

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

**Malic acid production by *Aspergillus oryzae***

Christoph Knuf

Department of Chemical and Biological Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2014

Malic acid production by *Aspergillus oryzae*  
Christoph Knuf  
ISBN 978-91-7385-984-4

© Christoph Knuf, 2014

Doktorsavhandlingar vid Chalmers tekniska högskola  
Ny serie nr: 3665  
ISSN 0346-718X

Department of Chemical and Biological Engineering  
Chalmers University of Technology  
SE-41296 Gothenburg  
Sweden  
Telephone +46 (0) 31-772 1000

Cover: Picture of *A. oryzae* conidia taken with an environmental scanning electron microscope.

Printed by Chalmers Reproservice  
Gothenburg, Sweden 2014

# Malic acid production by *Aspergillus oryzae*

Christoph Knuf

Department of Chemical and Biological Engineering  
Chalmers University of Technology

## Abstract

Malic acid is a C<sub>4</sub> dicarboxylic acid which is used as an acidulant in food and beverages. It is also considered as a bio-building block to replace petrochemically derived compounds in the post oil era. This organic acid can be biotechnologically derived from fermentation using renewable feedstocks as carbon source. Aspergilli are among the best producers of organic acid and *A. flavus/oryzae* is the best natural producer of malic acid.

The mechanism of malic acid production in *A. oryzae* was first assessed by transcriptome analysis. A nitrogen starvation response, probably regulated by a transcription factor related to the *S. cerevisiae* Msn2/4 transcriptional activator of stress related genes, was found to result in high malic acid production. Furthermore the pyruvate carboxylase reaction was identified as a metabolic engineering target. This gene, together with the malate dehydrogenase and a malic acid exporter was overexpressed in the strain 2103a-68, which was characterized in a second project. The overexpression led to an 80% increase in yield during the starvation phase ( $1.49 \text{ mol (mol gluc)}^{-1}$ ) and a triplication of the specific production rate. The increase in citric acid production in the engineered strain and its evaluation through model simulations led to the curation of the *A. oryzae* GEM. The existing model was curated with special emphasis on the mitochondrial transport reactions and let to a more defined network around the production of organic acids. Furthermore, the performance of the strain 2103a-68 on xylose as carbon source was evaluated as well and the good results led to the final project of manipulating the carbon source utilization by deleting the carbon catabolite repressor CreA.

This work contributed to the understanding of the regulation of malic acid production. This knowledge was used for the development of *A. oryzae* as an organic acid producer through metabolic engineering. Furthermore, the evaluation of xylose as an alternative carbon source paved the way towards the use of lignocellulosic feedstocks and showed the suitability of *A. oryzae* for the biorefinery of the future.

**Keywords:** systems biology, *Aspergillus oryzae*, malic acid, metabolic engineering, fermentation, C<sub>4</sub> dicarboxylic acids

## List of publications

This thesis is based on the following publications:

- I. **Knuf C**, Nookaew I, Brown SH, McCulloch M, Berry A, Nielsen J. 2013 Investigation of Malic Acid Production in *Aspergillus oryzae* under Nitrogen Starvation Conditions. *Appl Environ Microbiol* **79**(19):6050-6058 doi:10.1128/aem.01445-13 (published)
- II. **Knuf C**, Nookaew I, Remmers I, Khoomrung S, Brown S, Berry A, Nielsen J. (2014) Physiological characterization of the high malic acid-producing *Aspergillus oryzae* strain 2103a-68. *Appl. Microbiol. Biotechnol.*:1-11 (Epub ahead of print, doi: 10.1007/s00253-013-5465-x)
- III. **Knuf C**, Nookaew I, Nielsen J. (2014) The effect of *creA* deletion on the metabolism of a malic acid producing *Aspergillus oryzae* strain (manuscript)
- IV. **Knuf C**, Nielsen J. 2012. Aspergilli: Systems biology and industrial applications. *Biotechnology Journal* **7**:1147-1155. (Review, published)

Additional publications and patent applications not included in this thesis:

- V. Scalcinati G, **Knuf C**, Schalk M, Daviet L, Siewers V, Nielsen J. Modified microorganisms and use thereof for terpene production. United States patent application filed on June 27, 2011 and *PCT Patent Application EP11171612.2* filed June 28, 2011
- VI. Scalcinati G\*, **Knuf C\***, Partow S, Chen Y, Maury J, Schalk M, Daviet L, Nielsen J, Siewers V. 2012. Dynamic control of gene expression in *Saccharomyces cerevisiae* engineered for the production of plant sesquiterpene  $\alpha$ -santalene in a fed-batch mode. *Metab. Eng.* **14**:91-103.

(\*equal contribution)

## **Contribution summary**

- I. Designed the study, conducted the experiments, analyzed the data and wrote the manuscript
- II. Designed the study, conducted the experiments, analyzed the data and wrote the manuscript
- III. Designed the study, conducted the experiments, analyzed the data and wrote the manuscript
- IV. Reviewed the current literature and wrote the manuscript

## **Preface**

This thesis is submitted for the partial fulfillment of the degree doctor of philosophy. It is based on work carried out between 2009 and 2013 at the Department of Chemical and Biological Engineering, Chalmers University of Technology, under the supervision of Professor Jens Nielsen. The research was funded by Novozymes Inc., Davis, California, and the European Research Council.

Christoph Knuf

February 2014

bbl	barrel
BLAST	basic local alignment search tool
bp	base pairs
DW	biomass dry-weight
GEM	genome scale metabolic model
GO	gene ontology
GOI	gene of interest
$\text{h}^{-1}$	per hour
PCR	polymerase chain reaction
TCA	tricarboxylic acid cycle
WTI	West Texas Intermediate
3HP	3 hydroxypropionic acid

## Contents

1	Introduction .....	1
1.1	Aspergilli: Industrial workhorses .....	3
1.1.1	<i>A. niger</i> and citric acid production .....	4
1.1.2	C <sub>4</sub> dicarboxylic acids .....	7
1.2	Aspergilli and molecular biology .....	18
2	Malic acid production .....	21
2.1	Physiology .....	21
2.2	Transcriptome analysis .....	24
3	Engineering of the reductive TCA branch .....	29
4	<i>A. oryzae</i> GEM update .....	35
5	Manipulating the carbon source utilization .....	42
6	Conclusion .....	46
	Acknowledgements .....	49
	References .....	51



# 1 Introduction

This thesis is combining the findings and methods that emerged and are still under development in the fields of metabolic engineering and systems biology. The term metabolic engineering was shaped in the visionary papers of Bailey [8] and Stephanopoulos and Valliano [140]. Bailey defined metabolic engineering as follows:

“Metabolic engineering is the improvement of cellular activities by manipulation of enzymatic, transport, and regulatory functions of the cell with the use of recombinant DNA technology.” [8]

This definition did not change much in the last two decades. Furthermore Bailey explained the usual procedure of metabolic engineering, called the metabolic engineering cycle. This cycle consists of “a genetic change, an analysis of the consequences and a design of a further change”, as depicted in the center of Figure 1. The shown metabolic engineering cycle was extended by the action of systems biology. Systems biology describes the approach of integrating different omics data in order to describe or predict the behavior of a biological system. The function of systems biology in the metabolic engineering cycle is to allow a faster and broader analysis of the system under investigation and through integration of the obtained data predict the outcome of the genetic modification, thereby saving time and money in the wet lab.

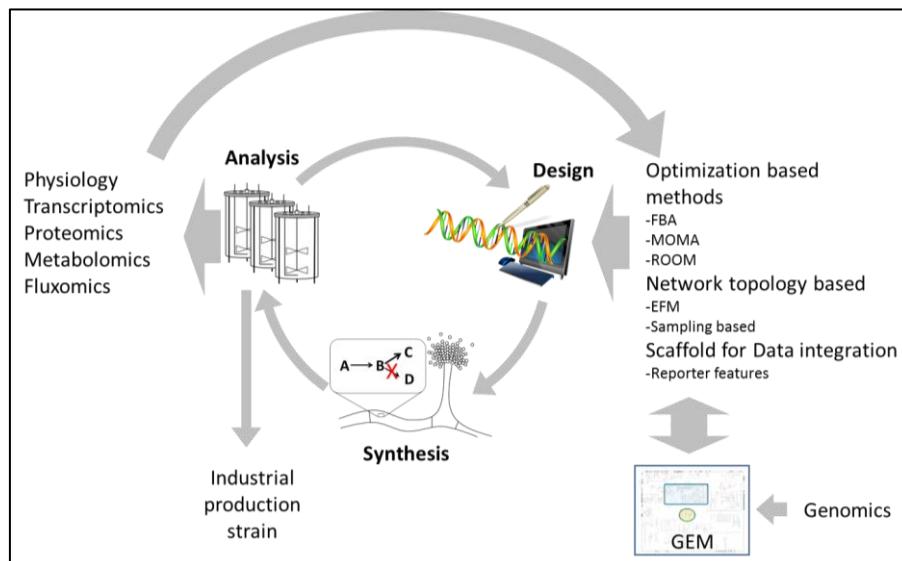


Figure 1: The metabolic engineering cycle (after Nielsen 2001 [96]), extended by the actions of systems biology, which integrates data from –omics studies and thereby facilitate the analysis and design part of the cycle. EFM, elementary flux mode; FBA, flux balance analysis; MOMA, minimization of metabolic adjustments; ROOM, regulatory on –off minimization

The aim of the cycle is an industrial production strain, which matches the industry's goals concerning yields, rates and titers. This strain can then be used in a biorefinery, which is in analogy to the petroleum refinery producing fuels, power and chemicals, just that the

biorefinery is not using crude oil as substrate, but fermentable sugars derived from biomass. As the metabolic engineering approach relies on the genetic modification of microorganisms, this field was initially applied using model organisms like *E. coli* and *S. cerevisiae* [97], which also made them the platform organisms for the first biotechnological production processes made possible through metabolic engineering. The same organisms were as well first used during the initial phase of systems biology, with the first genome scale metabolic models developed for these organisms in the early 2000s (*E. coli* [35]; *S. cerevisiae* [42]). With the availability of the genome sequences of several biotechnologically interesting Aspergilli (*A. nidulans* [43], *A. niger* [112], *A. oryzae* [83]), applying systems biology approaches in the field of filamentous fungi became possible. In case of *Aspergillus oryzae*, the first genome scale metabolic model was published in 2008 [153] and a microarray allowing for transcriptome analysis became available in the same year. Over the years, the molecular biology toolbox for Aspergilli was filled up with optimized transformation protocols and expression/deletion cassettes, so that rational metabolic engineering based on a detailed analysis through systems biology became possible.

## 1.1 Aspergilli: Industrial workhorses

Aspergilli have long been harnessed by mankind for human interests. *A. oryzae* for example, the subject of this study, has been used in the production of Japanese fermentation foods like Miso (soybean paste), Sake (rice wine) and Shoyu (soy sauce) for centuries. A patent on one of the first industrial enzymes produced by *Aspergillus oryzae* was issued in 1895 to Jokichi Takamine who is known as the father of American biotechnology. This diastatic enzyme called Taka-Diastase (US Patent 525,823) was one of the first enzymes marketed in the United States of America. Since then the product range originating from Aspergilli largely expanded. There is the big field of organic acid production with the most dominant example of citric acid, which is naturally produced by *Aspergillus niger* and mainly used as an acidulant in food and beverages. Another major product area comprises of heterologous enzymes like e.g. lipases for the detergent industry produced by *A. niger* and *A. oryzae*. These two fields demand high volume production, especially citric acid with an estimated market capacity of about 1,600,000 t per year [132]. On the other hand Aspergilli have the ability to produce a wide variety of secondary metabolites, of which some are used in the pharmaceutical industry e.g. the cholesterol lowering polyketide lovastatin that was isolated from *A. terreus* [14], but others are extremely carcinogenic, like aflatoxin, which is produced by *A. flavus* [53]. Though there are some Aspergilli that can be potentially harmful for humans, the industrially relevant ones, *A. niger* and *A. oryzae*, are in general considered as safe, as both have a long history of safe use in the fermentation industry [9, 136]. In the case of *A. niger*, some strains have the potential to produce ochratoxin and therefore it is needed to check strains for their ochratoxin production potential before further developing new isolates to production strains [136].

Another advantage of Aspergilli is the fact that they are able to utilize a broad substrate range. In nature Aspergilli live a saprophytic lifestyle, meaning that they are able to thrive on dead biomass. In order to have greatest flexibility Aspergilli are able to efficiently metabolize the monosaccharides abundant in this environment, like glucose, ribose, arabinose, xylose, rhamnose, mannose, galactose and fructose [36, 111]. This is already a big advantage over the model organisms *E. coli* and *S. cerevisiae*, which have to be engineered and evolved in the laboratory to be able to utilize all these sugars. These monosaccharides are not readily available in e.g. plant cell walls, but are incorporated in polymers like cellulose (glucose), starch (glucose) and hemicellulose (xylose, arabinose). For the hydrolysis of those polymers, the Aspergilli possess a big toolbox that enables them to find the right enzyme for breaking most bonds in the biomass. Among these enzymes are glucanases and xylanases, which are mainly responsible for the degradation of the most abundant polymers, cellulose and hemicellulose, respectively.

The availability of systems biology tools, the possibility of metabolic engineering, the current knowledge on large scale fermentation technology, the generally recognized as safe (GRAS)

status of many applications of Aspergilli, their ability to utilize and degrade a great variety of carbon sources and the natural ability of producing acids and enzymes makes them the perfect organism for the biorefinery of the future.

### **1.1.1 *A. niger* and citric acid production**

Citric acid is a well-known organic acid that is extensively used in the food and beverage industry as it combines a pleasant taste with low toxicity and palatability. It serves several functions in the food formulation, like sterilization, flavor fixation and enhancement, bacterial stabilization, and standardization of acid levels. Furthermore it can be used as a chelating agent as it efficiently complexes metal ions. As the acid was initially extracted from Italian citrus fruits, the market became an Italian monopoly. At around 1880 the firm Charles Pfizer developed a chemical synthesis route. In the early 20<sup>th</sup> century, surface culture methods using *Aspergillus niger* were implemented and in the middle of the century further developed towards a submerged process. The approximate yearly production in 2007 was 1.6 million tons. The majority of this is produced via fermentation using *A. niger* [60], but processes with *Yarrowia lipolytica* are reported as well [107].

The foundation for submerged cultures for citric acid production with *A. niger* were laid in 1948 [138], when the media composition promoting citrate accumulation was published. During the years the process was more and more defined and certain conditions for promoting citric acid accumulation were found. Firstly the carbon source should be provided in high initial concentration, as a direct correlation between the carbon source concentration and the specific production rate of citric acid was found [106]. Another important factor is the nitrogen source, which in industrial settings is usually provided through molasses that are used as carbon sources. For most studies in lab scale the medium is containing ammonium salts, like ammonium nitrate or ammonium sulfate [85]. In case of pH an optimum of around pH 2 or lower was found *in vivo* and *in silico* [6]. The aeration has an impact as well, as high aeration is lowering the CO<sub>2</sub> partial pressure in the medium. This leads to a reduction of available substrate for the carboxylation reaction of pyruvate to oxaloacetic acid, which is needed to replenish the carbon in the TCA cycle for high citric acid production. Though a certain amount of CO<sub>2</sub> is needed, sparging with CO<sub>2</sub> in the initial phase of citric acid fermentations has a negative influence on citric acid production [87]. Besides the hitherto discussed influences, trace-metals have to be present only in limiting conditions. Among those are manganese [64], zinc and iron [138]. Besides those external factors, it was reported that the cell morphology has an influence as well. Concerning the question whether pellet or filamentous growth promotes citric acid production the results are contradictory and probably strain specific. It was shown, that high concentrations of spores used for inoculation most likely yield in the dispersed morphology [105]. Common characteristics of the micro-morphology are short, swollen branches with swollen tips [104]. Through high shear forces the number of metabolically active tips can be increased, as

filaments that consist of old, vacuolated and metabolically inactive cells are more likely to break [106].

Though generations of researchers have investigated the interesting phenomenon of citric acid accumulation, Karaffa and co-authors came in 2003 to the conclusion that “only pieces of the puzzle are understood”[60]. The biochemical reactions that are leading towards citric acid accumulation were discovered in the 1950s. It was shown that the carbon is passing through the glycolysis, resulting in 2 moles pyruvate per mole glucose. One mole of pyruvate enters the mitochondrion and acetyl-CoA is formed in the pyruvate dehydrogenase reaction, releasing one CO<sub>2</sub>. In order to reach a molar yield of citric acid per glucose of above 66%, this CO<sub>2</sub> has to be recovered. That reaction takes place in the cytosol, where one mole of pyruvate is carboxylated to oxaloacetic acid, which subsequently is reduced to malic acid. The cytosolic malic acid can then be exchanged for a mitochondrial citric acid molecule that is hereby transported across the mitochondrial membrane. Once the malate is in the mitochondrion it is used to fuel the TCA cycle. It is oxidized to oxaloacetate, which condenses with the acetyl CoA to form citrate. During growth part of the citrate has to be used to generate cytosolic acetyl-CoA for lipid biosynthesis, but during non-growth conditions the molar yield of citric acid on glucose can be one (Figure 2).

By using this route, which recycles the lost CO<sub>2</sub> from the pyruvate dehydrogenase reaction from the mitochondrion in the cytosol, the whole process seems more like a bioconversion of glucose to citrate with maximum yields of about 95%. The above described pathway shows the origin of pyruvate as a direct result of glycolysis. But it was shown that considerable amounts of citric acid must have been formed from previously accumulated glycerol and erythritol, which most likely were a result of carbon overflow during the initial phase of citric acid fermentation [111]. To obtain high yields of citrate on glucose there is a net production of NADH which can be oxidized in the respiratory system. However, this results in ATP production and alternative oxidation of NADH is therefore important for obtaining high citrate yields on glucose [7, 120].

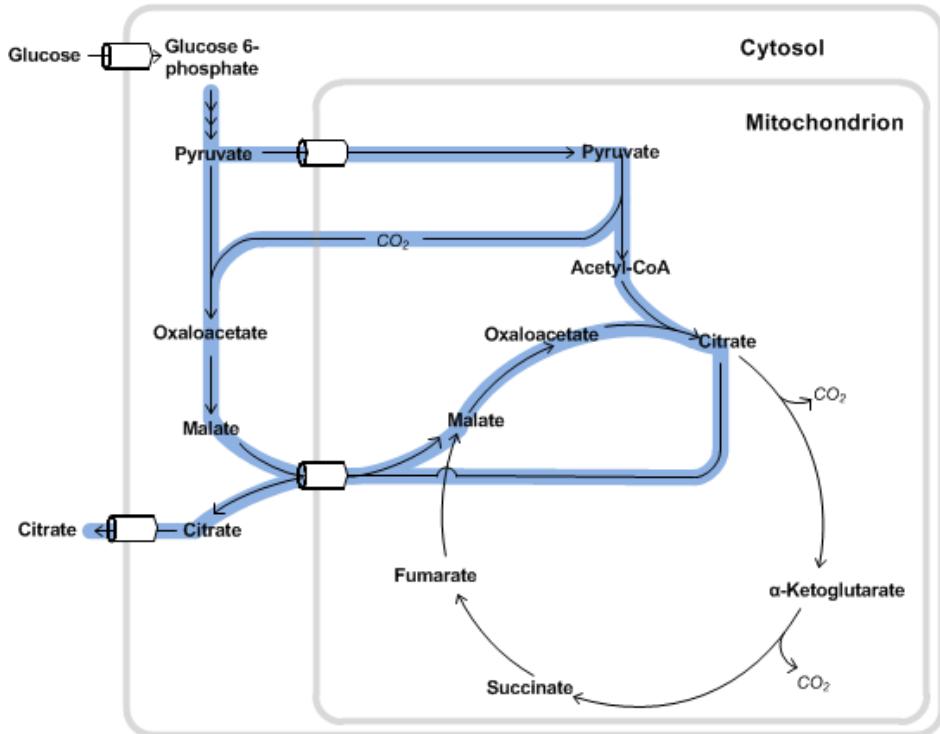


Figure 2: Central carbon metabolism and biochemical reactions leading to citric acid (after [60]).

A modelling approach identified the export of citrate from the mitochondrion as important factor for citric acid production in *A. niger* [44]. It is generally believed that this transport is based on the tricarboxylate transporter (TCT), which in yeast and mammalian cells exchanges cytosolic malate for mitochondrial citrate [38]. Therefore the malic acid accumulation that was measured prior to citrate accumulation in *A. niger* [129] might be a stimulus for the TCT to be more active. As the  $K_m$  for malic acid ( $K_m = 0.25\text{mM}$ ) is 10 times higher than for internal citrate ( $K_m = 0.027 \text{ mM}$ ) [103], quite high cytosolic malate concentrations are needed for efficient export of citrate. Once this transporter is active, it competes favorably with aconitase whose  $K_m$  for citrate ( $3.2 \text{ mM}$ ) is significantly higher than the one of the TCT [48], thereby pulling the citrate out of the TCA cycle without any further constraints needed downstream in the TCA cycle to force citrate accumulation.

This phenomenon was observed for *S. cerevisiae*, where it was shown, that increased titers of dicarboxylic acids in the cytosol positively influence the export of citrate from the mitochondrion [131]. Furthermore, the overexpression of malate dehydrogenase in *S. cerevisiae* increased not only the final malic acid titers significantly, but also increased the citrate titer by 33% [118]. 3D-structural predictions showed that there were similarities especially in the cytosol facing part between the *S. cerevisiae* and the *A. niger* putative TCT [34]. In order to utilize this mechanism for increased citrate production in *A. niger*, de Jongh et al. overexpressed heterologous genes of the reductive TCA branch in the cytosol in order to increase the cytosolic flux towards dicarboxylic acids. It was shown that the insertion of cytosolic malate dehydrogenase improved the production of citric acid and the glycolytic flux

[34]. Through the overexpression of Frds1 and fumRs, the yield could be improved to reach  $0.9 \text{ g Citrate (g glucose)}^{-1}$  in the later stage of the cultivation and even the abundance of  $\text{Mn}^{2+}$  ions did not prevent the accumulation of citrate.

