CeBiTec – Quarterly
Spring 2020

Methanol-Based Sustainable Production of Acetoin

Dr. Marta Irla (a former PhD student of the Wendisch lab, of the CLIB graduate cluster and of the board of CeBiTec), Eivind B. Drejer, Dennis Tin Chat Chan and Prof. Dr. Trygve Brautaset from NTNU (Norwegian University of Science and Technology, Trondheim; Norway) have recently established methanol-based acetoin production by metabolically engineered *Bacillus methanolicus* in a collaboration with Carsten Haupka and Prof. Dr. Volker F. Wendisch (CeBiTec and Faculty of Biology). Acetoin can be used as a food additive, a promoter of plant growth and a precursor for chemical synthesis. Its biosynthesis

was introduced to methylotrophic *B. methanolicus* by heterologous overexpression of *B. subtilis*-derived *alsSD* operon encoding acetolactate synthase and acetolactate decarboxylase, resulting in up to 0.26 ± 0.04 g L^{-1} of (*R*)-acetoin titer in shake flask cultivations. *B. methanolicus* was shown to possess a functionally active glyoxylate shunt, which was relevant for further strain optimization. In order to increase supply of the acetoin precursor pyruvate several approaches were tested and six different genes or operons were expressed in the acetoin producing strains. In particular, heterologous expression of a gene coding for malic enzyme from *Geobacillus stearothermophilus* in combination with the

endogenous isocitrate lyase gene increased acetoin titers 1.6-fold up to 0.42 ± 0.01 g L^{-1} which corresponds to a product yield on methanol of 0.07 g g^{-1}. The engineered *B. methanolicus* MGA3 strain overproduced 4 recombinant enzymes from two different plasmids with two distinct antibiotics resistance markers, demonstrating the increased complexity of metabolic engineering allowed by newly developed genetic tools for this organism. To our knowledge, this is the first demonstration of microbial production of acetoin from methanol.

The groups of Prof. Brautaset and Prof. Wendisch have collaborated on establishing methanol as feedstock for bioprocesses since 2009 due to their common interest in development of alternatives to sugar-based raw materials. The Drejer *et al.* 2020 paper is the most recent result of these groups who have continuously received European funding since more than a decade.

Drejer EB, Chan D, Haupka C, Wendisch VF, Brautaset T, Irla M (2020) Green Chemistry. 2020; 20(3):788-802. doi.org/10.1039/C9GC03950C
(M. Irla, V.F. Wendisch)

From a Research Visit to a Published Orphan Crop Genome Sequence in Eight Months

In July 2019, Christian Siadjeu visited the CeBiTec to learn about plant genome sequencing. Christian is from Cameroon, PhD student in the group of Prof. Dr. Dirk Albach at the University of Oldenburg, and works with trifoliate yam (*Dioscorea dumetorum*). *D. dumetorum* is a tuber crop plant grown in parts of Africa as a staple food. After intensive discussion of Mr. Siadjeu with researchers from the GGP (Genetics and Genomics of Plants) group we decided to cooperate with the goal of generating a high contiguity genome sequence of *D. dumetorum*. The CeBiTec infrastructure regarding Oxford Nanopore and Illumina sequencing as well as the computational resources, together with the skills of CeBiTec coworkers, allowed to produce and analyse the data quickly. Material and competent

yam specific knowledge was provided from Oldenburg. The result has just been published in "Genes" in a special issue designated "Recent Advances in Orphan Crop Genetics and Genomics" under the title "High Contiguity *de novo* Genome Sequence Assembly of Trifoliate Yam (*Dioscorea dumetorum*) Using Long Read Sequencing". The genome sequence generated as well as the gene annotation data have the potential to provide the basis for new ways to breed with *D. dumetorum*, for example by allowing to set up specific genetic markers, by making candidate gene approaches possible and by providing access to comparative genomics.

Siadjeu C, Pucker B, Viehöver P, Albach DC, Weisshaar B (2020) Genes 11(3): 274. doi.org/10.3390/genes11030274
(B. Weisshaar)

Animal, Fungi, and Plant Genome Sequences Harbour Different Non-Canonical Splice Sites

