcentral/fileexchange/49967-son-em),
Dinghy (http://dinghy.gforge.inria.fr/),
OptPipe (https://github.com/AndrasHartmann/OptPipe),
Sasita (http://sasita.gforge.inria.fr),
Totoro and Kotoura (http://hyperstories.gforge.inria.fr/),
MultiPus (http://multipus.gforge.inria.fr/),
MultiOpt (http://web.ist.utl.pt/~susanavinga/MultiOpt/).

In order to further exploit the achieved results, two informal consortia have been established between project partners in order to extend the R&D activities beyond the timeframe of BachBerry:

- Biotempo and Juelich will continue developing a promising strain for production of naringenin. The tools and strategies created within the BachBerry project are currently being used for production optimization using a new strain, developed by Juelich, which has already allowed a 20-fold production increase as compared to the strain tested within the Bachberry project.

- Biotempo, IBET, JHI, Juelich, TUD, and Evolva have established a collaboration aiming at generating an extensive article on the identification of RC-1 as a bioactive compound and its production through microbial fermentation. The production process will be later evaluated for its potential for commercialization.