As the pH of the environment is an important factor for citric acid production, the influence of pH changes was investigated using model predictions and transcriptional analysis. The first studies on modelling the citrate production were conducted in the 1990s when Torres used a model of the carbohydrate metabolism of *A. niger* for the prediction of citric acid production [144, 145]. Over the years the modelling approach was optimized with the first genome scale metabolic model of *A. niger* as an important milestone [7]. Using the genome scale modelling approach and combining it with the proton production capability, the production of organic acids was investigated. It was successfully shown that the secreted acids at certain pH values were the most effective for *A. niger* [6]. Furthermore by using transcriptional analysis, it was shown that 109 genes were directly corresponding to pH and candidate orthologues of the Pal/PacC pH controlling pathway known from *A. nidulans* were identified. The findings led to the conclusion, that the aggressive acidification of the microenvironment in combination with the storage of gluconic acid was an evolved strategy among *Aspergilli* in order to outcompete rival microorganisms [6].

### 1.1.2 $C_4$ dicarboxylic acids

The group of  $C_4$  dicarboxylic acids comprises of malate, fumarate and succinate (Figure 3). These acids, known as intermediates of the TCA cycle, are structurally very similar and can easily be interconverted chemically [151].

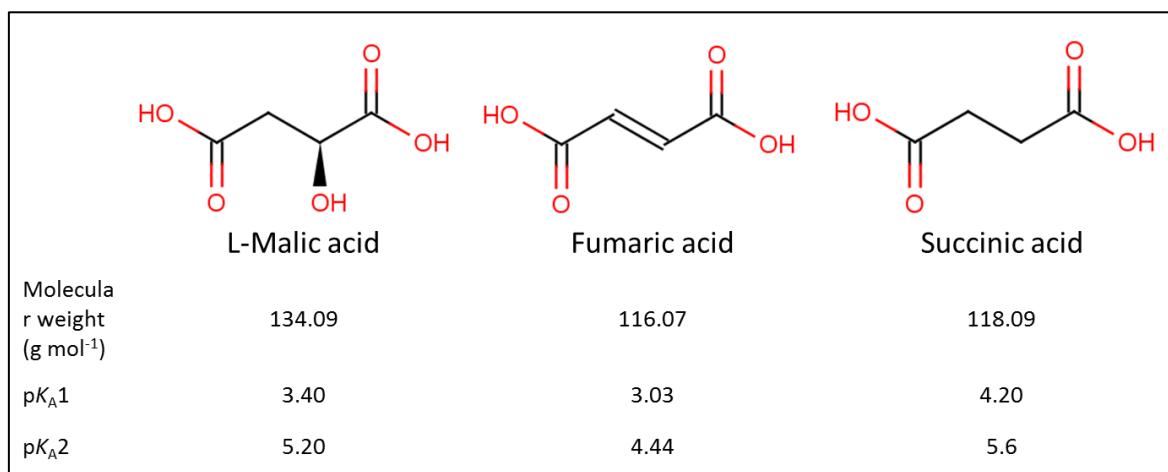


Figure 3: The group of  $C_4$  dicarboxylic acids, L-malic acid, Fumaric acid and Succinic acid, including their chemical properties.

The current way of producing  $C_4$  dicarboxylic acids consists mainly of chemical conversion of petrochemically derived maleic acid or its acid anhydride, maleic anhydride. Succinic acid is derived by catalytic hydrogenation of maleic anhydride to succinic anhydride and

subsequent hydration to succinic acid [73]. In the case of fumaric acid, maleic acid is converted using catalytic isomerization. This fumaric acid can then be converted by hydration to malic acid. The same conversion method can be directly applied to maleic acid to derive malic acid. The drawback of the synthetic method towards malic acid is the production of a racemic mixture, which is not desired for food purposes and it causes problems in further polymerization processes. Therefore a biological production of L-malic acid is preferred. In order to achieve this a common method is the enzymatic hydration using *Brevibacteria* species which are able to convert petrochemically derived fumarate using the fumarase reaction to L-malate and secrete this back into the medium [47]. Another way of producing L-malic acid is fermentation of sugars, which can be used to produce bio succinic and fumaric acid as well and which will be discussed in the next section.

The C<sub>4</sub> dicarboxylic acids have broad application possibilities and overlapping application fields, as they are easily interconvertible. Succinic acid or derivatives thereof are used for example in dairy products and fermented beverages, but can also be applied as specialty chemicals in polymers, food, pharmaceuticals, and cosmetics. Through catalytic processes it is possible to transform succinic acid into 1,4 butanediol, tetrahydrofuran or γ-butyrolactone, which are nowadays mainly derived from fossil resources. Furthermore succinate can be esterified to dimethyl succinate, which is known as an environmentally friendly solvent. Fumaric acid can be used in chemical synthesis, for polyester and other synthetic polymers and resins, as well for production of biodegradable polymers. Specialized applications are in the treatment of the skin condition psoriasis [4] and in cattle feed, where the addition of fumaric acid has achieved a decrease of methane emissions of up to 70% [86]. Malic acid can be used in biodegradable polymers as well, but the biggest market is currently the food industry where it is used as an acidulant and flavor enhancer. Malic acid produced through chemical synthesis is a racemic mixture, but for the flavor enhancement as well as for the polymerization application it is important to have an isomeric homogeneous product, preferably the L- isomer as this is the natural occurring form which can be obtained through fermentation. In a review from 2008 [132], the annual production of C<sub>4</sub> dicarboxylic acids was estimated to range between 10 000 t in case of malic acid, 12 000 t for fumaric and 16 000 t for succinic acid, all produced from petrochemicals. In the same study the annual market volume, if low cost biobased production could be established, was projected to be higher than 200 000 t for fumaric and malic acid and more than 270 000 t in case of succinic acid.

The possibility of microbial production of organic acids, especially the C<sub>4</sub> dicarboxylic acids, has been proven. The general argument that oil is getting scarce and a future society will need alternatives for petrochemical based materials is a generally accepted prophecy and therefore a good reason for funding agencies to sponsor public research in this field. But if scientists want to see their processes emerging in an industrial setting, it has to become interesting for the biotech companies, which means making profit, in order to invest in a process. This profitable efficiency has to happen in the next 20 years, which is the run time

of a patent. Therefore the important point concerning large scale production of bulk chemicals from renewable resources is the short term economic feasibility. The fabrication of a product from renewable chemicals in which the production process even binds CO<sub>2</sub>, which is generally demonized as the climate killer, is without a doubt desirable for our environment and climate and can be achieved through C<sub>4</sub> dicarboxylic acid production. Nevertheless the companies will only be able to sell hundreds of thousands of tons of a bio-based chemical if the price is comparable with the petrochemically derived product.

Many factors play a role in order to reach the bio-based/petrochemical prize equilibrium. On the one hand is the prize for crude-oil which has been rising over the last years. The bio-based equivalent of crude oil is glucose or other carbon sources, as it is usually used as the carbon source for a fermentative process. Therefore the prize of the carbon source is the first position to keep in mind for the cost calculation of a fermentative process. The refining of crude oil towards chemicals is a long optimized and established process and therefore the cost is more or less constant on that side. In case of the renewable building blocks this is the part the scientists have to optimize. There are numerous points to consider, some being the choice of organism, the mode of operation, the down-stream processing or the use of carbon sources.

In an attempt to calculate the feasibility of fumaric acid production, the authors of a study from 1990 came to the conclusion that the process will be viable in case the oil price reaches beyond 61 US\$ bbl<sup>-1</sup>. They assumed a productivity of 1.2 g L<sup>-1</sup> h<sup>-1</sup> and a yield of 0.74 g (g glucose)<sup>-1</sup>. The 2012 average oil prices for WTI (~94.05\$ bbl<sup>-1</sup>) and Brent oil (111.67\$ bbl<sup>-1</sup>) (<http://www.eia.gov/todayinenergy/detail.cfm?id=9530>) were both far beyond that threshold. Furthermore, current technology led to better fermentation performance, which should give an even better advantage for the bio-based production.

A more recent attempt of a techno-economic analysis of white biotechnology products from 2008 by Hermann and Patel [54] came to the conclusion that succinic acid among other bulk chemicals like 1,3 propanediol, polytrimethylene therphthalate (PPT) and ethanol would be economically viable for a crude oil price of 25\$ bbl<sup>-1</sup>. They also calculated a production costs plus profits (PCPP) ratio of the biological process over the petrochemical process of about 100% at a glucose price of 135€ t<sup>-1</sup> and a crude oil price of 25\$ bbl<sup>-1</sup>. That means at those oil and glucose prices the biological and petrochemical way are even at former technology stage. Unfortunately the glucose price elevated to a value in the range of 250.20 to 327.02 US\$ t<sup>-1</sup> (Last 52 weeks, from 15.01.2014 backwards, US sugar #11, <http://www.investing.com/commodities/us-sugar-no11>, 1\$=0.736€), meaning that it doubled. But luckily for the white biotechnology, the current oil price is 4 times higher than the value that was considered in the study. Nevertheless, a thorough calculation taking current state of the art and raw material prices into account is needed for a proper analysis of the opportunities for white biotechnology products like the C<sub>4</sub> dicarboxylic acids.

### 1.1.2.1 Microbial C<sub>4</sub> dicarboxylic acid production

A wide range of microorganisms naturally produces the C<sub>4</sub> dicarboxylic acids (Table 1) and the development of fermentative production processes dates back several decades. Initially researchers looked for natural producers and optimized the cultivation methods for these. During the development of metabolic engineering and molecular biology, the standard model organisms like *E. coli*, *C. glutamicum* and *S. cerevisiae* were targets of metabolic engineering. Heterologous gene expression and blocking of pathways by gene knock-out enabled those organisms to accumulate C<sub>4</sub> dicarboxylic acids. In parallel molecular biology tools were developed for the natural producers and the existing metabolic networks were optimized for organic acid production.

The best natural producers for **succinic acid**, [*Actinobacillus*] *succinogenes* [49], *Mannheimia succiniciproducens* [74] and *Basfia succiniciproducens* [70], were all isolated from bovine rumen, but the species affiliation is still questionable for [*Actinobacillus*] *succinogenes* and *Mannheimia succiniciproducens*. The gram negative, coccoidal, non-motile bacteria *B. succiniciproducens* was taxonomically classified as belonging to the family of *Pasteurellaceae* [70]. For *B. succiniciproducens*, a glycerol based fed-batch process has been developed which allows succinate production at a steady state rate of 0.094 g L<sup>-1</sup> h<sup>-1</sup> and a yield of 1.02g (g glycerol)<sup>-1</sup> for more than 80 days [135]. In wild-type *B. succiniciproducens* acetate and formate were identified as significant carbon sinks and therefore the pathway towards their formation was blocked by deleting the pyruvate formate lyase (*pflD*). This strategy worked out, but the flux was not completely directed towards succinate, but to a higher degree to lactate. This by-product was decreased by deleting lactate dehydrogenase as well. The double deletion strain showed a 45% increase in the molar yield of succinate on glucose (1.08 mol mol<sup>-1</sup>) [12].

As mentioned above, not only natural producers were used for succinate production, but model organisms were extensively engineered in order to secrete organic acids as well. The hitherto most engineered organism for succinate production is a *C. glutamicum* strain that carries 4 deletions and 6 over-expressions. The deletions were aiming at diminishing the secretion of by-products like acetate ( $\Delta pqr$ , pyruvate:menaquinone oxidoreductase;  $\Delta cat$ , Acetyl-CoA:CoA transferase;  $\Delta ackA$ , acetate kinase) and lactate ( $\Delta ldhA$ , L-lactate dehydrogenase). In order to enhance the production of succinic acid, the oxaloacetate pool ( $\uparrow pyc$ , pyruvate carboxylase;  $\uparrow ppc$ , phosphoenolpyruvate carboxylase) was increased, the glyoxylate pathway reconstructed ( $\uparrow aceA$ , isocitrate lyase (ICL);  $aceB$ , malate synthase (MS)), the initial reaction of the TCA cycle ( $\uparrow gltA$ , citrate synthase) enhanced and the export of succinate from the cell ( $\uparrow sucE$ , putative exporter) optimized [166]. All these modifications, together with a dual phase fermentation, characterized by an initial aerobic cultivation in shake-flasks and production in an anaerobic fed batch process, led to a final titer of 109 g L<sup>-1</sup>, a volumetric production rate of 1.11 g L<sup>-1</sup> h<sup>-1</sup> and an average yield of 1.32 mol (mol glucose)<sup>-1</sup>.

Table 1: Overview of natural C<sub>4</sub> dicarboxylic acid producers and their production performance regarding final yields, final titers, specific and volumetric production rates and the substrate used for the respective studies

Acid	Organism	Yield [g g <sup>-1</sup> ]	Titer [g L <sup>-1</sup> ]	r <sub>p</sub> [g (g DW) <sup>-1</sup> h <sup>-1</sup> ]	r <sub>p</sub> [g L <sup>-1</sup> h <sup>-1</sup> ]	C-source	comments
Succinate	<i>Mannheimia succiniciproducens</i> LPK7	0.86	15.4	-	-	Glucose	Batch, MBEL55E delta: IdhA, pflB, and pta-ackA [98]
	<i>Mannheimia succiniciproducens</i> MBEL55E	0.69	-	-	3.9	Whey	Continuous culture[75]
	<i>Basfia succiniciproducens</i> DD1	1.02	5.21	0.375	0.094	glycerol	Continuous culture [135]
	<i>Anaerobiospirillum succiniciproducens</i>	0.99	32			glucose	Non-ruminal, CO <sub>2</sub> sparging [95]
	<i>Basfia succiniciproducens</i> DD1 engineered	0.708	-	1.027	-	glucose	Batch $\Delta idhA \Delta pflD$ [12]
	<i>E. coli</i> SBS550MG	1.14	-	-	1.21	glucose/fructose	$\Delta idhA, \Delta adhE, \Delta iclR, \Delta ack-pta, \uparrow PYC$ <i>L. lactis</i> [148]
	<i>E. coli</i> AFP111-pyc	1.1	99.2	-	1.3	glucose	$\Delta idhA, \Delta pfl, \Delta ptsG$ , [149]
	<i>C. glutamicum</i>	0.87	109	-	1.11	glucose	4 deletions, 6 over-expressions [166]
Fumarate	<i>Rhizopus oryzae</i>	0.78	~25	-	-	glucose	$\uparrow PEPC$ [164]
	<i>S. cerevisiae</i> FMME-001		3.18	-	-	glucose	$\uparrow PYC2, \uparrow RoMDH$ [154]
	<i>E. coli</i> CWF812	0.389	28.2	-	0.448	glucose	Fed batch, 8 deletions + $\uparrow ppc$ [139]
	<i>R. oryzae</i> ATCC 20344,	0.85	92	-	4.25	glucose	Immobilized cells [21]
	<i>Rhizopus arrhizus</i> NRRL2582	0.65		-	-	glucose	20 L tank [125]
	<i>Rhizopus arrhizus</i> NRRL1526	0.8	97	-	-	glucose	[62]
Malate	<i>A. flavus</i> ATCC13697	0.938	113	-	0.59	glucose	[10]
	<i>A. oryzae</i> SaMF2103a-68	1.027	154	-	0.94	glucose	$\uparrow C4T318, \uparrow pyc, \uparrow mdh$ [17]
	<i>Rhizopus arrhizius</i> and <i>Paecilomyces varioti</i>	0.603	48	-	0.34	glucose	[142]
	<i>Monascus araneosus</i>	0.372	28	-	0.23	glucose	[81]
	<i>Schizophyllum commune</i>	0.357	18	-	0.16	glucose	[61]
	<i>Zygosaccharomyces rouxii</i>	0.387	75	-	0.54	glucose	[141]
	<i>Saccharomyces cerevisiae</i>	0.313	59	-	0.19	glucose	$\uparrow mae1, \uparrow pyc, \uparrow mdh3$ [161]
	<i>E. coli</i> XZ658	1.057	34	-	0.47	glucose	11 deletions, Two stage fermentation [165]
	<i>Toluopsis glabrata</i> T.G-PMS	0.19	8.5		0.18	glucose	$\uparrow RoPYC, \uparrow RoMDH, \uparrow SpMAE1$ [26]

The first fermentation processes for the production of **fumaric acid** were established in the 1940 in the US, but they were soon after replaced by chemical synthesis [126]. But as the oil prizes increase significantly and rising environmental awareness called for sustainable production processes, the interest in fermentation processes for fumarate were revived [46]. *Rhizopus* species are known to be the best natural producers for fumaric acid, with *Rhizopus oryzae* being the best reported producer so far [21]. Metabolic engineering approaches helped to increase the production of fumaric acid in comparison to the wild-type strain, overexpression of *pepc* in *Rhizopus oryzae* improved the fumaric acid yield (0.76 g g<sup>-1</sup> vs. 0.62 g g<sup>-1</sup> WT), whereas *pyc* overexpression hampered cell growth and decreased fumarate yield. Instead the malate yield in the latter strain was increased 3 fold. And of course the model organisms *E. coli* and *S. cerevisiae* were metabolically engineered for fumarate production [139, 154].

### 1.1.2.2 *A. flavus / oryzae* and malic acid production

Malic acid was first isolated in 1785 by carl Wilhelm Scheele [88] from unripe apples, hence the name malic from the latin word for apple, *malum*. Since then it was found in many living cells, as it is an important intermediate in cellular metabolism and a constituent of the TCA cycle. Four major metabolic routes exist towards malic acid (Figure 4). The first one involves the carboxylation of pyruvate to oxaloacetate and subsequent reduction to malate, which leads to the highest theoretical yield of 2 moles per mole glucose. The second route involves the classic TCA cycle. As two carbon dioxide molecules are cleaved off the six-carbon backbone of citrate during its course through the oxidative TCA cycle, the theoretical yield drops to only 1 mole per mole glucose. The third possible pathway utilizes the reactions of the glyoxylate shunt. This can either be cyclic in case the oxaloacetate is replenished by malate and result in a yield of 1 mol mol<sup>-1</sup> glucose or non-cyclic, which leads to a maximum yield of 1.33 mol mol<sup>-1</sup>, through replenishing the oxaloacetate by pyruvate carboxylation.

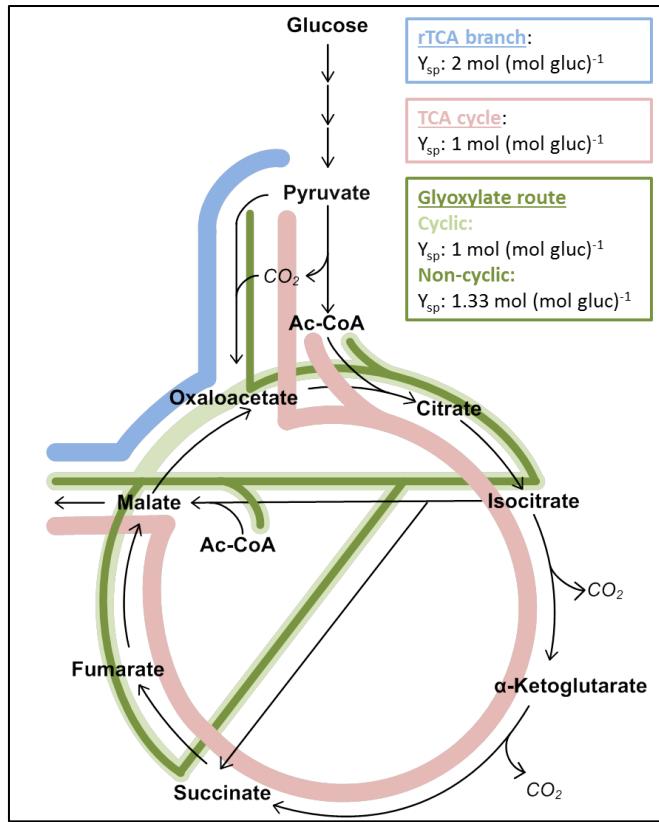


Figure 4: Possible pathways towards malic acid from glucose, the reductive branch of the TCA cycle, the oxidative TCA cycle and two versions of the glyoxylate cycle, cyclic and non-cyclic. Ac-CoA: Acetyl-CoA;  $Y_{sp}$ : maximal malic acid yield per substrate

Probably the first mentioning of malic acid as a product of microbial fermentation dates back to the year 1924 (Figure 5) when Dakin noticed malic acid as a by-product during ethanol fermentation [31] and also gave a hint towards the conditions for malic acid accumulation in microorganisms: "Some degree of nitrogen starvation seems to favor the production of malic acid." In the following years malic acid was mainly mentioned in physiological studies of cell metabolism, which led to the conclusion that malic acid is an important constituent of the TCA cycle [67]. In the early 1960s industrial interest increased and a patent was filed on microbial malic acid production [1]. Several Aspergilli were screened and *Aspergillus flavus* was found to be the best producer in an initial screening. Furthermore, the limitation of nitrogen was applied as well, with  $(\text{NH}_4)_2\text{SO}_4$  at a concentration of 0.2% as the best ammonium salt, yielding in malic acid concentrations of  $30.4 \text{ g L}^{-1}$  after 7 days and up to  $38.2 \text{ g L}^{-1}$  after 9 days of cultivation. The highest reported production rate was  $3.28 \text{ mmol L}^{-1} \text{ h}^{-1}$ , which was achieved with *A. parasiticus* after media optimization and a subsequent feed of  $\text{CaCO}_3$  [1]. A lot of research involving *A. flavus* was conducted in the late 1980s. This not only continued with further process optimization, but also elucidated the pathways leading to malic acid in *A. flavus* during intensive physiological studies. The first publication in the series from the Goldberg lab dealt with the biochemical aspects of acid biosynthesis.