Splice sites determine both ends of introns in eukaryotic transcripts. These splice sites are recognised by spliceosomes, complexes of proteins and RNAs, which remove the enclosed intron in order to produce mature transcripts. As crucial recognition points, splice sites are highly conserved across all eukaryotic species. However, there is some variation in these splice sites and the modifications are called non-canonical splice sites. Katharina Frey and Dr. Boas Pucker extended previous analyses of non-canonical splice sites to a systematic comparison across all available genome sequences of eukaryotes. One remarkable observation are high frequencies of GA-AG splice site combinations in two phylogenetically isolated animal species. Moreover, numerous fungal genome sequence annotations contain high numbers of CT-AC splice site combinations. In depth analysis revealed that these splice site combinations are most likely caused by annotation artifacts. As CT-AC is the

reverse complement of GT-AG, the canonical splice site combination, genes annotated on the wrong strand might be the explanation. This study supports previous findings of stronger conservation of the 5' splice site compared to the 3' splice site by providing a comprehensive data set. The excellent computational resources of the CeBiTec were crucial to process large numbers of genome sequences and RNA-Seq data sets.
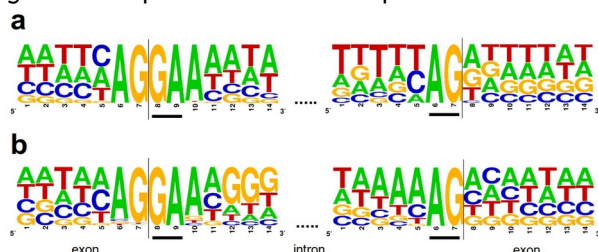
Flanking positions of GA-AG splice site combinations in *Eurytemora affinis* (a) and *Oikopleura dioica* (b). These two species show an exceptionally high frequency of this splice site combination in comparison to all other investigated species. Seven exonic and seven intronic positions are displayed at the 5′ and 3′ splice sites.

(B. Pucker, K. Frey)

A New Project Section Started that Addresses Improvement of Rapeseed Protein Quality

With funding from the Federal Ministry of Education and Research (BMBF) the RaPEQ-2 project is based on an earlier grant but has been implemented as a new project with now ten partners, reinforced orientation towards application, and coordination by the industrial partner *NPZ Innovation GmbH*. Rapeseed is an important oil crop cultivated in Germany. In addition to its high seed oil content, rapeseed also provides a high content of valuable protein. Currently, this protein is used mainly for feed. RaPEQ-2 aims at developing rapeseed protein for human nutrition. To achieve the introduction of high quality rapeseed protein to the German food market, the reduction of the dominant bitter off-taste component in rapeseed
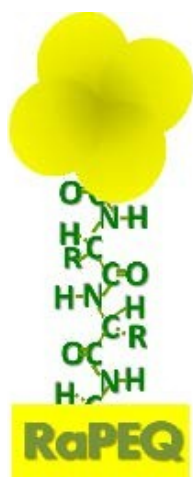
protein isolates, namely kaempferol-3-O-sinapoyl sophoroside (K3OSS), is necessary.

Two main objectives are pursued at CeBiTec in the context of a PhD project carried out by Hanna Schilbert: (I) Identification of key genes for the biosynthesis of K3OSS and (II) detection and characterisation of mutant lines defective in kaempferol biosynthesis. One key gene involved in the synthesis of the kaempferol aglycon is the flavonol synthase (FLS). Thus, the FLS gene family in rapeseed will be characterized with respect to e.g. sequence variations, expression patterns and functional activity. In addition, genes relevant for modification, glycosylation and derivatisation of the kaempferol aglycon will be studied and genotypic variation in these genes will be analysed. Relevant genes identified will be screened for defective alleles in a TILLING population (in cooperation with *NPZ Innovation GmbH*). Such mutants, or homozygous combinations of mutants in single genes, are expected to display reduced levels of K3OSS, which makes them useful for breeding and as improved starting material for rapeseed protein extraction.

(B. Weisshaar)

CeBiTec Researchers Participate in EU Consortium on Pangenomics

Bielefeld University is one of eleven project partners from Europe and North America participating in the EU H2020 RISE project PANGAIA. The full name of the project is 'Pan-genome Graph Algorithms and Data Integration'. It will run from January 2020 to December 2023. The European Union is funding PANGAIA through its Horizon 2020 research framework programme and the University of Milan (Italy) is coordinating the project. Other partners besides Bielefeld University are: the Netherlands Organization for Scientific Research (NWO), Comenius University Bratislava (Slovakia), the biotech companies Geneton (Slovakia) and Illumina Cambridge (Great Britain), the Institut Pasteur (France),