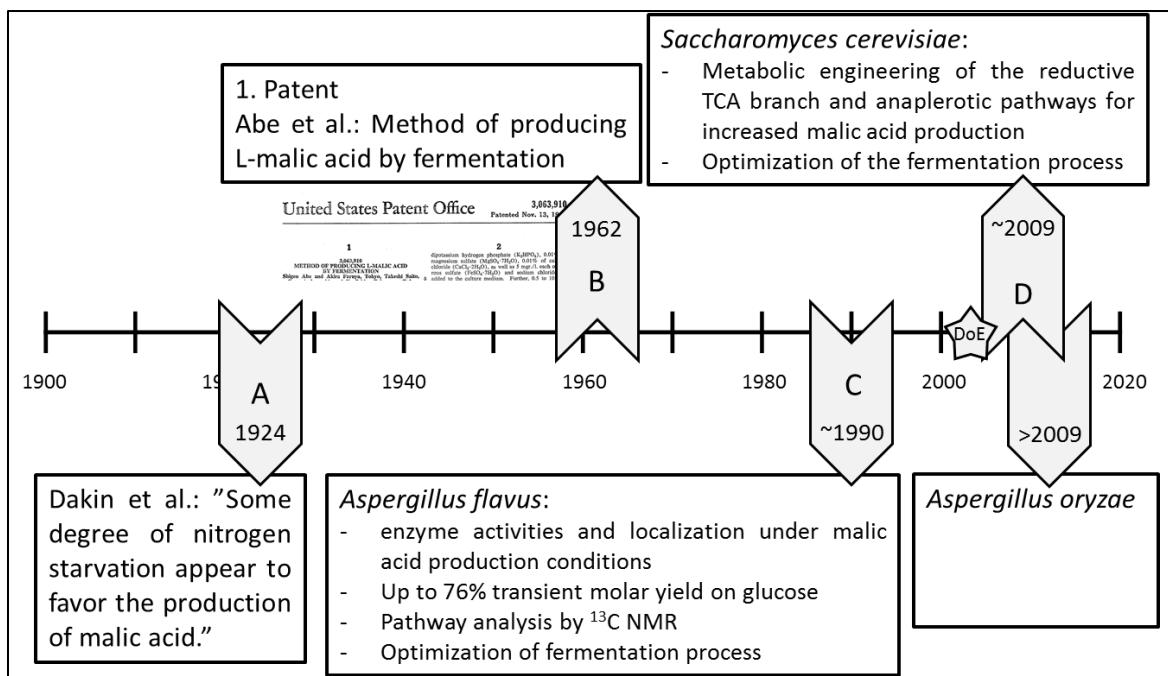


Figure 5: Historical timeline of research activities related to microbial malic acid production. A: [31]; B: [1]; C: [10, 13, 113, 114, 115, 117]; DoE: [151]; D: [160, 161, 162, 163]

In fermenters an overall molar yield of all  $\text{C}_4$  dicarboxylic acids of  $0.68 \text{ mol mol}^{-1}$  glucose was obtained. In addition, it was shown that malate dehydrogenase (Mdh) activity increased up to 10 fold and that this increase is due to *de novo* protein synthesis [117]. This led to the conclusion that malic acid production in *A. flavus* is a result of significant changes in the metabolic network due to nitrogen starvation and that the reductive cytosolic branch of the TCA cycle is carrying the main flux towards malic acid. This result was further confirmed by  $^{13}\text{C}$  NMR measurements, which confirmed that the labeling of  $[1-^{13}\text{C}]$  glucose was mainly incorporated in the third position of malate. This result proved, that malate originated directly from oxaloacetate and did not take the route through the TCA cycle [113]. Another indication for that was given after further process optimization, which resulted in overall  $\text{C}_4$  dicarboxylic acid yields of  $1.55 \text{ mol mol}^{-1}$  glucose (malate yield of  $1.28 \text{ mol mol}^{-1}$  glucose), which is beyond the  $1.33 \text{ mol mol}^{-1}$  that could be achieved through the non-cyclic glyoxylate route. Along with the increased yield, the production rate was increased to reach  $0.59 \text{ g L}^{-1} \text{ h}^{-1}$  ( $4.4 \text{ mmol L}^{-1} \text{ h}^{-1}$ ). The main factors that increased the yield and titer of  $\text{C}_4$  dicarboxylic acids were an increased agitation and  $\text{Fe}^{2+}$  ion concentration [10]. During fermentations yielding in high concentrations of  $\text{C}_4$  dicarboxylic acids it was also observed that crystals were forming during the fermentation process, which consist of calcium salts of the secreted acids [114]. This fact shows the importance of the excess calcium carbonate. Through the formation of calcium acid salts, the product is constantly removed from the solution. This facilitates the transport of the acids out of the cells, as they do not have to pump the acid against a concentration gradient.

In order to better understand the metabolic network around the  $\text{C}_4$  dicarboxylic acids the localization of enzymes was evaluated as well. In the case of *A. nidulans* pyruvate

carboxylase was found to be localized in the cytosol. Malate dehydrogenase and fumarase on the other hand showed unique isoenzymes in the cytosol and mitochondria. The presence of pyruvate carboxylase and malate dehydrogenase in the cytosol led to the postulation of the reductive cytosolic pathway to malic acid [101]. But the investigated *A. nidulans* strain did not produce malic acid under the given conditions. Later the presence of unique isoenzymes for malate dehydrogenase and fumarase were also confirmed in the malic acid producing *A. flavus* strain Kyowa A-114 (ATCC 13697) [113]. The presence of pyruvate carboxylase in *A. flavus* was investigated in a study with several Aspergilli. For the majority of the tested *Aspergillus* species, among them *A. flavus*, pyruvate carboxylase activity was only found in the cytosol fraction. For *A. oryzae* activity was detected in both, the cytosol and mitochondria [13]. The difference in localization of pyruvate carboxylase in *A. flavus* and *A. oryzae* is puzzling, as genome sequencing of both species indicate an extremely close relation between these two and it is even suggested, that the two are ecotypes of which one was selected by mankind (*A. oryzae*) to produce foodstuffs because of its inability to produce the carcinogen aflatoxin [109].

After metabolic engineering [8, 140] took off during the 1990s, pathways towards malic acid were first manipulated in model organisms like *S. cerevisiae* and *E. coli*. The first modification in *S. cerevisiae* was the overexpression of the *FUM1* gene under the control of the strong inducible *GAL10* promoter [116], which enabled *S. cerevisiae* to efficiently convert fumarate into malate. In order to produce malate from glucose the cytosolic malate dehydrogenase (MDH2) was overexpressed [118]. The expression of MDH under the control of the *GAL10* or the PGK promoter led to a 6 to 16 fold increase in cytosolic MDH activity and a 3.7 fold increase of malic acid accumulation in production medium. This relatively simple engineered strain was able to accumulate up to 11.8 g L<sup>-1</sup>. A modern highly engineered *S. cerevisiae* strain overexpressing pyruvate carboxylase, malate dehydrogenase and a malate exporter in a pyruvate decarboxylase deletion background produced 59 g L<sup>-1</sup> with a malate yield of 0.42 mol mol<sup>-1</sup> glucose and a productivity of 0.19 g L<sup>-1</sup> h<sup>-1</sup> [161]. These values were obtained in CaCO<sub>3</sub> buffered shake-flasks. While scaling up the process, the culture pH and CO<sub>2</sub>/O<sub>2</sub> concentrations were identified as key process parameters. An optimization of these parameters led to a 19% improvement of the malate yield on glucose [160]. In the case of *E. coli* metabolic engineering for malic acid production got a kick start by adding modifications on a strain that was already engineered for the production of succinic acid. The final malic acid production strain was able to accumulate malate to a final titer of 32 g L<sup>-1</sup>, in a two stage fermentation process, which is significantly lower than the titer obtained with *S. cerevisiae*, but the yield (1.42 mol mol<sup>-1</sup>) was more than three times as high and the productivity (0.47 g L<sup>-1</sup> h<sup>-1</sup>) twice as high [165]. Another *E. coli* strain was engineered in a way as to generate ATP during the production process by overexpressing the *Mannheimia succiniciproducens* PEP carboxykinase [94], which enable this strain to produce malic acid to a final titer of 9.25 g L<sup>-1</sup> after 12h of aerobic cultivation. Concerning the yield (0.75 mol mol<sup>-1</sup>), this strain could not compare with the previously mentioned *E. coli* strain, but the increase in production rate (0.74 g L<sup>-1</sup> h<sup>-1</sup>) was a significant improvement.

### 1.1.2.3 Downstream processing

The downstream processing in which a biotechnological product is recovered from the fermentation broth is roughly estimated to account for 50% to 70% [11] of the total production cost and therefore has a high impact on the economic feasibility of the whole process. In order to obtain C<sub>4</sub> dicarboxylic acids in high purity from fermentation broth, three main steps are undertaken. In the first step the biomass is separated mainly by membrane filtration or centrifugation. The second step comprises the removal of impurities and primary separation. The last step finally purifies the acid by vacuum evaporation and crystallization. The main approaches for the separation of C<sub>4</sub> dicarboxylic acids are from the fermentation broth include direct crystallization, precipitation, membrane separation, solvent extraction, chromatography, and *in situ* separation [27], these are briefly discussed below using succinic acid as an example.

**Direct crystallization** aims at directly purifying the acid of choice from fermentation broths. When using different temperatures, 60°C for by-product removal through vacuum distillation and 4°C for the actual crystallization of succinic acid, a yield of 75% and a purity of 97% was obtained when using a simulated broth. Applying this method to real fermentation broths, the yield and purity dropped to 45% and 28%, respectively [78, 82]. Another approach used the different dissociation states of the organic acids in the fermentation broth and reached a yield of 70% and a purity of 90% [77]. This method is probably the oldest and a fairly efficient process, but the purity will have to be improved if the product is to be used for polymerization.

The **Precipitation** method uses different ions, mainly calcium, but also ammonium, in order to precipitate the acid from the fermentation broth. This method has been applied successfully in the purification of citric acid and lactic acid. During the process the acid is first precipitated by the addition of Ca(OH)<sub>2</sub>, CaO or ammonium and then the salt is separated by filtration. The filtrate reacts with sulfuric acid and the free organic acid is obtained. The drawback of using calcium ions is the production of gypsum in equal molar amounts to the organic acid [32]. An advantage is the already existing infrastructure and knowledge on this process. The use of ammonium has the advantage that it can be partly recovered after the process, though the recovery by pyrolysis uses a lot of energy. The recovery yield with ammonium is reported to be 93.3% [156].

Another approach is **membrane separation**, including membrane filtration, like microfiltration, ultrafiltration and nano-filtration, or electro-dialysis. Yao et al. were able to obtain a purity of >99.5% with a yield of 75% in a process that applied micro-centrifugation, ultrafiltration and active charcoal adsorption, followed by further purification and final crystallization [155]. Using two stages of electro-dialysis, Zeikus et al. achieved a yield of 60% [159]. Nevertheless, the drawback of this method is the cost of the device and a relatively low yield.

The development of a simple **solvent extraction** method turned out not to be satisfactory [63]. Nevertheless a reactive extraction method in which the product is first converted into a compound without carboxylic groups and then recovered was developed and the use of aliphatic amines yielded in 95% recovery for a simulated solution and 78%-85% using *E. coli* fermentation broth [72]. By using acidification and esterification, recovery yields of 95% were obtained for both model solutions and actual fermentation broth [100].

Ion exchange resin, alumina, silica, and zeolite molecular sieve adsorption are some of the methods used for **chromatography** and recovery yields of >95% were reported for e.g. the resin XUS 40285 and when applying a method involving alkaline-type anion exchange resins the yield was 99% [27]. The challenge in method development here lies in the evaluation of the ideal sorbent that shows a high capacity, a complete and stable regenerability and specificity for the desired acid.

A method that would be desirable to apply in case the product inhibits the growth of the production host is ***in situ* product recovery (ISPR)**. A successful implementation of this strategy, using the anion exchange resin NERCB 09, was reported for an *E. coli* process. The production process was extended from 48h (fed-batch process) to 126h. Though the yield and rate was slightly lower for the integrated fermentation system,  $1.3 \text{ g L}^{-1} \text{ h}^{-1}$  and  $0.52 \text{ g g}^{-1}$  compared to  $1.54 \text{ g L}^{-1} \text{ h}^{-1}$  and  $0.57 \text{ g g}^{-1}$  in the fed batch case, stable production could be achieved over a longer period, leading to the overall production of  $145.2 \text{ g L}^{-1}$  [76].

Though the above mentioned results all refer to succinic acid as the target product, the same methods can be applied to fumaric and malic acid as well. For the fermentation broth containing mainly malic acid obtained after cultivation of *A. oryzae* a procedure was patented that uses concentrating electro-dialysis in a first step and bipolar electro-dialysis in order to convert the acid salt into the free acid [57].

## 1.2 Aspergilli and molecular biology

In order to improve the production performance of a microbial host strain through metabolic engineering one needs molecular biology tools for directed modifications. Whereas the molecular biology studies and the resulting possibilities for the standard microbial hosts like *E. coli* and *S. cerevisiae* are numerous and easy to apply, the development of molecular biology tools for Aspergilli is lagging behind. Nevertheless, some strategies were inferred from other organisms and optimized for an *Aspergillus* host. Important aspects for efficient strain construction are the use of the transformation strategy, marker genes, plasmid availability and promoters and terminators. In order to introduce genetic material into the cell, different **transformation** strategies are available, among these are *Agrobacterium* mediated transformation [91], electroporation [25], biolistic methods [41] and probably mostly used protoplast mediated transformation [50].

In order to select for positive transformants **selection markers** are needed. Those markers can be divided into dominant and auxotrophic selection markers. Frequently used dominant markers are the prokaryotic markers conferring resistance against hygromycin B (*hph*, *E. coli*) [121] and phleomycin (*ble*, *Streptoalloteichus hindustanus*). Another dominant nutritional marker which can be used to select not only for the existence of a marker, but also for the frequency of integration is acetamidase (*amdS*, *A. nidulans*) [29]. Among the auxotrophic markers, orotidine-5'-phosphate decarboxylase (*pyrG*, *A. niger/oryzae*) [52, 84] and orotidine-phosphoribosyl transferase (*pyrE*, *A. niger*) [150] are well known and resemble the *URA3* auxotrophic marker from *S. cerevisiae* [15]. Both enzymes are part of the pyrimidine biosynthesis pathway and thereby involved in the *de novo* UMP biosynthesis and *de novo* pyrimidine base biosynthesis, which makes mutants auxotroph for uracil or uridine. Transformants can be selected on uracil-lacking minimal medium. Another advantage of using *pyrG* or *pyrE* markers is that they can be counter-selected for by 5-FOA (5-fluoroorotic acid). 5-FOA will be converted to the toxic compound fluorouracil through the pyrimidine pathway, thereby allowing only cells to survive which lack the *pyrG* or *pyrE* gene. The positive and negative selection makes it also possible to rescue the marker and thereby allows recycling of it. In order to do so, there are two strategies available. In the first case about 300 bp of direct repeat are flanking the marker gene. This sequence is a target for spontaneous recombination, which leaves only one copy of the flanking sequence in the genome [68]. In the second approach, the marker gene is flanked by loxP sites (34bp) [40], which recombine after recognition and action of the Cre recombinase. Other than the fairly easy approach of looping out the marker gene through direct repeats, the loxP/Cre system requires additional expression of the Cre recombinase. Though advanced transformation cassettes have been developed [93], it takes additional work and time until the next round of modification can take place. Other markers providing both positive and negative selection are *niaD* (selecting for growth without nitrate, negative selection using chlorate toxicity) [20] and *sC* (selecting for growth without sulfate, negative selection using selenite toxicity) [19].

In order to facilitate the cloning process, a range of **plasmids** was developed as well, which are mainly integrative plasmids, designed for integration into the genome. The drawback of Aspergilli concerning strain construction is their low frequency of homologous recombination. The random integrations of the DNA fragments into the genome make targeted gene deletion and specific integration of overexpression fragments into highly transcribed sequences troublesome. In some strains regularly used for strain construction components of the non-homologous end joining (NHEJ) process were deleted, which led to significantly increased homologous recombination in *Neurospora crassa* and was also adapted for *A. oryzae* [92]. Another obstacle for integration of DNA fragments is the fact that Aspergilli are multinucleate. Staining of conidia of the *A. oryzae/flavus* group showed that they are multinucleate [158]. The appearance of multinuclear cells and the high frequency of random integrations call for an additional confirmation of correct integration of the desired fragment into all nuclei by Southern blot analysis, even after confirmation of the integration of the deletion cassette at the desired locus by analytical PCR. An interesting and fast cloning method was developed and applied in *A. nidulans*. The method is based on USER cloning and allows for expression in a combined targeting-expression cassette. The vector set is designed for defined integration into the genome and expression of the GOI, either under the constitutive *gpdA* or the inducible *alcA* promoter [51].

For the overexpression of a gene or a whole pathway it would therefore be desirable to use episomal plasmids. Unfortunately episomal vectors do not naturally occur in Aspergilli, furthermore it was speculated that Aspergilli do not even possess the enzymatic machinery for handling small circular DNAs [3]. Nevertheless it was shown for *A. nidulans*, that genes could be expressed from episomal plasmids that contain an AMA1 sequence, which represents an inverted repeat of mobile Aspergillus transformation enhancers (MATEs) [2, 45]. Those plasmids are structurally stable and do not recombine with the chromosome. While comparing the transformation efficiencies of these plasmids with integrative plasmids, a 1000 times higher transformation efficiency was reported for the episomal plasmids ( $50.000 \text{ colonies (ng DNA)}^{-1}$  against  $70 \text{ colonies (ng DNA)}^{-1}$ ). It was also shown, that there are about 10 copies of the plasmid per nucleus [3], nevertheless, long term stability could not be proven for these vectors, which makes them less useful for an industrial application in which the productivity is supposed to last for many generations [39].

In order to control the expression values, the choice of **promoter** is important as well. A great variety of promoters has been evaluated and used for protein expression in Aspergilli. On the one hand **inducible promoters** like *PglaA* and *PalcA* are used in case the expression of the desired protein has to happen at a certain stage of the cultivation. The expression of *glaA* is repressed on xylose, but highly induced when cultivated on maltose or starch. A CCAAT box has been identified which led to a gradual increase in transcriptional activity when multiple copies were added in the promoter region [79]. The *alcA* gene expression is

induced by various substrates like e. g. ethanol or threonine and strongly repressed by the presence of glucose. Expression of the *alcA* gene is highly depending on the activator AlcR [71] and a co-inducer. The performance of the *alcA* promoter for high level protein expression was enhanced by expression of AlcR under control of a constitutive promoter and multiple integrations of the *alcR* gene in the expression strain. Furthermore, the inhibition mediated through CreA was reduced through binding site modifications in the promoter [39]. Other inducible promoters are *PalcC* which controls the *A. nidulans* alcohol dehydrogenase gene, *PexlA* from *A. awamori* which is xylose inducible, *A. oryzae*'s thiamine dependent *thiaA* promoter [102] or the promoter controlling *A. niger*'s *sucA*, which is inulin or sucrose inducible. An interesting artificial promoter system is the Tet-On system which is based on the *E. coli* tetracycline resistance operon. It was evaluated in *A. nidulans* and has shown tight regulation, fast response within minutes after addition of the inducer and the ability of being fine-tuned depending on the inducer concentration [90]. On the other hand, **constitutive promoters** allow a stable expression of the gene of interest during the cultivation. Examples for constitutive promoters are the *A. nidulans* *PgpdA* (glyceraldehyde-3-phosphate), *PadhA* controlling expressing of aldehyde dehydrogenase and the mid-level expression promoter of the *tpiA* gene encoding triosephosphate isomerase. Furthermore *A. niger*'s *pkiA* (protein kinase A), the glutamate dehydrogenase A promoter (*PgdhA*) from *A. awamori* and the mitochondrial promoters of *oliC* (ATP synthase) from *A. nidulans* and *A. oryzae*'s *Ptef1* (translation elongation factor 1a) are known to be constitutive promoters.

Promoters that have been successfully used in *A. oryzae* for high protein expression are *PamyA* (*taka*) [143], *PglA* [146], *PsodM* [56] and *Ptef1* [65]. A constitutive promoter used for metabolic engineering of *A. oryzae* is *Ppgk* (phosphoglycerate kinase) [17, 130]. It showed stable and high expression during a 150 h cultivation for malic acid production in a wild type strain and was therefore used for the overexpression of the reductive TCA pathway towards malic acid [17].

## 2 Malic acid production

The first questions posed during my thesis work was **IF** *A. oryzae* is producing malic acid to same high amounts as *A. flavus* under nitrogen starvation and if it does **WHY** is it doing so. The project was therefore kicked off with a detailed study on the physiological response of *A. oryzae* under nitrogen starvation conditions, which were as well suggested for fumaric acid production using *Rhizopus nigricans* [127]. In order to study this, the wild-type strain NRRL3488 was cultivated in batch mode, supplied with a limited amount of either peptone or ammonium sulfate as nitrogen source and glucose in excess. The reason for choosing these two nitrogen sources was that peptone as a complex carbon source is closer to the natural environment of Aspergilli. They grow on decaying biomass, which supplies nitrogen in the form of polypeptides which have to be hydrolyzed by the cells in order to be usable for peptide synthesis. Peptone is available in varying qualities from different sources and is a relatively expensive nitrogen source, which is a disadvantage for the use in industrial large scale fermentations. Ammonium sulfate on the other hand is a defined and cheap nitrogen source.

### 2.1 Physiology

The use of nitrogen as the limiting substrate leads to a fermentation profile which in theory is divided into two stages, an initial stage of biomass formation, until the nitrogen source is depleted in the medium, and a second phase in which the existing biomass is converting the remaining glucose into malic acid and a small fraction into energy for maintenance of the cellular functions. With this setup, the overall malic acid yield is supposed to be increased, as carbon is not “wasted” for the synthesis of cell constituents and the associated increased energy demand. The expected fermentation profile can be seen in Figure 6. This figure also includes the concentration of ammonia measured over the time course of a representative cultivation to show the nitrogen limitation during the second phase (stationary/starvation phase) of the cultivation.

The question if *A. oryzae* is producing malic acid in the same conditions than shown before for *A. flavus* was investigated by cultivating two different wild-type strains in shake-flasks with peptone as nitrogen source and  $50 \text{ g L}^{-1}$  glucose. Both strains were accumulating malic acid in the fermentation broth, whereas NRRL3488 was secreting malic acid at a volumetric rate almost double of NRRL3485 with values of  $0.563 \pm 0.020 \text{ g L}^{-1} \text{ h}^{-1}$  and  $0.299 \pm 0.011 \text{ g L}^{-1} \text{ h}^{-1}$ , respectively. The higher volumetric production rate was also reflected in the final titers of  $38.86 \pm 2.8 \text{ g L}^{-1}$  for NRRL3488 and  $23.12 \pm 0 \text{ g L}^{-1}$  for NRRL3485.

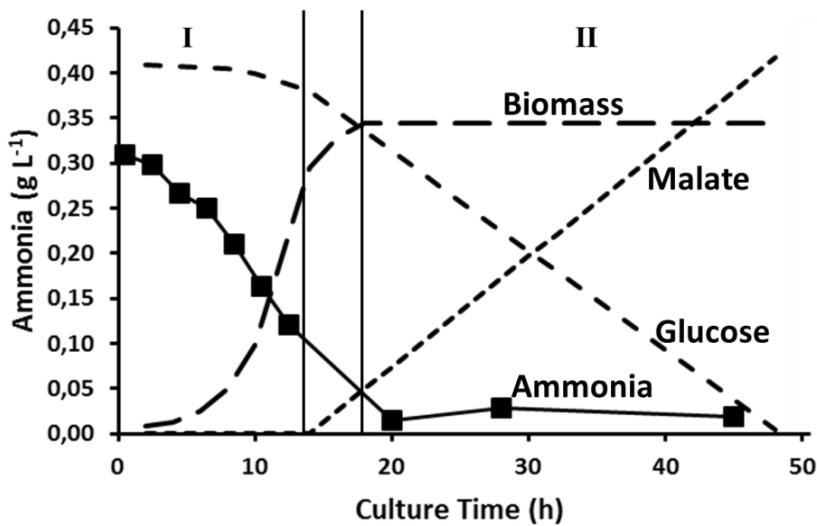


Figure 6: Theoretical fermentation profile for malic acid production and measured Ammonia concentrations of a representative fermentation. I: exponential phase; II: stationary phase

The better producing strain NRRL3488 was chosen for further investigation of the malic acid production mechanisms in *A. oryzae* and therefore cultivated in fermenters with either ammonium sulfate or peptone as nitrogen source and 50 g L<sup>-1</sup> glucose as carbon source. The initial concentrations of the nitrogen source were adjusted in order to obtain a similar exponential growth phase, which led to initial concentrations of 6 g L<sup>-1</sup> peptone and 1.4 g L<sup>-1</sup> ammonium sulfate. The fermentation profiles of the quadruplicate cultivations of each condition can be seen in Figure 7. From the steeper slope of the glucose graph in the peptone fermentation one can see that the glucose uptake rate is significantly higher than in the ammonium sulfate condition (exact data can be seen in Table 2).

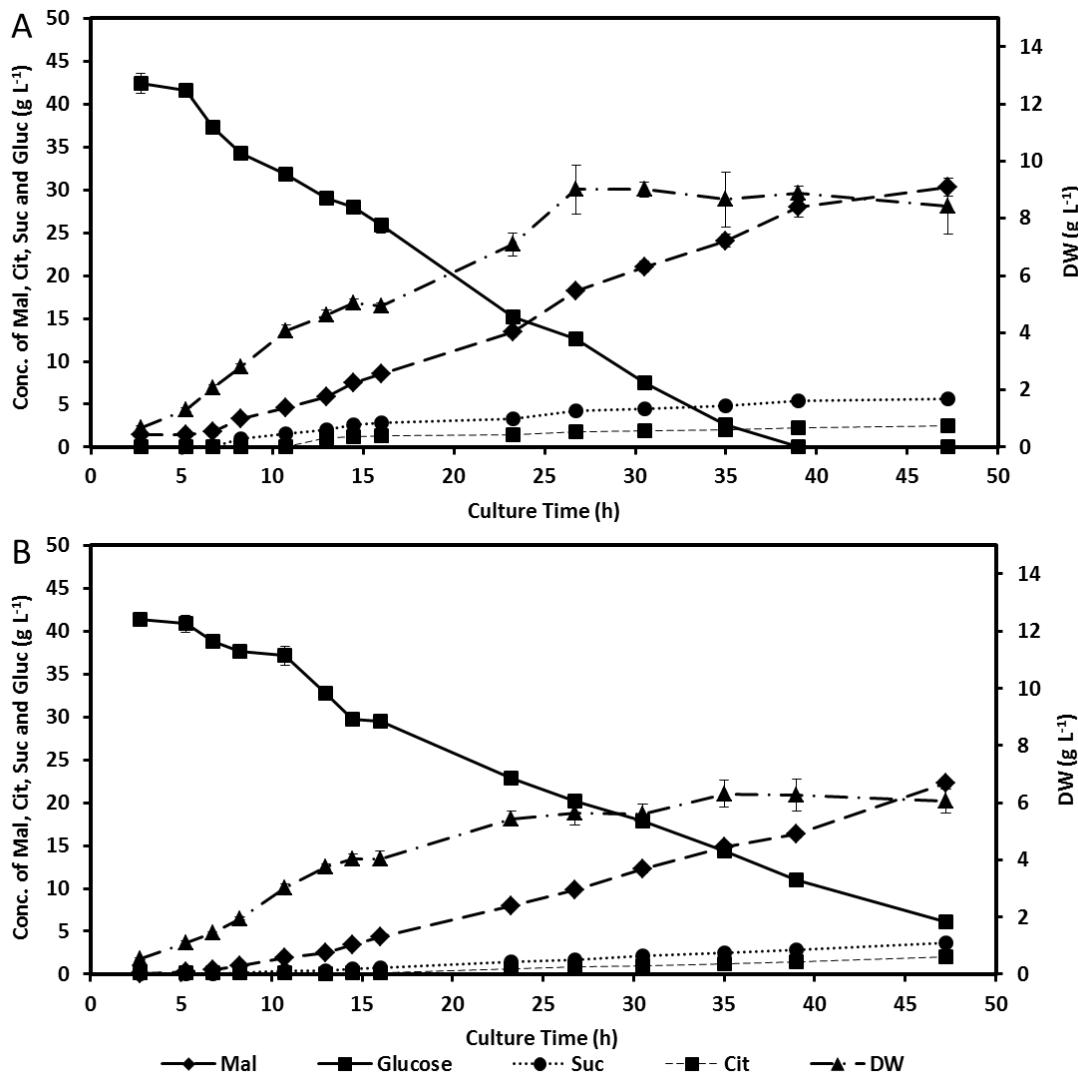


Figure 7: Profiles of malic acid production cultivations of NRRL3488 in MAF medium. The shown extracellular metabolite profiles are averages and standard deviations of 4 reactors each. The upper part shows the fermentation results on peptone (A) and ammonium sulfate (B). Mal, malate; Cit, citrate; Suc, succinate; Gluc, glucose and DW, dry weight.

Malate was detected in significant amounts in both fermentations. The volumetric production rate of malic acid was lower in the ammonium sulfate condition. But it showed the same trend in both fermentations, the volumetric malic acid production rate during the stationary phase is increased in comparison to the exponential phase. When it comes to the molar yields on a glucose basis, the values are almost the same for both nitrogen sources. In both conditions the yield increased from about  $0.33 \text{ mol mol}^{-1}$  to around  $1 \text{ mol mol}^{-1}$ . The second most abundant acid in the fermentation broth was succinic acid, which is following the same trend than malic acid, but in much lower concentrations. As a result of the increased volumetric glucose uptake and malic acid secretion rate, the final titers of malic acid in the fermentation broth were significantly higher for the peptone condition,  $30.27 \pm 1.05 \text{ g L}^{-1}$ , in comparison to the ammonium sulfate cultivations,  $22.27 \pm 0.46 \text{ g L}^{-1}$ .

Table 2: Physiological data for NRRL3488 grown on MAF medium supplemented with either peptone or ammonium sulfate as nitrogen source.

Nitrogen source	final titer of malate (g L <sup>-1</sup> )	Phase <sup>a</sup>	$\mu_{\max}^b$ h <sup>-1</sup>	$r_{\text{malate}}^c$ (mmol L <sup>-1</sup> h <sup>-1</sup> )	$r_s^d$ (mmol L <sup>-1</sup> h <sup>-1</sup> )	Yields on glucose (mol mol <sup>-1</sup> )			
						Citrate	Malate	Succinate	Pyruvate
peptone	30.27±1.05	exp.	0.23±0.01	4.22±0.25	8.10±0.81	n.d. <sup>e</sup>	0.33±0.05	n.d.	n.d.
		stat.	--	6.61±0.57	6.13±0.34	0.03±0.01	0.98±0.13	0.14±0.03	0.02±0.01
ammonium	22.27±0.46	exp.	0.21±0.05	1.59±0.15	3.71±0.44	0.01±0.01	0.34±0.06	0.07±0.02	n.d.
		stat.	--	4.36±0.14	3.92±0.10	0.07±0.01	1.09±0.05	0.20±0.01	0.03±0.00

The numbers stated are means of four individual bioreactors ± standard errors.

a) exp.: exponential growth phase; stat.: stationary phase

b)  $\mu_{\max}$ : maximum specific growth rate

c)  $r_{\text{malate}}$ : specific malate production rate

d)  $r_s$ : substrate consumption rate

e) n.d.: not determined

In comparison to *A. flavus*, the results concerning the ammonium fermentation are comparable in case of the volumetric rates (0.59 g L<sup>-1</sup> h<sup>-1</sup>, *A. oryzae*; 0.58 g L<sup>-1</sup> h<sup>-1</sup>, *A. flavus*), whereas the yield is significantly lower in *A. oryzae* (1.09 mol mol<sup>-1</sup>, *A. oryzae*; 1.26 mol mol<sup>-1</sup>, *A. flavus*). Yield and titer are both higher than for an engineered *S. cerevisiae* strain [161] and concerning *E. coli* strains either higher in yield compared to WGS-10 [94] or higher in the production rate compared to XZ658 [165].

## 2.2 Transcriptome analysis

After verifying *A. oryzae*'s ability to secrete high amounts of malic acid into the fermentation medium, especially under nitrogen starvation conditions, the further underlying mechanisms were investigated that lead to the optimization of the metabolism towards malic acid production. Therefore the transcriptional state of the cells was analyzed using Aspergillus tri-species Affymetrix microarrays.

The first approach to assess the transcriptional state of the cell was the use of the reporter feature algorithm of the Biomet toolbox concerning reporter metabolites and biological process GO-term analysis Figure 8. The reporter metabolites analysis revealed 59 metabolites with significant transcriptional changes around them (distinct directional *P*-value <0.001). Among the up-regulated reporter metabolites, intra- and extracellular ammonia, end-products of purine metabolism, allantoate and urate, and three metabolites of the glutathione metabolism were found. The 51 metabolites connected to down-regulation contained for example metabolites from the amino acid synthesis, TCA cycle metabolites and energy or reduction equivalents.

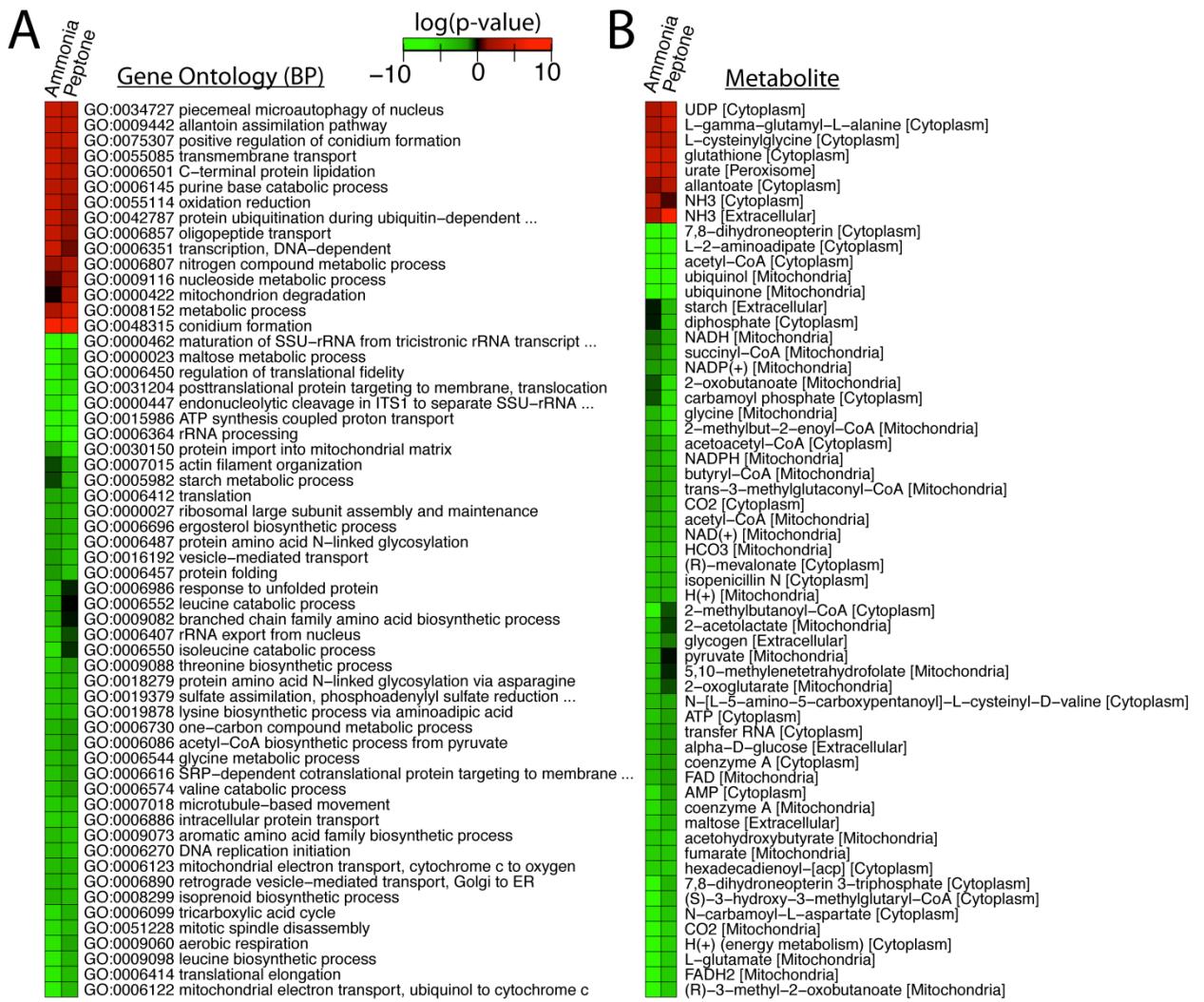


Figure 8: Heat maps of overrepresented GO-terms concerning biological processes (bp)(A) and reporter metabolites (B) depicting changes in transcription from the growth phase to starvation phase. The p-values of the shown GO-terms and reporter metabolites are smaller than 0.005 in either of the conditions (ammonium or peptone). SSU, small subunit; CoA, coenzyme A; SRP, signal recognition particle; Golgi, Golgi apparatus; ER, endoplasmic reticulum; acp, acyl carrier protein

The analysis concerning biological process GO terms revealed 15 GO terms that show a positive distinct directional p-value of less than 0.001 under at least one of the two conditions. Among the GO terms that are characterized by general up-regulation of the corresponding genes under nitrogen starvation are piecemeal microautophagy of the nucleus, purine base catabolic process, protein ubiquitination, or conidium formation. On the down-regulated side, amino acid synthesis related GO terms, as well as protein synthesis, translation/translational elongation, protein folding and intracellular protein transport, are found. As the largest sink of energy was removed through the stop of cellular growth, energy supplying processes like aerobic respiration, mitochondrial electron transport, or ATP synthesis coupled proton transport were correlated with transcriptional down-regulation.

Taken together, both, the reporter metabolite and GO term analysis, indicate that the cells are degrading cellular components and nitrogen containing compounds in order to recycle the nitrogen, furthermore the incorporation of nitrogen was reduced by general down-regulation of the protein production machinery. The up-regulation of the conidium formation and autophagy of the nucleus GO term lead to the assumption that the cells are in a severely stressed condition in which they struggle to survive. In this context it is even more intriguing that they continue taking up glucose from the medium and convert it to malic acid with a yield of about 1 mol mol<sup>-1</sup>. Though the pathway to malic acid must be extremely active during nitrogen starvation conditions, no indication could be found using the initially discussed methods of transcriptome analysis.

Therefore a more targeted approach was chosen and the *P*-values and directions of transcriptional changes were plotted onto the central carbon metabolism network depicted in Figure 9. Following the color code (red & up-arrow, up regulated; green & down-arrow, down regulated), it becomes obvious, that on a transcriptional level the glycolysis is up-regulated. On the other hand, the TCA cycle seems to be down-regulated, which is in accordance with the results of the GO-term analysis. The connection of transcriptional regulation and malic acid production can be established by looking at the values for the pyruvate carboxylase and malate dehydrogenase reaction, which form the cytosolic reductive TCA branch from pyruvate via oxaloacetate to malic acid. The genes encoding for the reductive TCA branch are as well up-regulated, which forms a generally up-regulated direct connection between glucose and malic acid. In order to understand the regulation mechanism behind this, the promoter sequences of the up-regulated genes were searched for conserved sequences.

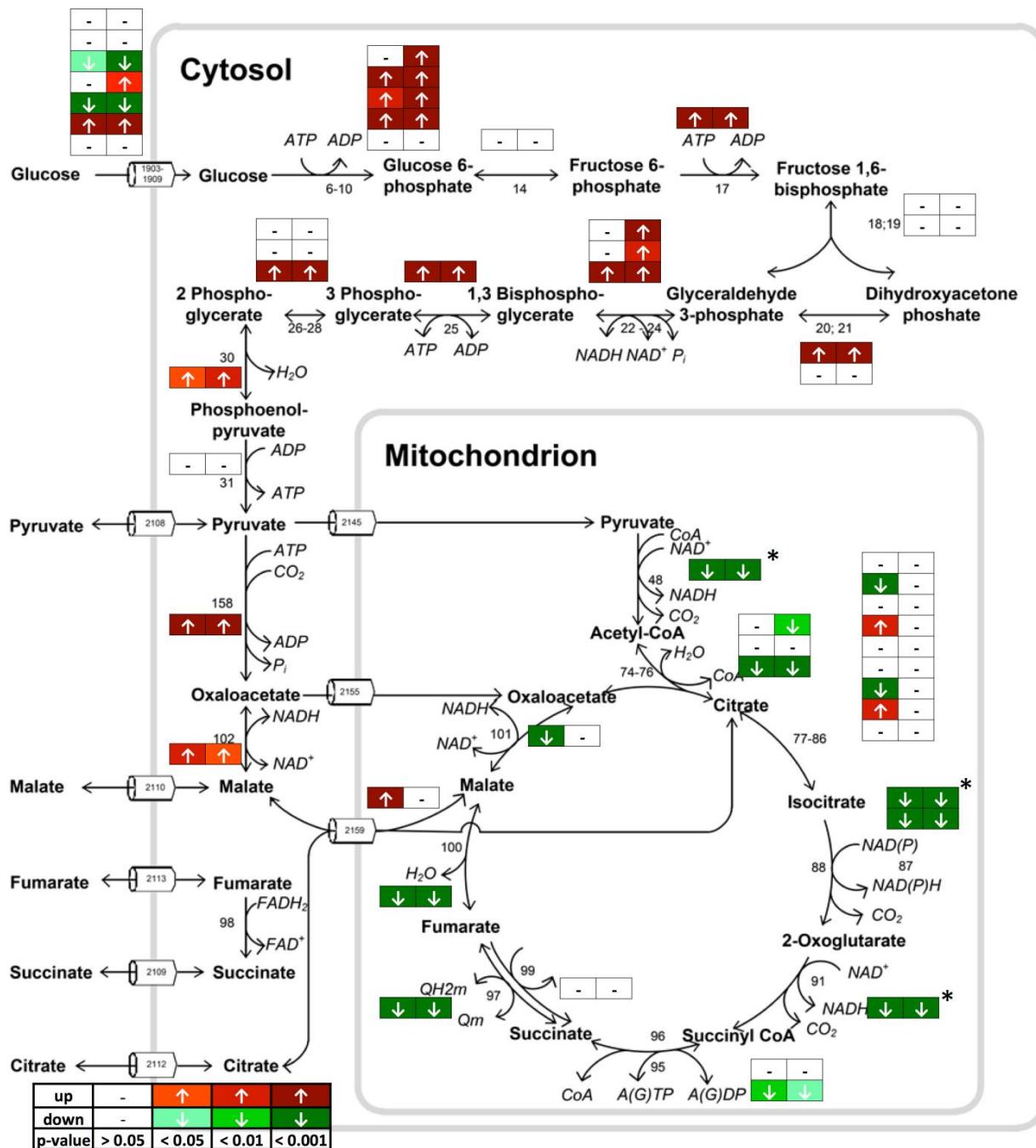


Figure 9: Schematic drawing of the central carbon metabolism of *A. oryzae* and reactions related to the production of malic acid. Transcriptional changes of the genes encoding for enzymes catalyzing the depicted reactions are shown in the boxes next to the reaction. The darker the color, the more significant is the transcriptional change between the stationary and the exponential phase of the cultivation. The direction of changes is indicated in the color and the direction of the arrow in the boxes. Red and up arrow: transcriptionally up-regulated in stationary phase; green and down arrow: transcriptionally down-regulated in the stationary phase. Asterisks, the shown data is taken from the most significantly changed genes of an enzyme complex.