Simon Fraser University (Canada), University of Tokyo (Japan), Cornell University, and Pennsylvania State University (both USA). At Bielefeld University, the project is carried out by Prof. Dr. Alexander Schönhuth (left) and Prof. Dr. Jens Stoye (right).
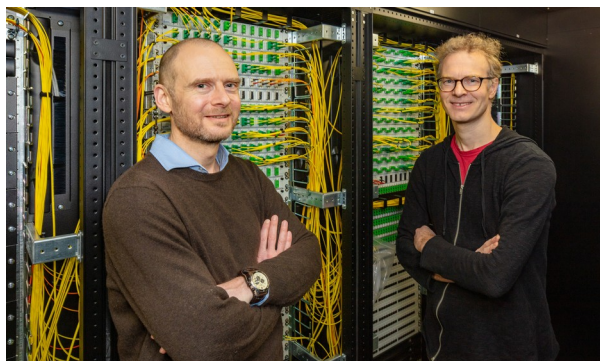
Pangenomics enables a simultaneous, integrated analysis of many strains of the same organism. These can be viruses, bacteria, and sometimes even higher organisms. This makes it possible to highlight the similarities and differences between the individual members. In the case of pathogens, it is often even possible to understand and predict the processes that led to the development of particularly infectious strains. The technology can also be used to detect hereditary diseases in humans or to determine which mutations in a tumour have led to strong, abnormal growth.

Over the next few years, the project aims to develop new algorithms and data structures that will make computer-assisted pangenomics faster and more user-friendly. One goal is to develop algorithms for variation graphs and coloured de Bruijn graphs. Such graphs enable a rapid and high-resolution differentiation of pathogenic and harmless variants of a virus. In particular, they also allow to identify completely novel mutations that may lead to resistance to the usual medications, such as those that have presumably occurred in the coronavirus variant causing the current pandemic.

(A. Schönhuth, J. Stoye)

5th Anniversary of de.NBI

In February 2020, the German Network for Bioinformatics Infrastructure – de.NBI celebrated its 5th anniversary at the *Seminaris CampusHotel* in Berlin. Since its inception the network is operating very successfully and employs more than 250 scientists distributed all over Germany. To mark the occasion of the 5th year anniversary of the de.NBI network, the Administration Office presented a brochure titled *Von der Datenanalyse zum Verstehen komplexer biologischer Systeme* [from data analysis to understanding complex biological systems]. This brochure showcases the contribution of de.NBI as a national bioinformatics infrastructure in the field of life sciences and provides solid information about the advancement of research conducted in the field of big data on a national level. In order to target the largest amount of interested readers, the brochure was presented in the German language. An English version is intended.

In general, the brochure is divided into four large chapters, namely the plant, microbial and medicine bioinformatics followed by the description of the de.NBI network and its activity in the field of service, training, cloud

computing, and industrial partnership. The interplay of research activities from biotechnology, molecular biology, genome research, systems biology, medicine and bioinformatics is presented from different perspectives. Moreover, the brochure also includes an interview with the de.NBI coordinator and the head of the administration office and discusses the structure and organization of the network as well as the many activities that have been carried out since de.NBI establishment.

www.denbi.de/images/Downloads/
deNBI_highlight_brochure_German.pdf
(I. Maus)

A Workshop of *teutolab* biotechnology Wins LeLa-Award

The workshop "Understanding and calculating lactose intolerance" was awarded the second place of the LeLa-award in the category experiment of the year by the *LernortLabor – Bundesverband der Schülerlabore e.V.* [place of learning laboratory – federal association of student labs]. The award ceremony was held in Dresden during the LeLa-convention on March 9, 2020.

The goal of this workshop is to make metabolic processes, which are relevant to the curriculum, tangible and understandable for students. The subject matters of the workshop can be found in the biology curriculum for the upper secondary level in North Rhine-Westphalia under the content field energy metabolism. During the workshop, the students deal with the enzyme kinetics of ß-galactosidase. The focus is on the photometric analysis of the enzyme activity depending on the substrate concentration and the evaluation of the experimentally data. The context of the workshop is lactose intolerance. This context was chosen because of its reference to the living environment of the students.

This project's innovation lays within the workshop's interdisciplinary character. In addition to the biological contents, mathematical competences are taught as well. Mathematics is used as a problem-based and application-oriented tool. In this way, students get a closer insight into the importance of mathematics in a real context. Furthermore, the computer-based evaluation, especially with Excel, is a new aspect of student's lab workshops.

(A. Wenzel)

Upcoming Events

▸ September 14 – 15, 2020 | *Landwirtschafts-zentrum Haus Düsse, Bad Sassendorf*
7th CeBiTec Retreat

▸ September 21 – 23, 2020 | Center for Interdisciplinary Research (ZiF), Bielefeld University
10th International CeBiTec Research Conference (ICRC) on Advances in Industrial Biotechnology

▸ further events are announced on the CeBiTec web page