One pattern that came out of the analysis was the motif CCCTC, which showed an occurrence *P*-value of 6.6E-06. This motif is the recognition site of the *S. cerevisiae* transcription factor Msn2/4. This transcription factor is known to be a transcriptional activator of the multi-stress response [134] and concluding from the results obtained here, is

likely to be responsible for the switch between ATP generation under unlimited growth and malic acid production under nitrogen starvation in *A. oryzae*. This direct conversion of glucose to malic acid makes sense from an ecological and evolutionary point of view, as (i) the glucose availability can be reduced, as glucose consumption via the NADH and ATP neutral reductive pathway can continue even without growing. (ii) *A. oryzae* shows optimal growth over a wide range of pH (pH 3-7) [23] and can therefore outgrow competing microorganisms, whose growth is oppressed by low pH. (iii) It has been shown by sequence analysis, that *A. oryzae* shows the largest extension for hydrolytic enzymes, working at low pH, in comparison to *A. nidulans* and *A. fumigatus* [83].

Having established an explanation of the regulatory mechanism leading to malic acid production above, the transcription data was also used for the identification of metabolic engineering targets to further enhance the production capacity. Therefore the transcriptional changes were linked to changes in reaction rates, which were calculated through a random sampling approach, using physiological data from the ammonium fermentation as constraints. One of the reactions identified was the carboxylation of pyruvate to oxaloacetate. In order to further check the feasibility of the overexpression of the pyc gene, the enzyme activities of pyruvate carboxylase and malate dehydrogenase (the two enzymes in the reductive TCA branch) were investigated in cells harvested from shake-flask cultivations during the exponential growth phase and the starvation phase. The activity of pyruvate carboxylase increased from  $0.024 \pm 0.004$  units  $\text{mg}^{-1}$  total protein to  $0.033 \pm 0.007$  units  $\text{mg}^{-1}$  total protein. Malate dehydrogenase activity decreased from  $4.848 \pm 0.828$  units  $\text{mg}^{-1}$  total protein to  $4.304 \pm 0.358$  units  $\text{mg}^{-1}$  total protein. As the increase of pyruvate carboxylase activity was expected to be higher, the possibility of increased protein degradation of pyruvate carboxylase was investigated. As seen in the GO term analysis, the protein ubiquitination was subject to transcriptional up-regulation, therefore the UbPred programme [122] was used to predict the ubiquitination sites of *pyc*, *mdh* and *mae3*. Two sites were predicted with high confidence for *pyc*, whereas none was predicted for *mdh* or *mae3*. As the enzyme activity of Pyc is significantly lower than for Mdh, this step is considered to be the flux controlling step of the reductive TCA branch. Furthermore, the maximum flux that can be achieved with the above mentioned enzyme activity during the stationary phase was calculated to be  $1.01 \text{ mmol (g DW)}^{-1} \text{ h}^{-1}$ . This correlated well with the calculated malic acid production rate during that phase and supports the theory of Pyc being the flux controlling step in the pathway towards malic acid and therefore makes overexpression of *pyc* a promising target.

### 3 Engineering of the reductive TCA branch

After the natural ability to secrete high amounts of malic acid has been proven in the previous section, the engineered *A. oryzae* strain 2103a-68 was further investigated, which carries, among the overexpression of malate dehydrogenase and the malate exporter, the suggested strategy of overexpressing the cytosolic pyruvate carboxylase. The effect of overexpressing the reductive TCA branch and the malate exporter was investigated using glucose as carbon source. As a first step to using lignocellulosic material as carbon source for renewable chemical production, the engineered strain was then further analyzed on xylose containing medium and a mixture of glucose and xylose as carbon source.

The wild-type strain NRRL3488 and the engineered strain 2103a-68 were first compared concerning the expression of the additionally expressed genes and the enzyme activity of pyruvate carboxylase and malate dehydrogenase.

The strain 2103a-68 was created by transformation with DNA fragments containing the genes of interest under control of the phosphoglycerate kinase (*pgk*) promoter [130] and the *glaA* terminator. After transformants grew on selective plates, they were screened for the fastest acidification of the medium and 2103a-68 was the best performing transformant. The integration of all three fragments was confirmed by PCR, but the integration events were not quantified. Therefore a first approach to compare the engineered to the wild-type strain was the quantification of transcripts of each gene of interest and also the enzyme activity of pyruvate carboxylase and malate dehydrogenase.

As the focus is now on the production stage, the relative transcription of the GOIs was determined during the stationary phase (48h). Though all genes were under the control of the same promoter, the transcription level varied in relation to the wild-type strain. In case of *pyc*, the transcriptional level was 3.6 time higher in the engineered strain, *mdh* transcription was increased 9.6 times and the malate exporter was transcribed 7 times more frequent than in the wild-type. Though the transcription relies to a great extent on the location on the genome, one could speculate, that the *pyc* fragment was integrated once, the transporter fragment twice and the *mdh* fragment three times, as the relative transcription increases in increments of about 3.5.

In order to check if the increased transcription of the GOI translates into enzyme activity, pyruvate carboxylase and malate dehydrogenase activities were evaluated during exponential growth and starvation phase in both, the wild-type and engineered strain, from shake-flask cultures. Whereas the enzyme activities were not differing significantly during the exponential growth phase, the activity increased significantly in the starvation phase samples. The pyruvate carboxylase activity was twice as high in 2103a-68 compared to the wild-type and the malate dehydrogenase activity increased even four times, which is consistent with the trend of the relative transcription shown in Figure 10.

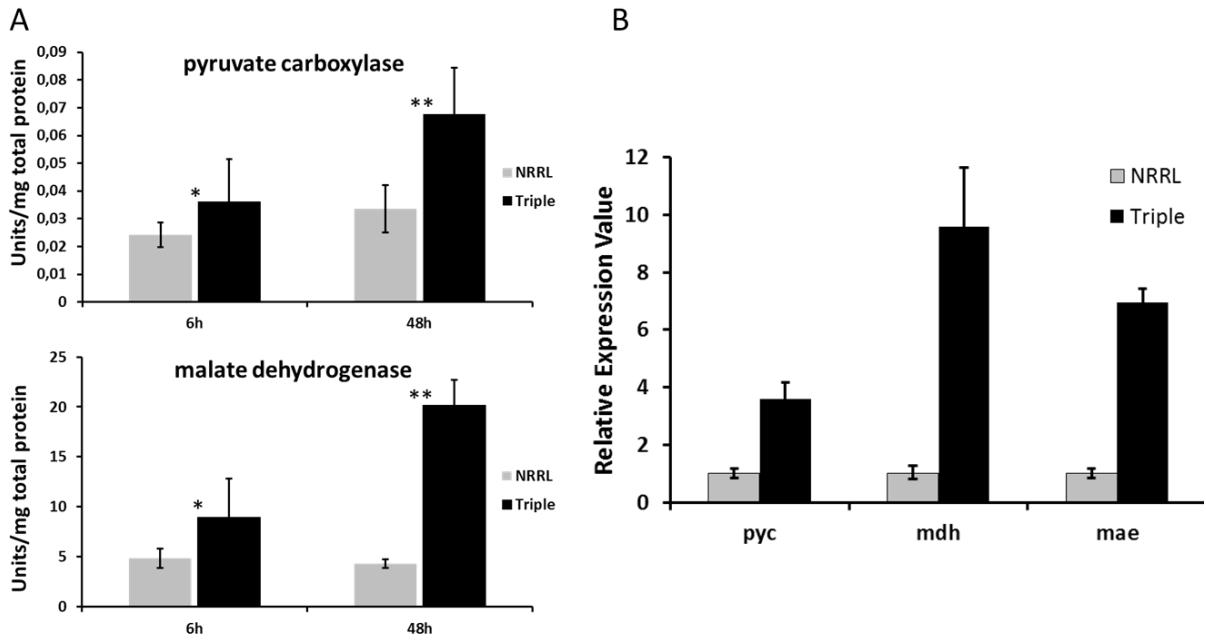


Figure 10: Enzyme activities (A) of pyruvate carboxylase (upper part) and malate dehydrogenase (lower part) and relative expression levels (B) of *pyc*, *mdh* and *mae* after 48h.

In the previous study it was already determined, that the enzyme activity of pyruvate carboxylase is two orders of magnitude lower than for malate dehydrogenase, this ratio is still unchanged and got even slightly worse. Therefore the pyruvate reaction is still considered to be the flux controlling step. And again, the overall specific production rate that was calculated using the enzyme activity (NRRL3488,  $0.842 \text{ mmol (g DW)}^{-1} \text{ h}^{-1}$ ; 2103a-68,  $1.709 \text{ mmol (g DW)}^{-1} \text{ h}^{-1}$ ) correlated well with the physiological data obtained during lab-scale cultivations Table 3. Though it was argued before, that the resulting variation of integration events and the subsequent selection of the best acidifying strain results in the optimal expression ratio of the GOIs [17], the pyruvate carboxylase step still seems to be the flux controlling step, and it would be worthwhile to consider additional integration of *pyc* expression fragments.

As the increased activity of the reductive TCA branch and over-expression of the malate transporter were proven by transcription analysis and enzyme assays, the final proof of the engineering strategy followed in lab-scale fermenters. The engineered strain and the wild-type were cultivated in MAF medium containing  $100 \text{ g L}^{-1}$  glucose as carbon source and ammonium sulfate as nitrogen source. As can be seen in Figure 11, the glucose uptake rate is increased in the engineered strain and the most significant difference is the steep slope of the malic acid graph. The wild-type and the engineered strain produced malate to final concentrations of  $26.77 \pm 0.197 \text{ g L}^{-1}$  and  $66.3 \pm 2.36 \text{ g L}^{-1}$ , respectively.

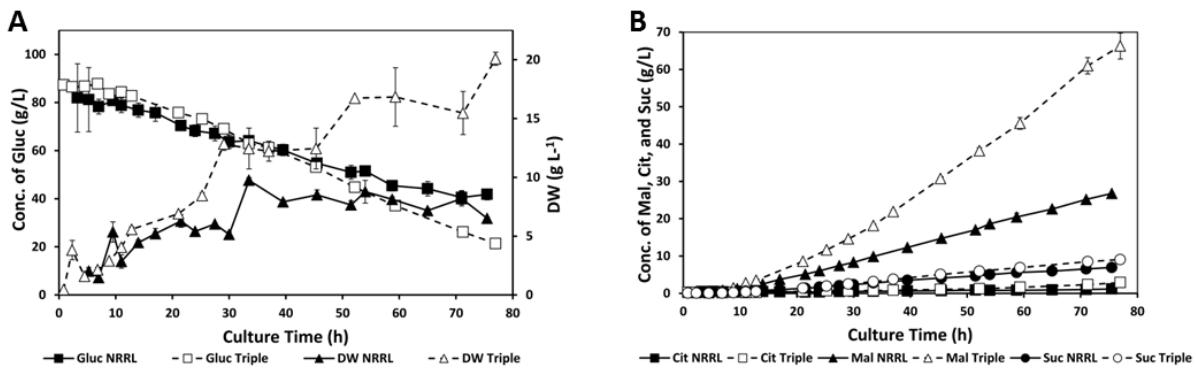


Figure 11: Fermentation profiles of NRRL3488 and 2103a-68 in MAF medium containing glucose from triplicate cultivations. Dry weight and carbon source concentration (A) and extracellular metabolite concentrations (B). Gluc, glucose; DW, dry weight; Cit, citrate; Mal, malate; Suc, succinate; NRRL, NRRL3488; Triple, 2103a-68.

The significantly higher final concentration in 2103a-68 was a result of a 70% increased glucose uptake rate and an 80 % increased malate yield of the engineered strain of  $1.49 \text{ mol (mol glucose)}^{-1}$ . These factors led to a malic acid production rate of  $1.87 \text{ mmol (g DW)}^{-1} \text{ h}^{-1}$  in the engineered strain, which is tripled compared to the parental strain. The malic acid production performance of the engineered strain exceeds all values obtained with other organisms and even beats most strains when comparing the 2103a-68 malic acid values with succinic acid production performance. The only strains performing better are the *Mannheimia* and *Basfia succiniciproducens* wild-type and engineered strains.

As the results obtained with glucose as carbon source were so promising, the performance of the engineered strain using xylose and a glucose/xylose mixture was investigated as well (Figure 12). In the xylose only case, the condition were exactly as mentioned above, just that  $100 \text{ g L}^{-1}$  xylose were used instead of glucose. The malic acid production rates in this setting cannot reach the high values obtained on glucose, but are still higher than for the wild-type strain cultivated on glucose. The cultivation on the glucose/xylose mixture was a first attempt to move towards a process for second generation biochemical production using lignocellulosic feedstocks, which contain glucose and xylose as main sugars. In the fermentation profile in Figure 12 it can be seen that glucose is the preferred carbon source, as it was first taken up and xylose consumption followed. The rates and yields were calculated for the two carbon sources separately. Though the values are not as divergent as shown before for the single carbon source cultivations, the values represent the same trend, the strain performs better on glucose. During the glucose phase (16-35h), the engineered strain showed lower values compared to the single carbon source, but still produces malic acid at a rate of  $1.46 \text{ mmol (g DW)}^{-1} \text{ h}^{-1}$ . The values for the xylose phase (43-66h) on the other hand are slightly increased compared to the xylose only cultivation and reached a malate production rate of  $1.08 \text{ mmol (g DW)}^{-1} \text{ h}^{-1}$ . The interesting part is the carbon uptake rate, which stayed almost constant for the glucose, xylose and glucose/xylose mixture cultivations.

Table 3: Physiological data for NRRL3488 and 2103a-68 (Triple) grown in malic acid fermentation medium with varying carbon sources during acid production phase.

Strain	Carbon-source	Specific rates [mmol (g DW) <sup>-1</sup> h <sup>-1</sup> ]				Volumetric rates [g L <sup>-1</sup> h <sup>-1</sup> ]				Yields [mol mol <sup>-1</sup> C-source]			Yields [mmol Cmol <sup>-1</sup> ]		
		r <sub>PMal</sub> <sup>a</sup>	r <sub>PCit</sub> <sup>b</sup>	r <sub>PSuc</sub> <sup>c</sup>	r <sub>s</sub> <sup>d</sup>	r <sub>PMal</sub>	r <sub>PCit</sub>	r <sub>PSuc</sub>	r <sub>s</sub>	Y <sub>SMal</sub> <sup>e</sup>	Y <sub>SCit</sub> <sup>f</sup>	Y <sub>SSuc</sub> <sup>g</sup>	Y <sub>SMal</sub>	Y <sub>SCit</sub>	Y <sub>SSuc</sub>
NRRL3488	Glucose	0.61±0.01	0.02±0	0.16±0.03	0.73±0.06	0.34±0.06	0.01±0	0.08±0.01	0.55±0.05	0.83±0.07	0.02±0	0.22±0.02	138.33±10.74	3.33±0.53	36.67±3.22
Triple	Glucose	1.87±0.23	0.06±0.01	0.28±0.04	1.26±0.11	1.05±0.13	0.05±0.01	0.14±0.02	0.95±0.08	1.49±0.05	0.05±0	0.22±0.01	247.67±8.27	8±0.79	36.49±1.74
Triple	Xylose	0.89±0.14	0.05±0.01	0.16±0.02	1.27±0.09	0.5±0.08	0.04±0.01	0.08±0.01	0.8±0.06	0.7±0.06	0.04±0.01	0.12±0.01	140.83±12.54	7.57±1.53	24.75±2.04
Triple	Glucose phase (16-35h)	1.46±0.16	0.05±0.01	0.14±0.02	1.23±0.05	0.82±0.09	0.04±0.01	0.07±0.01	0.93±0.04	1.18±0.08	0.04±0	0.12±0.01	196.74±13.27	6.76±0.66	19.32±1.47
	Xylose phase (43-66h)	1.08±0.19	0.03±0.01	0.02±0.01	1.26±0.06	0.61±0.11	0.03±0.01	0.01±0.01	0.79±0.04	0.86±0.11	0.03±0	0.01±0.01	171.97±4.59	5.25±0.19	2.9±0.41

The numbers stated are means of four individual bioreactors ± standard errors.

a) r<sub>PMal</sub>: malic acid production rate

b) r<sub>PCit</sub>: citric acid production rate

c) r<sub>PSuc</sub>: succinic acid production rate

d) r<sub>s</sub>: substrate consumption rate

e) Y<sub>SMal</sub>: Yield of malate per substrate

f) Y<sub>SCit</sub>: Yield of citrate per substrate

g) Y<sub>SSuc</sub>: Yield of succinate per substrate

This is an important trait for the development of new strains for second generation chemical production, which is often limited by low carbon uptake rates. In this connection it is important that *A. oryzae* is not only able to metabolize xylose efficiently, but also to take it up from the medium with the same rate as glucose. Therefore *A. oryzae* does not need to be extensively engineered just to be able to utilize this pentose sugar, as the usual platform organisms like *E. coli*, *S. cerevisiae* and *C. glutamicum* needed to be [18, 80, 137].

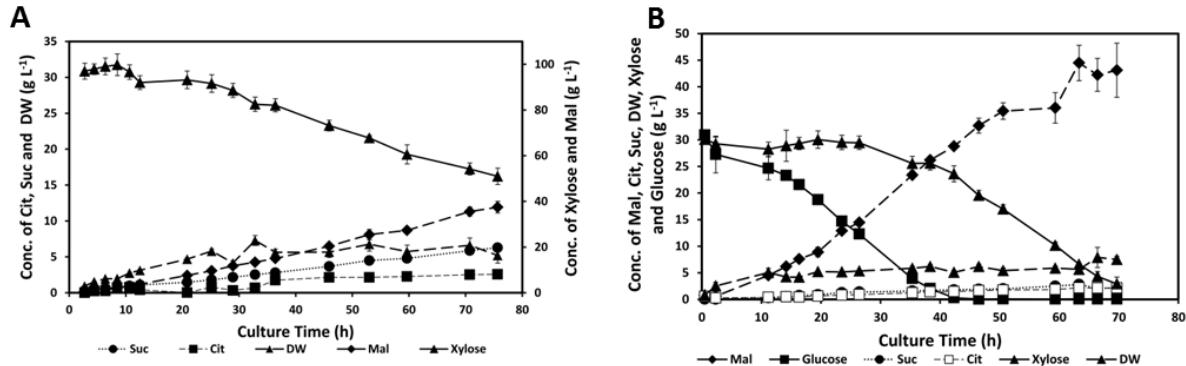


Figure 12: Fermentation profiles of 2103a-68 in MAF medium containing xylose (A) or a glucose/xylose mixture. DW, dry weight; Cit, citrate; Mal, malate; Suc, succinate.

In order to further characterize the engineered strain <sup>13</sup>C flux analysis was used, in which the intracellular fluxes are fitted towards the measured external metabolites and the labeling pattern of amino-acids, which was determined by GC-MS. As this method is limited to measuring *de-novo* synthesized amino-acids, intracellular fluxes could only be determined during the exponential growth phase. In order to obtain labeled samples from exponentially grown cells, shake-flask cultivations were performed and samples were taken after 7.5h of cultivation. In order to be able to fit the fluxes with special respect to organic acids, a compartmentalized flux model needed to be constructed. As the existing *A. oryzae* flux model [133] was not compartmentalized, an *A. niger* model [89] was extended by the reductive TCA branch in the cytosol. The so calculated fluxes show an increase of carbon flow through the rTCA branch in the cytosol and an increased flux of malate and oxaloacetate into the mitochondrion, in order to fuel the TCA cycle (Figure 13). These results show that the overexpression of the rTCA branch already has an impact on the malate production during the exponential growth phase and allows increased malate production compared to the wild-type during cellular growth. Though the efficiency of this strain is not optimal during the growth phase, the parallel growth and increased production allows for use of the engineered strain even in a continuous process. Thereby carbon containing waste streams could be used for the production of renewable chemicals.

Taken together, this strain allows for high level production of malic acid from both, glucose and xylose. Therefore it is very well suited for the biorefinery of the future. Though it already performs very well concerning malic acid production, it might still be optimized through

metabolic engineering. One example might be the engineering of the pyruvate carboxylase step, which still seems to be a flux controlling step.

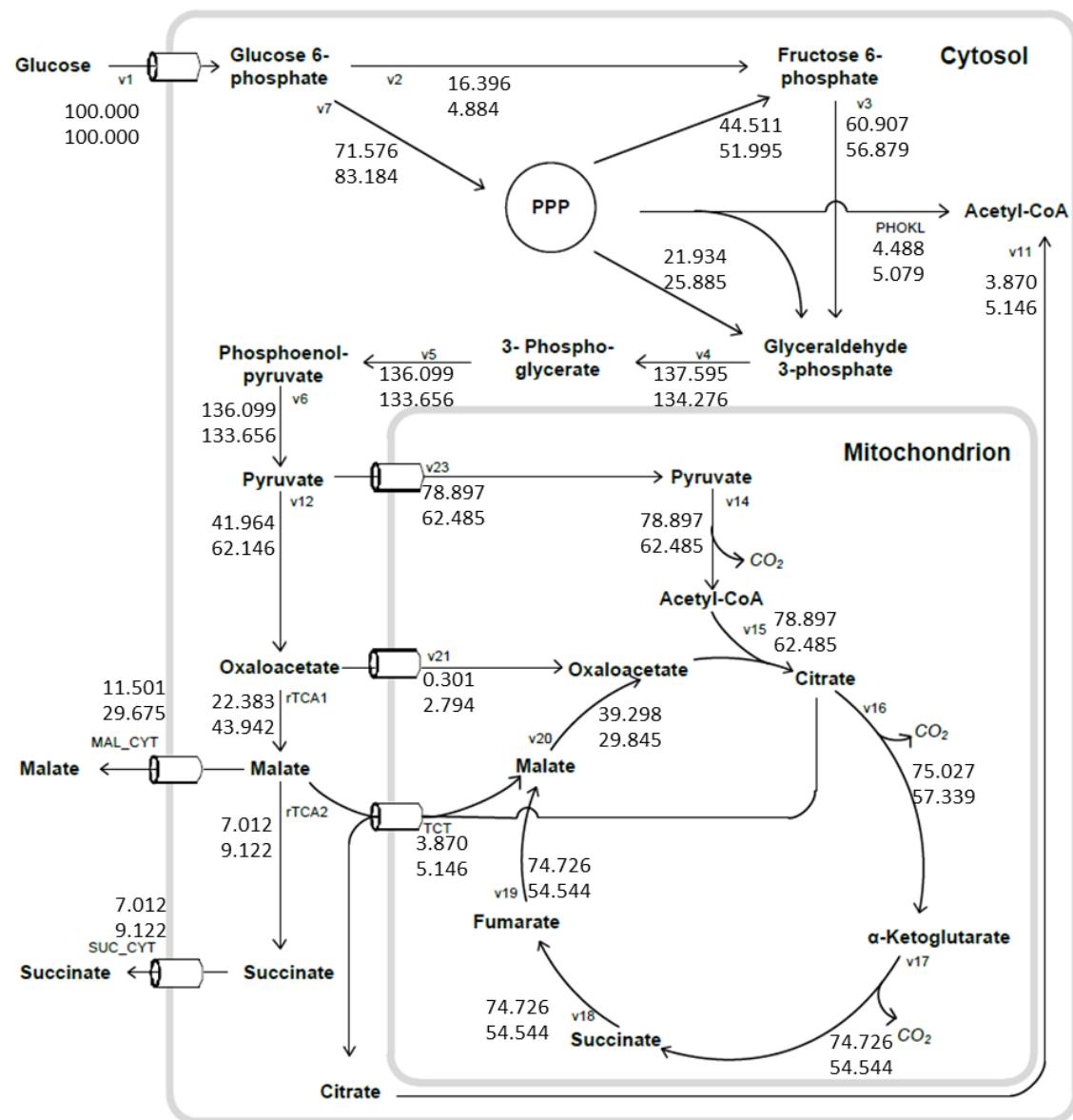


Figure 13: Central carbon metabolism of *A.oryzae* including the intracellular fluxes of NRRL3488 (upper values) and 2103a-68 (lower values). The samples were taken in mid exponential phase (7.5h) from shake-flask cultivations with MAF medium containing 25 g L<sup>-1</sup> glucose (slightly modified after Knuf et al. [66]).

## 4 *A. oryzae* GEM update

Genome scale metabolic models (GEMs) are used for a great variety of systems biology applications. They can be used for data generation as well as interpretation. For transcription data for example, the metabolites that participate in certain reactions are linked through a GEM to the genes that encode for the reaction. Thereby the transcriptional changes around a certain metabolite can be determined and metabolic hot spots can be identified in the network [99, 108]. GEMs are available for the most important Aspergilli, *A. niger* [7], *A. nidulans* [33] and *A. oryzae* [153].

After the publication of the genome sequence of *A. oryzae* [83] in 2005 work on a GEM could start. As the annotation of only about 50% of the identified genes was too poor, this was improved by sequencing an expressed sequence tag (EST) library. After that it was possible to assemble a GEM. This first genome scale *A. oryzae* model iWV1314 was published by Vongsangnak et al. in 2008. The GEM was validated by comparing the model predictions for maximum specific growth rate ( $\mu_{max}$ ;  $h^{-1}$ ) in batch cultivations and biomass yield ( $Y_{sx}$ ; g DW (mmol Substrate) $^{-1}$ ) during chemostat cultivations with experimental data. The model was able to accurately predict the growth rates for given carbon source uptake rates in batch cultivations using glucose, maltose, glycerol and xylose as carbon sources with an average accuracy of 98%. The biomass yields during chemostat cultivations could be calculated successfully as well. The model was used in a protein production project where it aided the investigation of amino acid requirements of an  $\alpha$ -amylase overexpression strain [152]. As the requirements for tyrosine, aspartate cysteine and threonine significantly increased, the pathways leading to increased amounts of these amino acids were identified as possible targets for improving  $\alpha$ -amylase production in an industrial setting. For this kind of simulation the model was well suited, as the reactions that are important for the amino acid and protein production are well annotated. On the other hand, information about transporters, especially the annotation of mitochondrial transport reactions is poor in almost any GEM. This is the same case in this model, as out of the 161 unique transport reactions only 53 are annotated, leaving the existence of 108 mitochondrial transport reactions open for speculation. Furthermore, most genes are annotated according to their degree of homology to *A. niger*, *A. fumigatus* or *S. cerevsiae* genes, which means, that the transport characteristics might vary as well.

In order to further elucidate the organic acid production potential and to find further engineering targets, an accurate model for this purpose is needed. The malic acid secretion is very much dependent on transport reactions, especially the exchange of metabolites between the mitochondrion and the cytosol. As mentioned above, one model that explains the high malic acid secretion obtained with *A. niger* is based on the assumption, that malic acid production in the cytosol is preceding citric acid production [69]. The malic acid titers in the cytosol act as a trigger for the tricarboxylate transporter [129].

During cultivations of the high malic acid producing strain 2103a-68 citric acid accumulation was detected along with an increased malic and succinic acid production. Assuming this mechanism to be active in *A. oryzae* as well, the knock-out of the TCT encoding gene in *A. oryzae* might diminish secretion of the by-product citric acid and thereby positively influence the yield of the desired C<sub>4</sub> dicarboxylic acids. The TCT reaction is one of the few mitochondrial membrane transport proteins that are annotated. According to the model the gene AO090020000012 encodes for the anti-port of malic acid and citric acid. In order to verify this strategy, knock out simulations were performed, in which the experimental external fluxes of malate, succinate and citrate were used as constraints. These calculations always resulted in a possible solution, indicating that the network around the organic acid transport between mitochondrion and cytosol offers too much flexibility. When looking at the mitochondrial membrane exchange possibilities in the model, it almost seemed as if there was no border for organic acids.

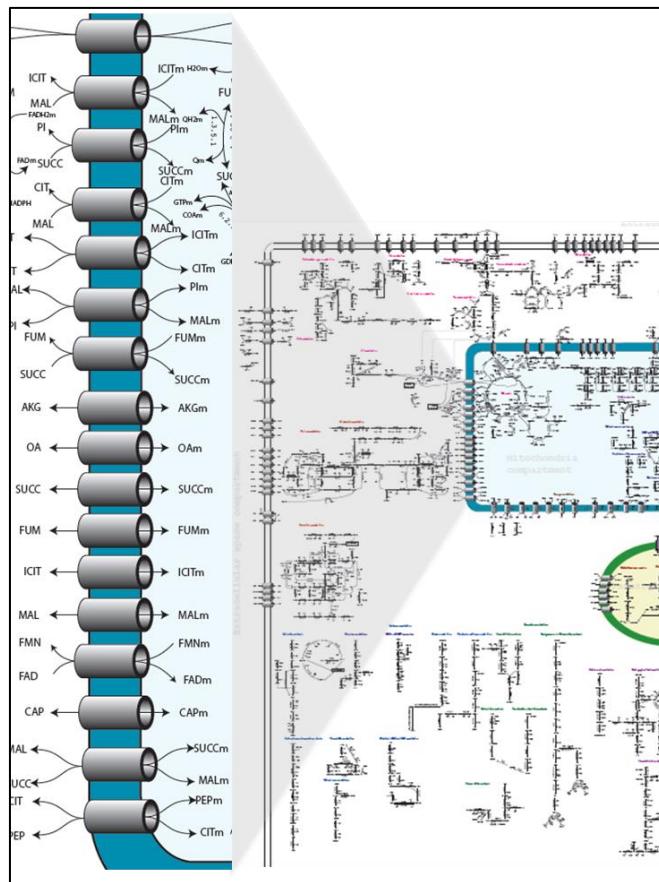


Figure 14: Detail of the graphical representation of the mitochondrial transport reactions of iWV1314.

As this would not be feasible in nature, the model was curated with special emphasis on the mitochondrial transporters. Through BLAST comparison of *A. oryzae* genes with other Aspergilli and *S. cerevisiae*, using KEGG [58], Uniprot [28], Cello [157] and NCBI BLAST [5], the latest annotations were taken into account for improved annotation of the *A. oryzae* model.

The resulting model *iLC1407* contains net 18 additional reactions compared to the available model from the Sysbio Toolbox [30], which is an update in which some bugs of *iWV1314* were fixed. Furthermore 38 gene annotations were added and 46 relocated. 28 of the transport reactions were removed, as no homolog genes were found in *A. oryzae* (Table 4).

Table 4: Model properties of *iWV1314*, the update available on Sysbio.se and the latest update *iLC1407*

Characteristics	<i>iWV1314</i>	Vongsangnak; Ågren Updates	<i>iLC1407</i>
Reactions	1846	2328	2346 (+18)
Metabolites	1073	1264	1265 (+1)
Genes	1314	1369	1407 (+38 46 relocated)
Transport reactions	281	457	429 (-28)
Compartments	4	4	4

Predictions of biomass yields on different carbon sources were made using *iLC1407* and very good fits were obtained using pure and mixed carbon sources (Figure 15). Then the metabolic functions of *A. oryzae* and *A. niger* were analyzed in order to investigate the differences of both species that diverge the metabolism for the production of malic acid and citric acid under similar culture conditions. Through the analysis of gene orthologs, it has been observed that both organisms possess almost the same metabolic machinery including proton transport and electron transport chain. Hence the *in-silico* production of TCA intermediates of both GEMs showed no differences. These results suggest that regulatory functions or varying enzyme transport capabilities are related to the production of malic acid in *A. oryzae* and citric acid in *A. niger*.

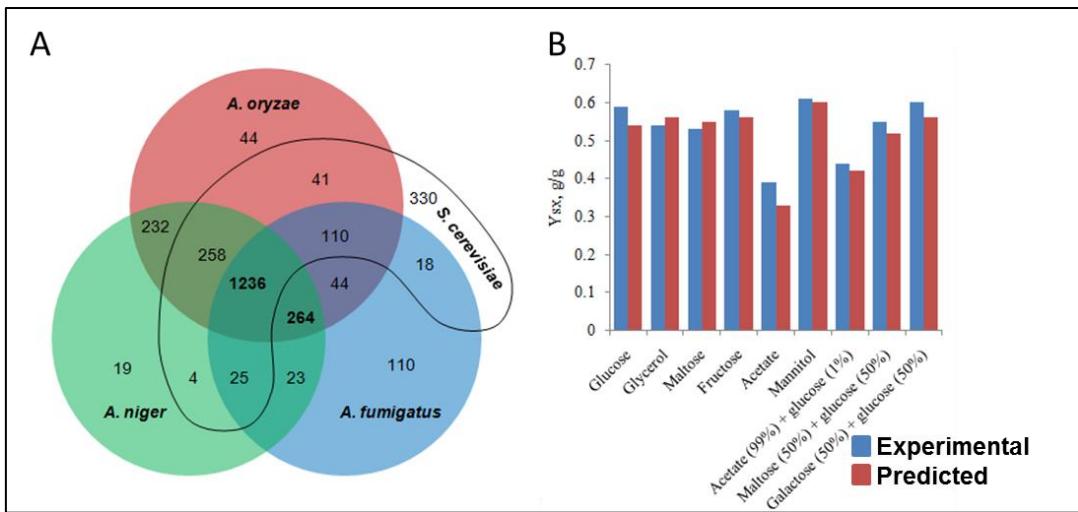


Figure 15: Venn diagram of homolog genes in *A. niger*, *A. oryzae*, *A. fumigatus* and *S. cerevisiae* (A) and comparison of predicted biomass yields ( $Y_{SX}$ ) using the updated model iLC1407 and experimental data on different carbon sources [22, 110, 123]

Furthermore different knock out scenarios were calculated using the random sampling algorithm [16] and experimental data from malic acid fermentations that were used as external fluxes. *In silico* the TCT (citrate/malate antiporter), the aspartate/malate shuttle, the oxaloacetate/alpha keto glutarate transporter and combinations of these were knocked out. Other than expected, all *in silico* mutants were able to grow and fulfill the constraints. The growth ability should have been hampered significantly by the TCT deletion, as the cytosolic acetyl-CoA supply was supposed to be hampered. Cytosolic acetyl-CoA is needed for fatty acid and sterol synthesis and protein acetylation [55]. There are two main pathways considered for the supply of cytosolic acetyl-CoA. The first route is mainly used by e.g. *S. cerevisiae* or *Candida albicans*, which converts acetate that originates from pyruvate via acetyl-CoA synthetase to acetyl-CoA. The second route utilizes ATP-citrate lyase which converts cytosolic citrate to acetyl-CoA and oxaloacetic acid. As shown for *A. nidulans*, the deletion of the ATP-citrate-lyase, which converts citrate to acetyl-CoA, is greatly diminishing growth on carbon sources that do not result in cytosolic acetyl-CoA [55], indicating that the acetyl-CoA synthetase is not able to supply acetyl-CoA in e.g. glucose containing medium.

As the main transporter for citrate out of the mitochondrion into the cytosol was removed, there should be no possibility to generate cytosolic acetyl CoA in our calculations. But instead the model used the pyruvate decarboxylase route, as the enzymes needed for that route are present in *A. oryzae*. This result shows the limitations of FBA and stoichiometric models, they do not consider regulation, but give you the best solution possible with the set of reactions/genes the model/organism contains. On the one hand this could be assessed as a flaw, as it is not accurately reflecting the natural behavior. On the other hand it opens the possibility to speculate about evolutionary engineering a strain that is not able to generate acetyl-CoA as the citrate route is blocked to relief repression on the acetyl-CoA synthetase

route, thereby generating a strain that is not producing citrate, but increased malate and can still grow on glucose as sole carbon source.

The above mentioned tricarboxylate transporter is not described for Aspergilli, but the *S. cerevisiae* proteins Yhm2p and Ctp1p were shown to have citrate transport abilities [37]. Yhm2p is supposed to be a component of the citrate-oxoglutarate NADPH redox shuttle without malate transport capabilities and to play a role in replication and segregation of the mitochondrial genome [24]. For Ctp1p citrate and malate transport abilities have been proven, whereas the  $K_m$  is significantly higher than reported for mammalian systems [59]. A BLAST search of Cpt1p against the *A. oryzae* RIB40 genome returned several positive genes. Out of these, two genes, AO090020000012 and AO090102000454 showed a significant up-regulation comparing starvation against the growth phase (Data obtained in the first study, Table 5). As AO090020000012 was among the genes most probably encoding for a transporter with citrate-malate antiport function, this gene was deleted in the *A. oryzae* strain NRRL3488. But in initial shakeflask cultivations, the deletion strain did not show a significant difference in the profile of secreted organic acids.

Table 5: Comparison of transcriptional changes of *A. oryzae* genes, which show a certain degree of homology to the *S. cerevisiae* Ctp1p, a tricarboxylate transporter gene. Transcriptional comparison of expression levels between the starvation and the growth phase.

Gene	Sequence identity to <i>S. cerevisiae</i> Ctp1p	adj,P,Val	logFC
AO090020000012	49.8%	0.000317967	0.501386012
AO090023000454	49.0%	1.27854E-05	0.454117181
AO090005000048	39.1%	0.159778469	-0.061435612
AO090102000125	32.4%	0.436944134	-0.076882676

Another question that arose from the project on the engineered strain was the question of the origin of the detected succinate. 2103a-68 was not only producing increased levels of malate, but also increased amounts of succinate. This led to the speculation, that there is a direct continuation of the reductive TCA branch from malate via fumarate to succinate in the cytosol. This question was addressed by measuring the summed fractional labelling of the secreted malic acid and succinic acid by GC-MS. Furthermore the summed fractional labelling for cytosolic and mitochondrial derived malate and succinate were simulated (Figure 16).

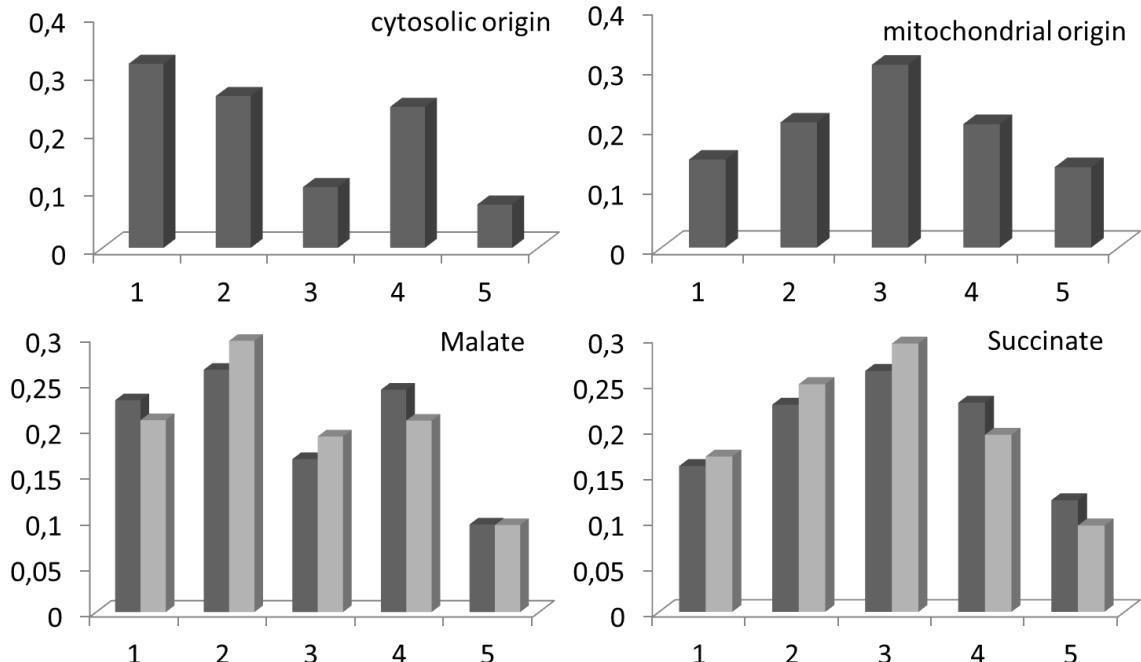


Figure 16: Pattern of the summed fractional labelling. Upper part, simulated pattern for cytosolic and mitochondrial derived malic acid. Lower part, fitting of simulated and experimentally derived SFLs for malate and succinate. Dark bars, fitted SFLs; bright bars, measured SFLs.

The pattern of summed fractional labelling of malate resembled the one of the cytosolic route to a large extend, with minor contributions from the TCA cycle. The labelling pattern of the succinic acid mainly resembled the simulated pattern of the TCA cycle derived succinic acid. Therefore, the idea of a functional pathway from malic acid to succinic acid was dropped and the model structure concerning the reductive cytosolic TCA branch was kept as in iWV1314.

A simpler flux model with the summed fractional labelling of malate and succinate and the external fluxes for malate, succinate, citrate and glucose as input was used to calculate internal fluxes Figure 17. These indicate that 22% of the secreted malate originated from the mitochondrion.

In summary, the model curation led to a “cleaner” picture of the current knowledge on mitochondrial exchange reactions, it is important to stress the current knowledge part, as the knowledge about mitochondrial transport reactions in general and in *A. oryzae* in particular are not that well studied. Based on the current model, it was shown that metabolic possibilities of *A. oryzae* and *A. niger*, extrapolated from the data obtained from the corresponding GEMs, are not that different, which leads to the conclusion, that regulatory mechanisms or enzyme properties make the difference between malic acid and citric acid secretion.

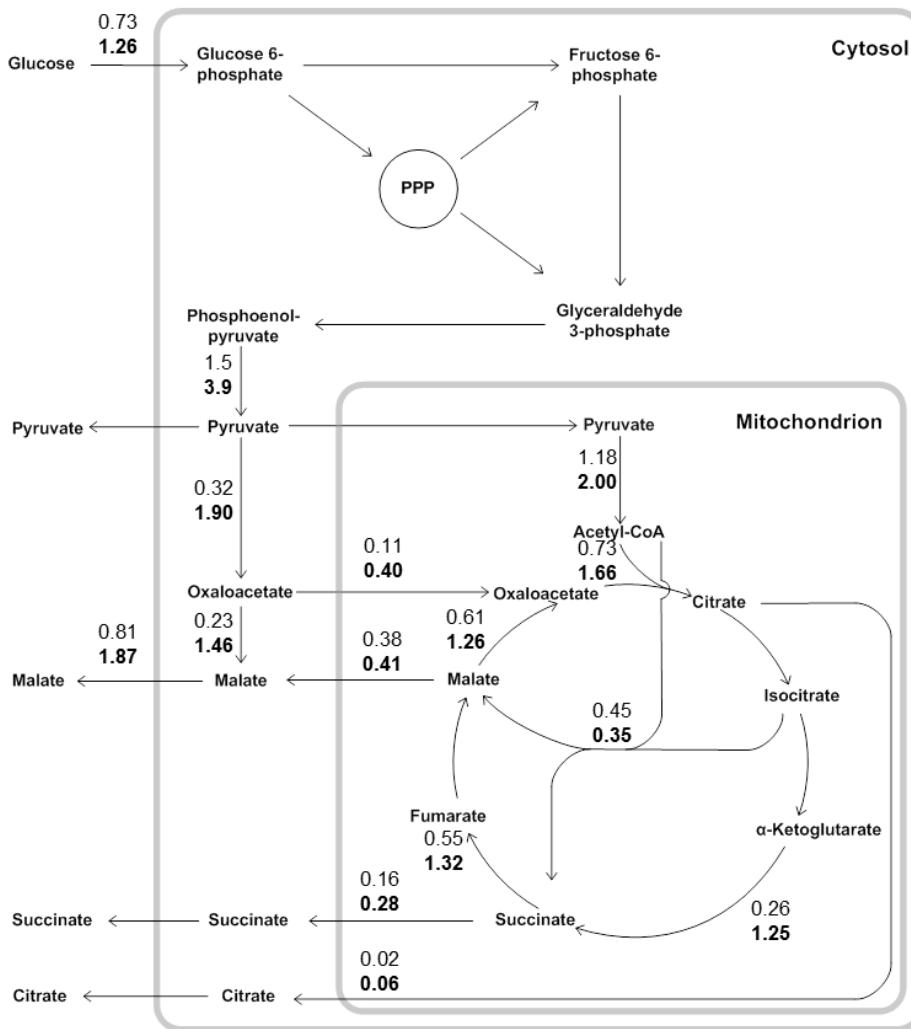


Figure 17: Central carbon metabolism of *A. oryzae* and internal fluxes calculated according to external fluxes determined for NRRL3488 and 2103a-68 during the stationary phase of a bioreactor cultivation and the summed fractional labelling of malate and succinate from the final sample of a shakeflask cultivation.

This allows for speculations about turning *A. niger*, which can grow at even lower pHs than *A. oryzae*, into an efficient malic acid producer. The deletion of the gene AO090020000012, which was identified to possibly encode for a TCT in *A. oryzae*, did not lead to significant changes in the organic acid production and will require future investigation. The additional deletion of AO090102000454, which might be another gene encoding for a TCT, could be an additional task for the future.

## 5 Manipulating the carbon source utilization

Most microorganisms prefer to metabolize glucose over any other carbon source [128] and so does *A. oryzae*, as shown in the section 3. Enzymes which are needed for the catabolism of less preferred carbon sources are usually transcriptionally inhibited. This inhibition also contains enzymes participating in the degradation of sugar polymers, like cellulose, hemicellulose and pectin. In Aspergilli, the transcriptional inhibitor CreA is known to be a major player in the complex regulation mechanisms [128]. CreA has been studied intensively in *A. nidulans*. Like *S. cerevisiae*'s Mig1, it contains zinc fingers of the Cys2His2 type, which probably bind to the 5'-SYGGRG-3' consensus sequence. When deleting CreA in *A. nidulans*, Prathumpai et al. [119] reported previously that the subsequently metabolized sugars glucose and xylose were consumed in parallel. Furthermore, culture supernatant of the mutant strain showed elevated xylanase activity, which would be beneficial for a simultaneous saccharification and fermentation process. This would also cut down costs, as the expenses for enzyme mixes for hydrolysis would be omitted.

As shown above, *A. oryzae* is able to consume xylose at high rates and efficiently converts this pentose to malate. Lignocellulosic feedstocks, which would be the preferred carbon source for a truly sustainable production of chemicals through a biorefinery, contain both glucose and xylose. In a batch-cultivation setup the subsequent utilization does not cause problems, but in a continuous cultivation setup, only the glucose fraction would be used for the conversion and the carbon from xylose will not be metabolized. This is of course a major drawback and a parallel utilization, as shown for the *A. niger* CreA mutant, would be desirable. Therefore the CreA gene was deleted in a uracil auxotrophic descendent of NRRL3488.

This pyrG deleted NRRL3488 mutant was transformed with the deletion cassette. This cassette contained the *A. niger* pyrG gene flanked by 1kb DNA fragments which are homologous to the upstream and downstream region of the creA gene. Prototrophic mutants were selected on minimal medium and colonies were subsequently purified on minimal medium. DNA from purified colonies was isolated and PCR was performed using the DNA as template. Primers were constructed to bind upstream and downstream of the integration cassette (Figure 18) and inside the pyrG sequence, running towards the ends. The two primer pairs for upstream and downstream verification amplified a 1.5 kb and 2 kb fragment, respectively. All but one of the in Figure 18 displayed mutants showed the expected bands. As Aspergilli are known to integrate DNA fragments in an ectopic manner, the correct single integration of the deletion fragment was furthermore verified by Southern blotting. As only mutant AOMCK01.09 showed a clear single band, this transformant was used for further analysis of the effect of the *creA* deletion in *A. oryzae*.

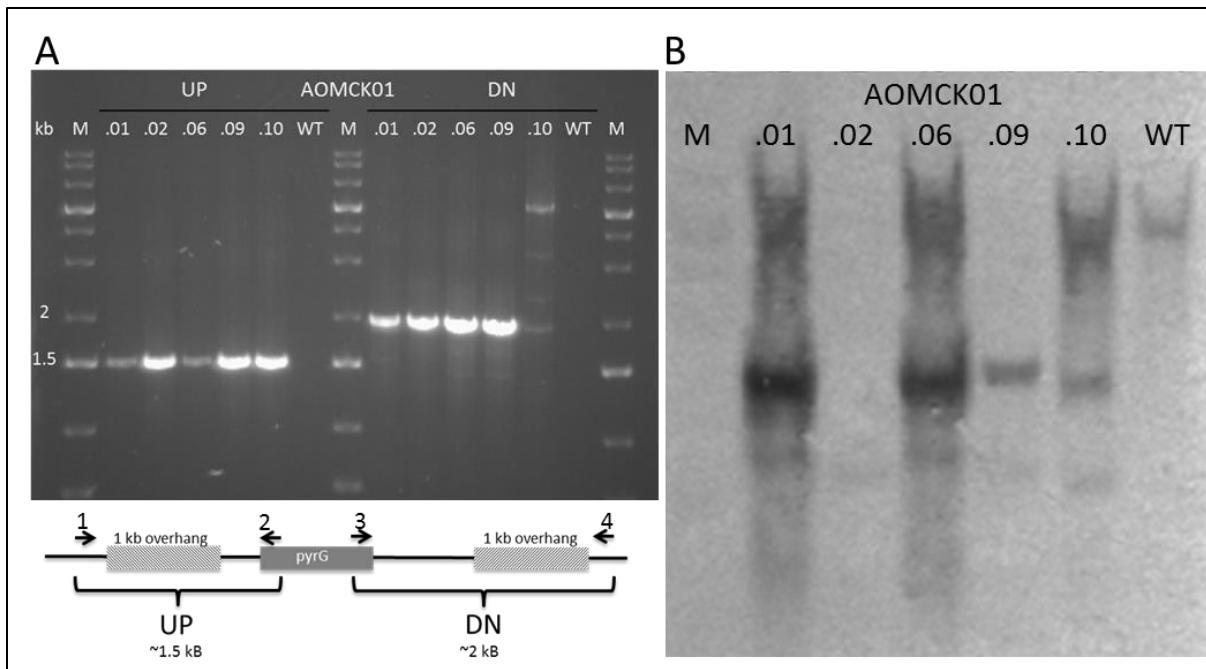


Figure 18: PCR (A) and southern blotting (B) confirmation of the integration of the deletion fragment. M: GeneRuler™ 1 kb Plus Ladder; WT: NRRL3488

The resulting *creA* deletion strain showed reduced mycelial growth compared to the wild-type. The mycelium is very dense in the center of the colony and an uncoordinated network of hyphae is growing outwards, whereas the wild-type spreads evenly in straight radial hyphae Figure 19. Furthermore the mycelium is growing upwards as well in case of NRRL3488, whereas the mutant hyphae are flat on top of the agar. Concerning sporulation, AOMCK01.9 takes much longer time to sporulate and the spores are concentrated more towards the center, whereas in the case of the wild-type, sporulation occurs more evenly after around 4 days.

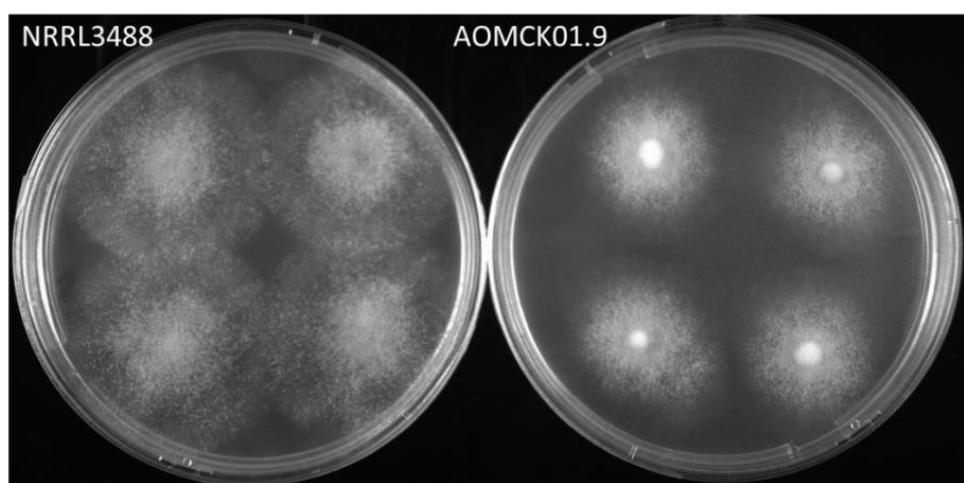


Figure 19: Growth comparison of NRRL3488 and AOMCK01.9 on spore propagation plates.

The first thing to check on the mutant strain was the effect of the *creA* deletion on the carbon source uptake. Therefore the wild-type and mutant strain were cultivated in MAF

medium containing increased nitrogen source, in order to exclude the effect of nitrogen starvation. Both strains consumed the glucose first (Figure 20), until a glucose concentration of about 2 g L<sup>-1</sup> was reached. This happened at about 10h of cultivation, thereafter glucose and xylose were consumed in parallel and after 13h, when the glucose was completely exhausted, xylose consumption went on until about 25h of cultivation. This result is in disagreement with the results obtained in *A. nidulans*.

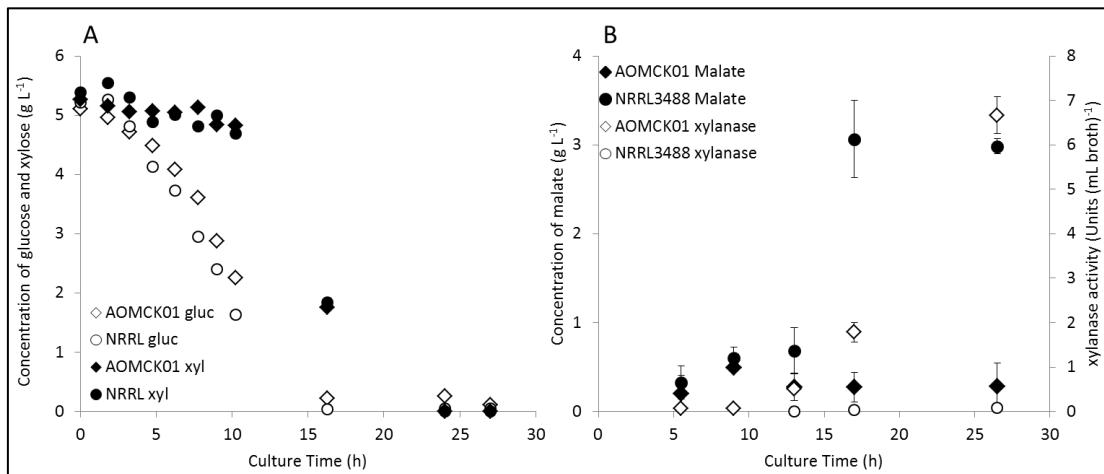


Figure 20: Fermentation profiles of NRRL3488 and AOMCK01.09 during batch cultivations in MAF medium containing 6.4 g L<sup>-1</sup> ammonium sulfate and 5g of each carbon source. Carbon source concentration (A), malate concentration and xylanase activity (B).

Another trait previously found for an *A. nidulans creA* deletion mutant was the positive influence on the secretion of xylanases. In order to check for the same effect, the same fermentation conditions as above were applied and the xylanase activity of the supernatant was measured. For NRRL3488, no activity could be detected in the first samples, in the sample at 17h, the activity reached 0.045 Units mL<sup>-1</sup>. In the following samples the activity kept rising to finally reach 0.090 Units mL<sup>-1</sup>. For the mutant strain on the other hand, xylanase activity could already be detected after 13 h. From there on the activity increased constantly from 0.510 Units mL<sup>-1</sup> to finally reach 6.668 Units mL<sup>-1</sup>. This final activity is 74 times higher than measured for the wild-type cultivations. Though the deletion obviously affected the secretion of xylanases positively, the effect of glucose repression either on the xylanase secretion or on the xylose metabolism could not be confirmed.

The initial aim was to construct a strain that could be used in a continuous consolidated bioprocess. As the strain NRRL3488 was previously shown to be able to produce malic acid, the *creA* deleted strain was supposed to be used in a process to produce this C<sub>4</sub> dicarboxylic acid. Therefore the ability of AOMCK01.9 to produce malic acid was evaluated as well. As can be seen in Figure 20, there was hardly any malate detected in the supernatant of the mutant cultivation, whereas about 3 g L<sup>-1</sup> were accumulated in the NRRL3488 cultivation. The initial characterization was performed using high amounts of nitrogen in the fermentation broth. In order to simulate the production medium and investigate the malic acid production potential further, the wild-type and the mutant strain were cultivated in MAF medium

containing the usual  $1.4 \text{ g L}^{-1}$  of ammonium sulfate. The carbon source composition was also varied. In the first case a glucose/xylose mixture, in the second pure glucose and in the third case pure xylose were used at initial concentrations of  $30 \text{ g L}^{-1}$ , each. The final concentrations of malate are shown in Figure 21. Though the final concentration of malate increased significantly and also the ratio towards the wild-type is not as disproportionate as in the initial comparison, the final malic acid concentration in the mutant broth is less than half of the wild-type.

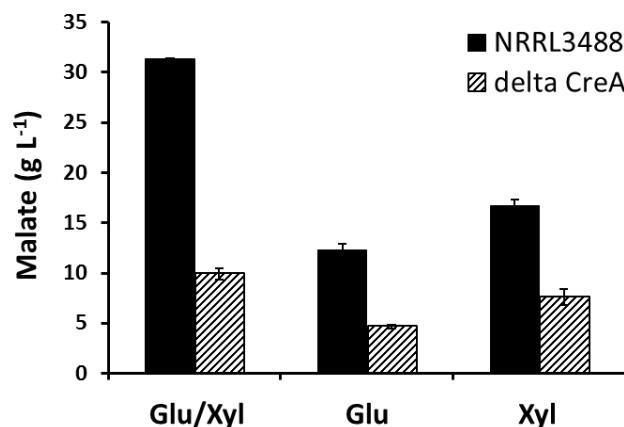


Figure 21: Final malate concentrations of NRRL3488 and AOMCK01.9 after cultivation in three different carbon source compositions. Glu/Xyl, glucose and xylose mixture; Glu, glucose only; Xyl, xylose only; intial concentration of wach carbon source was  $30 \text{ g L}^{-1}$ .

Taken together, the deletion of *creA* did not lead to the expected effect of creating a strain that can produce malic acid from parallel metabolized glucose and xylose. The strain AOMCK01.9 was producing less than half of the final titer of malate than the wild-type NRRL3488 and consumed glucose before xylose. But the strain showed significantly elevated xylanase activity in the fermentation broth. As the reduced malate production indicates an important role of the deleted gene AO090026000464 in the regulation of the central carbon metabolism, further investigation through for example transcriptome analysis of this strain, would be an interesting follow up project.

## 6 Conclusion

The objective of the work was to investigate the malic acid production potential of the filamentous fungus *A. oryzae*. Therefore the ability of the wild-type strain NRRL3488 concerning malic acid production was evaluated on different nitrogen sources and the transcriptional changes between the exponential growth phase and the malic acid production phase, which is characterized by nitrogen starvation, was evaluated (**Paper I**). Subsequently an engineered *A. oryzae* strain (2103a-68) was characterized in lab scale fermenters (**Paper II**). This strain also produced significant amounts of citrate as a by-product. In order to perform *in-silico* knock out evaluations the existing model needed to be updated and was refined concerning reactions connected to organic acid production. During the characterization of 2103a-68, the performance of this strain was not only evaluated on the commonly used carbon source glucose, but also on xylose. As the positive results for both carbon sources paved the way towards the use of *A. oryzae* as a platform organism for the production of renewable chemicals in the biorefinery of the future, the carbon source utilization was sought to be de-regulated by the deletion of AO090026000464, which is supposed to encode for the carbon repressor *creA* (**Paper III**).

From the first study we learned that *A. oryzae* wild-type strains have a natural ability to produce significant amounts of malic acid and that they convert glucose to malic acid with half of the maximum theoretical yield, when nitrogen becomes limiting. This is an important trait in order to reach high yields, as the carbon is not incorporated into biomass, but into the product, in this case malic acid. The regulation of this mechanism was investigated through transcriptome analysis, which revealed a general up-regulation of genes involved in the glycolysis, a synchronistic down-regulation of TCA genes and an up-regulation of the rTCA genes in the starvation phase compared to the growth phase. Binding sites for the *S. cerevisiae* transcriptional activator Msn2/4 were found when analyzing the up-stream sequences of the up-regulated genes of the glycolysis and rTCA. This suggests that *A. oryzae* uses malic acid production as a stress response towards nitrogen starvation. The metabolic changes evoked by the nitrogen starvation stress, leading to continued conversion of glucose to malate also opens possibilities to utilize this mechanism to produce other products. One could for example speculate about developing metabolic engineering strategies that divert the flux away from malic acid at the various branch-points on the way towards malic acid. The most important hub is the cytosolic pyruvate pool. A product easily derived from pyruvate with only one enzymatic reaction is lactate, for which the flux would just have to be diverted through the lactate dehydrogenase reaction. Another product with increased interest is 3 hydroxypropionic acid. There are several theoretical pathways from pyruvate to 3HP [147], for most of them the enzymes that would be needed were not reported to occur in nature, yet. One of the working pathways has been reconstructed in *E. coli* and goes in three steps from Pyruvate via Acetyl-CoA and Malonyl-CoA to 3HP [124]. A very interesting but so far only theoretical pathway employs a hitherto unknown malate decarboxylase. Once such an enzyme is found or engineered, it would allow production of 3HP in a single

step from malic acid and in this case the *A. oryzae* NRRL3488 strain would be the perfect production platform.

An even better strain to build on would be the engineered *A. oryzae* strain 2103a-68, which was further characterized in paper 2 and shown to be an even better malic acid producer than the wild-type. The overexpression of the cytosolic TCA branch and the malic acid transporter let to production rates, yields and final titers that are among the highest ever reported for a microbial system. Nevertheless, the pyruvate carboxylase reaction might still be the controlling step in this pathway. In order to overcome this problem, inserting more copies of the *pyc* expression cassette and resulting higher expression, or enzyme engineering might be helpful in this respect. Another observation was the increased citric acid production, which might be a result of the increased activity of the tri-carboxylate transporter, which is supposed to be triggered by high cytosolic malic acid concentrations. A deletion of the corresponding gene in *A. oryzae* could lead to diminishing the citric acid production and to higher yields of malate. During this second project, the performance of the engineered strain was furthermore evaluated on xylose and glucose/xylose mixtures and the positive results confirmed the suitability of this strain to be used in a biorefinery, using e.g. lingo-cellulosic biomass, which mainly consists of those two sugars, as carbon source. Pretreated renewable feedstocks usually contain additional inhibitors and Aspergilli are generally known to be more resistant than other organisms, nevertheless follow-up studies would have to confirm the positive results concerning malic acid production on pretreated biomass.

The issue of alternative carbon source utilization was also addressed in this thesis. As seen from the cultivations with the glucose/xylose mixture, the engineered strain prefers to consume glucose over xylose. For *A. niger* it was shown that the deletion of *creA*, a carbon catabolite repressor, let to simultaneous consumption of both sugars and increased secretion of hydrolytic enzymes. This effect could not be fully confirmed for the deletion of the gene AO090026000464, which is supposed to encode for CreA. The knock-out led to increased xylanase activity, but the carbon sources were still consumed subsequently and the malic acid production was also negatively affected. These results indicate that the regulation of the central carbon metabolism is affected in a way, as well as the carbon repression of secretion of hydrolytic enzymes. In order to further investigate the role of AO090026000464, transcription and metabolite analysis would be interesting.

The initial *A. oryzae* model was constructed with respect to protein production. In order to be able to confidently predict the metabolism of organic acid production, the transport reactions between the mitochondrion and the cytosol were of special interest. As these and other reactions connected to organic acid production were poorly annotated, the annotation and localization of reactions in the *A. oryzae* GEM were revised. This curation led to the deletion of 28 transport reactions for which no annotation could be found, annotation of 38 reactions and re-localization of 46 reactions. The resulting model was able to accurately

predict previously reported results and could be used as a scaffold for further development of metabolic engineering targets for not only protein production, but also organic acid production.

Nevertheless, the GEMs of the Aspergilli and the one of *A. oryzae* in particular are not as concise as the one of the model organisms like *S. cerevisiae* or *E. coli*. This is mainly due to bad annotation or annotation entirely based on homology to other organisms. Bad gene ontology definitions are a result of the incomplete or inaccurate gene annotation and this hampers high throughput analysis of transcription data. In order to bring systems biology in Aspergilli to a level seen in e.g. *S. cerevisiae*, the community has to make an effort in order to reach more reliable gene annotation and localization in the future.

Taken together, the presented work shows the great opportunities that *A. oryzae* offers for biotechnological applications. The work on malic acid production in this organism extended the possible use not only for the production of enzymes, but also bulk chemicals like organic acids. *A. oryzae* combines several advantages. It has been used safely for several centuries in the food industry. Furthermore large scale production processes have been established for the production of enzymes and the experience can be applied for the production of organic acids as well. In addition it is able to secrete large amounts of hydrolytic enzymes, which can help making the carbon from renewable feedstock accessible for conversion to the desired product. All these advantages make *A. oryzae* the organism of choice for the future biorefinery, which will aid in making the world a more sustainable place. People might think that *A. oryzae* is not as “sexy” as *E. coli* or *S. cerevisiae* and projects might take longer time, but *A. oryzae* has a great potential for the sustainable production of chemicals. After the era of “proof of principle” metabolic engineering in model organisms I hope that *A. oryzae* will be attended to as production host in order to achieve the required yields, titers and rates needed for the economic feasibility of industrial production processes.

## Acknowledgements

In order to get to this point ones needs a lot of support. Concerning the professional side my biggest thanks go to Jens Nielsen for guiding me through the “odd project” and giving me the freedom to develop my own ideas and carry them out in the lab. I also have to thank him for gathering this diversity of great people from all over the world that form the Systems and Synthetic Biology Group here at Chalmers.

I would like to thank my closest collaborators, Intawat Nookaew for his bioinformatics support and his troubleshooting abilities, Sergio Bordel Velasco for his help with the flux calculations, Luis Caspeta for the model curation, Sakda Khoomrung for his support with the analytics and Stefan Tippmann, Marieke Buffing and Ilse Remmers for their help in the lab.

Many thanks go to our collaborators at Novozymes Inc., Alan Berry and Steven Brown for sharing their results on malic acid production by *Aspergillus oryzae*, the provision of the engineered strain and their support during the gene deletions. Thanks to Michael McCulloch for introducing me to cultivations of filamentous fungi.

I already started as a master student in the former isoprenoid group and many thanks go to Verena Siewers, who was my first contact here and who taught me metabolic engineering and strain construction of *S. cerevisiae*. Gionata Scalcinati introduced me to the DasGip fermenter system and Siavash Partow and Yun Chen were always willing to answer questions concerning practicalities in the lab.

Many thanks to the technical staff who keeps the lab running, Marie Nordqvist, Suwanee Jansa-Ard, Malin Nordvall, Emma Ribbenhed and Ximena Rozo Sevilla and the office staff who were so helpful in organizing life in Sweden, Erica Dahlin and Helena Janveden and Martina Butorac for help with the graphical design of posters and figures.

And of course many thanks to the members of the Sys2Bio and IndBio group who created such a great atmosphere in the lab and the kitchen that made me enjoy most days at work. Special thanks to Maurizio Bettiga and Dina Petranovic for discussing academic life. Special thanks to my friends and office mates, Tobias Österlund for spending so many days with me outdoors in the mountains and on the sea, Bouke de Jong for the hours in the gym and sauna and many more social activities, Rahul Kumar for all the kilometers we ran together, Francesco Gatto for his positive attitude, Sakda Khoomrung for the whisky evenings, Nina Johansson, Rakesh Kopram, Lina Lindberg, Hampus Sunner, Rasmus Ågren, Stefan Tippmann, Kaisa Thorell, Petri-Jaan Lahtvee, Alexandra Bergman, Adil Mardinoglu, Jichen Bao, Eduard Kerkhoven, Florian David, José L. Martínez, Clara Navarrete, Klaas Buijs, Saeed Shoae, Fredrik Karlsson and Leif Väremo for being around and making the time in Sweden so enjoyable.

Additional thanks go to the Thai community, Kanokarn Kocharin, Natapol Pornputtapong, Kwanjeera Wanichthanarak, Sakda Khoomrung, Suwanee Jansa-Ard, Pramote

Chumnanpuen, Wanwipa Vongsangnak and Intawat Nookaew. Though my only contribution to the cooking was to peel onions and garlic, they and especially Intawat kept providing me with Thai food and we spend many enjoyable evenings together. This was also the time when Wanwipa lured me into this *Aspergillus* project, thanks for that.

And last but not least my greatest thanks go to my Mom, Dad and Sister for their endless support and trust over the last decades.

## References

1. Abe S, Furuya, Akira, Takayama, Ken-Ichiro (1962) Method of producing L-malic acid by fermentation. In: KYOWA HAKKO KOGYO KK, United States patent 3,063,910
2. Alekseenko A, Clutterbuck AJ (1996) The plasmid replicator AMA1 in *Aspergillus nidulans* is an inverted duplication of a low-copy-number dispersed genomic repeat. Mol. Microbiol. 19:565-574
3. Alekseenko A, Nikolaev I, Vinetski Y et al. (1996) Gene expression from replicating plasmids in *Aspergillus nidulans*. Mol Gen Genet 253:242-246
4. Altmeyer PJ, Mattlies U, Pawlak F et al. (1994) Antipsoriatic effect of fumaric acid derivatives: Results of a multicenter double-blind study in 100 patients. Journal of the American Academy of Dermatology 30:977-981
5. Altschul SF, Gish W, Miller W et al. (1990) Basic local alignment search tool. J. Mol. Biol. 215:403-410
6. Andersen M, Lehmann L, Nielsen J (2009) Systemic analysis of the response of *Aspergillus niger* to ambient pH. Genome Biology 10:R47
7. Andersen MR, Nielsen ML, Nielsen J (2008) Metabolic model integration of the genome, metabolome and reactome of *Aspergillus niger*. Mol Syst Biol 4
8. Bailey J (1991) Toward a Science of Metabolic Engineering. Science 252:1668-1675
9. Barbesgaard P, Heldt-Hansen H, Diderichsen B (1992) On the safety of *Aspergillus oryzae*: a review. Appl. Microbiol. Biotechnol. 36:569-572
10. Battat E, Peleg Y, Bercovitz A et al. (1991) Optimization of L-malic acid production by *Aspergillus flavus* in a stirred fermentor. Biotechnol. Bioeng. 37:1108-1116
11. Bechthold I, Bretz K, Kabasci S et al. (2008) Succinic Acid: A New Platform Chemical for Biobased Polymers from Renewable Resources. Chemical Engineering & Technology 31:647-654
12. Becker J, Reinefeld J, Stellmacher R et al. (2013) Systems-wide analysis and engineering of metabolic pathway fluxes in bio-succinate producing *Basfia succiniciproducens*. Biotechnol. Bioeng. 110:3013-3023
13. Bercovitz A, Peleg Y, Battat E et al. (1990) Localization of pyruvate carboxylase in organic acid-producing *Aspergillus* strains. Appl. Environ. Microbiol. 56:1594-1597
14. Bizukojc M, Ledakowicz S (2009) Physiological, morphological and kinetic aspects of lovastatin biosynthesis by *Aspergillus terreus*. Biotechnology Journal 4:647-664
15. Boeke J, Croute F, Fink G (1984) A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol Gen Genet 197:345-346
16. Bordel S, Agren R, Nielsen J (2010) Sampling the Solution Space in Genome-Scale Metabolic Networks Reveals Transcriptional Regulation in Key Enzymes. PLoS Comput Biol 6:e1000859

17. Brown SH, Bashkirova L, Berka R et al. (2013) Metabolic engineering of *Aspergillus oryzae* NRRL 3488 for increased production of L-malic acid. *Appl Microbiol Biotechnol* 97:8903-8912
18. Buschke N, Schröder H, Wittmann C (2011) Metabolic engineering of *Corynebacterium glutamicum* for production of 1,5-diaminopentane from hemicellulose. *Biotechnology Journal* 6:306-317
19. Buxton FP, Gwynne DI, Davies RW (1989) Cloning of a new bidirectionally selectable marker for *Aspergillus* strains. *Gene* 84:329-334
20. Campbell E, Unkles S, Macro J et al. (1989) Improved transformation efficiency of *Aspergillus niger* using the homologous *niaD* gene for nitrate reductase. *Curr. Genet.* 16:53-56
21. Cao N, Du J, Gong CS et al. (1996) Simultaneous Production and Recovery of Fumaric Acid from Immobilized *Rhizopus oryzae* with a Rotary Biofilm Contactor and an Adsorption Column. *Appl. Environ. Microbiol.* 62:2926-2931
22. Carlsen M, Nielsen J (2001) Influence of carbon source on  $\alpha$ -amylase production by *Aspergillus oryzae*. *Appl. Microbiol. Biotechnol.* 57:346-349
23. Carlsen M, Spohr AB, Nielsen J et al. (1996) Morphology and physiology of an  $\alpha$ -amylase producing strain of *Aspergillus oryzae* during batch cultivations. *Biotechnol. Bioeng.* 49:266-276
24. Castegna A, Scarcia P, Agrimi G et al. (2010) Identification and Functional Characterization of a Novel Mitochondrial Carrier for Citrate and Oxoglutarate in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 285:17359-17370
25. Chakraborty BN, Kapoor M (1990) Transformation of filamentous fungi by electroporation. *Nucleic Acids Res.* 18:6737
26. Chen X, Xu G, Xu N et al. (2013) Metabolic engineering of *Torulopsis glabrata* for malate production. *Metab. Eng.* 19:10-16
27. Cheng K-K, Zhao X-B, Zeng J et al. (2012) Downstream processing of biotechnological produced succinic acid. *Appl. Microbiol. Biotechnol.* 95:841-850
28. Consortium TU (2013) Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Res.* 41:D43-D47
29. Corrick CM, Twomey AP, Hynes MJ (1987) The nucleotide sequence of the *amdS* gene of *Aspergillus nidulans* and the molecular characterization of 5' mutations. *Gene* 53:63-71
30. Cvijovic M, Olivares-Hernández R, Agren R et al. (2010) BioMet Toolbox: genome-wide analysis of metabolism. *Nucleic Acids Res.* 38:W144-W149
31. Dakin HD (1924) THE FORMATION OF I-MALIC ACID AS A PRODUCT OF ALCOHOLIC FERMENTATION BY YEAST. *J. Biol. Chem.* 61:139-145
32. Datta R, Glassner D, Jain M et al. (1992) Fermentation and purification process for succinic acid US Patent 5:168,055. In:
33. David H, Hofmann G, Oliveira A et al. (2006) Metabolic network driven analysis of genome-wide transcription data from *Aspergillus nidulans*. *Genome Biology* 7:R108

34. De Jongh WA, Nielsen J (2008) Enhanced citrate production through gene insertion in *Aspergillus niger*. *Metab. Eng.* 10:87-96
35. Edwards JS, Palsson BO (2000) The *Escherichia coli* MG1655 in silico metabolic genotype: Its definition, characteristics, and capabilities. *Proceedings of the National Academy of Sciences* 97:5528-5533
36. El-Enshasy H, Hellmuth K, Rinas U (2001) GpdA-promoter-controlled production of glucose oxidase by recombinant *aspergillus niger* using nonglucose carbon sources. *Appl. Biochem. Biotechnol.* 90:57-66
37. El Moualij B, Duyckaerts C, Lamotte-Brasseur J et al. (1997) Phylogenetic Classification of the Mitochondrial Carrier Family of *Saccharomyces cerevisiae*. *Yeast* 13:573-581
38. Evans CT, Scragg AH, Ratledge C (1983) A Comparative Study of Citrate Efflux from Mitochondria of Oleaginous and Non-oleaginous Yeasts. *Eur. J. Biochem.* 130:195-204
39. Fleißner A, Dersch P (2010) Expression and export: recombinant protein production systems for *Aspergillus*. *Appl. Microbiol. Biotechnol.* 87:1255-1270
40. Forment J, Ramón D, Maccabe A (2006) Consecutive gene deletions in *Aspergillus nidulans*: application of the Cre/loxP system. *Curr. Genet.* 50:217-224
41. Fungaro MHP, Rech E, Muhlen GS et al. (1995) Transformation of *Aspergillus nidulans* by microprojectile bombardment on intact conidia. *FEMS Microbiol. Lett.* 125:293-297
42. Förster J, Famili I, Fu P et al. (2003) Genome-Scale Reconstruction of the *Saccharomyces cerevisiae* Metabolic Network. *Genome Res.* 13:244-253
43. Galagan JE, Calvo SE, Cuomo C et al. (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* 438:1105-1115
44. García J, Torres N (2011) Mathematical modelling and assessment of the pH homeostasis mechanisms in *Aspergillus niger* while in citric acid producing conditions. *J. Theor. Biol.* 282:23-35
45. Gems D, Johnstone IL, Clutterbuck AJ (1991) An autonomously replicating plasmid transforms *Aspergillus nidulans* at high frequency. *Gene* 98:61-67
46. Goldberg I, Rokem JS, Pines O (2006) Organic acids: old metabolites, new themes. *Journal of Chemical Technology & Biotechnology* 81:1601-1611
47. Gong CS, Cao N, Sun Y et al. (1996) Production of l-malic acid from fumaric acid by resting cells of *Brevibacterium* sp. *Appl. Biochem. Biotechnol.* 57-58:481-487
48. Grotjohann N, Huang Y, Kowallik W (2001) Tricarboxylic Acid Cycle Enzymes of the Ectomycorrhizal Basidiomycete, *Suillus bovinus*. *Z. Naturforsch.* 56 c:334-342
49. Guettler MV, Rumler D, Jain MK (1999) *Actinobacillus succinogenes* sp. nov., a novel succinic-acid-producing strain from the bovine rumen. *Int. J. Syst. Bacteriol.* 49:207-216
50. Hahm YT, Batt CA (1988) Genetic Transformation of an argB Mutant of *Aspergillus oryzae*. *Appl. Environ. Microbiol.* 54:1610-1611
51. Hansen BG, Salomonsen B, Nielsen MT et al. (2011) Versatile Enzyme Expression and Characterization System for *Aspergillus nidulans*, with the *Penicillium brevicompactum*

Polyketide Synthase Gene from the Mycophenolic Acid Gene Cluster as a Test Case. Appl. Environ. Microbiol. 77:3044-3051

52. Hartingsveldt W, Mattern I, Zeijl CJ et al. (1987) Development of a homologous transformation system for *Aspergillus niger* based on the *pyrG* gene. Mol Gen Genet 206:71-75
53. Hedayati MT, Pasqualotto AC, Warn PA et al. (2007) *Aspergillus flavus*: human pathogen, allergen and mycotoxin producer. Microbiology 153:1677-1692
54. Hermann BG, Patel M (2007) Today's and tomorrow's bio-based bulk chemicals from white biotechnology. Appl. Biochem. Biotechnol. 136:361-388
55. Hynes MJ, Murray SL (2010) ATP-Citrate Lyase Is Required for Production of Cytosolic Acetyl Coenzyme A and Development in *Aspergillus nidulans*. Eukaryot. Cell 9:1039-1048
56. Ishida H, Hata Y, Kawato A et al. (2004) Isolation of a Novel Promoter for Efficient Protein Production in *<I>Aspergillus oryzae</I>*. Biosci., Biotechnol., Biochem. 68:1849-1857
57. Jump J, Brown SH (2011) Process for separating and recovering C4 dicarboxylic acids. WO2011002895 A3. In: Novozymes North America I, Novozymes I (eds)
58. Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 28:27-30
59. Kaplan RS, Mayor JA, Gremse DA et al. (1995) High Level Expression and Characterization of the Mitochondrial Citrate Transport Protein from the Yeast *Saccharomyces cerevisiae*. J. Biol. Chem. 270:4108-4114
60. Karaffa L, Kubicek C (2003) *Aspergillus niger* citric acid accumulation: do we understand this well working black box? Appl. Microbiol. Biotechnol. 61:189-196
61. Kawagoe M, Hyakumura K, Suye S-I et al. (1997) Application of bubble column fermentors to submerged culture of *Schizophyllum commune* for production of l-malic acid. J. Ferment. Bioeng. 84:333-336
62. Kenealy W, Zaady E, Du Preez JC et al. (1986) Biochemical Aspects of Fumaric Acid Accumulation by *Rhizopus arrhizus*. Appl. Environ. Microbiol. 52:128-133
63. Kertes AS, King CJ, Harvey W. Blanch IB (2009) Extraction chemistry of fermentation product carboxylic acids. Biotechnol. Bioeng. 103:431-445
64. Kissel M, Kubicek CP, Röhr M (1980) Influence of manganese on morphology and cell wall composition of *Aspergillus niger* during citric acid fermentation. Arch. Microbiol. 128:26-33
65. Kitamoto N, Matsui J, Kawai Y et al. (1998) Utilization of the TEF1-a gene (TEF1) promoter for expression of polygalacturonase genes, *pgaA* and *pgaB*, in *Aspergillus oryzae*. Appl. Microbiol. Biotechnol. 50:85-92
66. Knuf C, Nookae I, Brown SH et al. (2013) Investigation of Malic Acid Production in *Aspergillus oryzae* under Nitrogen Starvation Conditions. Appl. Environ. Microbiol. 79:6050-6058
67. Kornberg HL, Krebs HA (1957) Synthesis of Cell Constituents from C2-Units by a Modified Tricarboxylic Acid Cycle. Nature 179:988-991

68. Krappmann S, Braus GH (2003) Deletion of *Aspergillus nidulans* aroC using a novel blaster module that combines ET cloning and marker rescue. Mol. Genet. Genomics 268:675-683
69. Kubicek C (1988) The role of the citric acid cycle in fungal organic acid fermentations. Biochemical Society symposium 54:113-126
70. Kuhnert P, Scholten E, Haefner S et al. (2010) *Basfia succiniciproducens* gen. nov., sp. nov., a new member of the family Pasteurellaceae isolated from bovine rumen. Int. J. Syst. Evol. Microbiol. 60:44-50
71. Kulmburg P, Judewicz N, Mathieu M et al. (1992) Specific binding sites for the activator protein, ALCR, in the alcA promoter of the ethanol regulon of *Aspergillus nidulans*. J. Biol. Chem. 267:21146-21153
72. Kurzrock T, Weuster-Botz D (2011) New reactive extraction systems for separation of bio-succinic acid. Bioprocess Biosystems Eng. 34:779-787
73. Kurzrock T, Weuster-Botz D (2010) Recovery of succinic acid from fermentation broth. Biotechnol. Lett. 32:331-339
74. Lee P, Lee S, Hong S et al. (2002) Isolation and characterization of a new succinic acid-producing bacterium, *Mannheimia succiniciproducens* MBEL55E, from bovine rumen. Appl. Microbiol. Biotechnol. 58:663-668
75. Lee PC, Lee SY, Hong SH et al. (2003) Batch and continuous cultures of *Mannheimia succiniciproducens* MBEL55E for the production of succinic acid from whey and corn steep liquor. Bioprocess Biosystems Eng. 26:63-67
76. Li Q, Wang D, Hu G et al. (2011) Integrated bioprocess for high-efficiency production of succinic acid in an expanded-bed adsorption system. Biochem. Eng. J. 56:150-157
77. Li Q, Wang D, Wu Y et al. (2010) One step recovery of succinic acid from fermentation broths by crystallization. Separation and Purification Technology 72:294-300
78. Lin SKC, Du C, Blaga AC et al. (2010) Novel resin-based vacuum distillation-crystallisation method for recovery of succinic acid crystals from fermentation broths. Green Chemistry 12:666-671
79. Liu L, Liu J, Qiu RX et al. (2003) Improving heterologous gene expression in *Aspergillus niger* by introducing multiple copies of protein-binding sequence containing CCAAT to the promoter. Lett. Appl. Microbiol. 36:358-361
80. Liu R, Liang L, Chen K et al. (2012) Fermentation of xylose to succinate by enhancement of ATP supply in metabolically engineered *Escherichia coli*. Appl. Microbiol. Biotechnol. 94:959-968
81. Lumyong S, Tomita F (1993) L-malic acid production by an albino strain of *Monascus araneosus*. World J Microbiol Biotechnol 9:383-384
82. Luque R, Lin C, Du C et al. (2009) Chemical transformations of succinic acid recovered from fermentation broths by a novel direct vacuum distillation-crystallisation method. Green Chem 11:193-193
83. Machida M, Asai K, Sano M et al. (2005) Genome sequencing and analysis of *Aspergillus oryzae*. Nature 438:1157-1161

84. Mattern I, Unkles S, Kinghorn J et al. (1987) Transformation of *Aspergillus oryzae* using the *A. niger* *pyrG* gene. *Mol Gen Genet* 210:460-461
85. Mattey M (1992) The Production of Organic Acids. *Crit. Rev. Biotechnol.* 12:87-132
86. McGinn SM, Beauchemin KA, Coates T et al. (2004) Methane emissions from beef cattle: Effects of monensin, sunflower oil, enzymes, yeast, and fumaric acid. *J. Anim. Sci.* 82:3346-3356
87. McIntyre M, McNeil B (1997) Effects of elevated dissolved CO<sub>2</sub> levels on batch and continuous cultures of *Aspergillus niger* A60: an evaluation of experimental methods. *Appl. Environ. Microbiol.* 63:4171-4177
88. Meek JS (1975) The determination of a mechanism of isomerization of maleic acid to fumaric acid. *Journal of Chemical Education* 52:541
89. Meijer S, Otero J, Olivares R et al. (2009) Overexpression of isocitrate lyase—glyoxylate bypass influence on metabolism in *Aspergillus niger*. *Metab. Eng.* 11:107-116
90. Meyer V, Wanka F, Van Gent J et al. (2011) Fungal Gene Expression on Demand: an Inducible, Tunable, and Metabolism-Independent Expression System for *Aspergillus niger*. *Appl. Environ. Microbiol.* 77:2975-2983
91. Michielse C, Hooykaas PJ, Hondel CMJJ et al. (2005) Agrobacterium-mediated transformation as a tool for functional genomics in fungi. *Curr. Genet.* 48:1-17
92. Mizutani O, Kudo Y, Saito A et al. (2008) A defect of LigD (human Lig4 homolog) for nonhomologous end joining significantly improves efficiency of gene-targeting in *Aspergillus oryzae*. *Fungal Genet. Biol.* 45:878-889
93. Mizutani O, Masaki K, Gomi K et al. (2012) Modified Cre-loxP Recombination in *Aspergillus oryzae* by Direct Introduction of Cre Recombinase for Marker Gene Rescue. *Appl. Environ. Microbiol.* 78:4126-4133
94. Moon SY, Hong SH, Kim TY et al. (2008) Metabolic engineering of *Escherichia coli* for the production of malic acid. *Biochem. Eng. J.* 40:312-320
95. Nghiem N, Davison B, Suttle B et al. (1997) Production of succinic acid by *anaerobiospirillum succiniciproducens*. *Appl. Biochem. Biotechnol.* 63-65:565-576
96. Nielsen J (2001) Metabolic engineering. *Appl. Microbiol. Biotechnol.* 55:263-283
97. Nielsen J, Keasling JD (2011) Synergies between synthetic biology and metabolic engineering. *Nat Biotech* 29:693-695
98. Oh IJ, Kim DH, Oh EK et al. (2009) Optimization and Scale-Up of Succinic Acid Production by *Mannheimia succiniciproducens* LPK7. *J. Microbiol. Biotechnol.* 19:167-171
99. Oliveira A, Patil K, Nielsen J (2008) Architecture of transcriptional regulatory circuits is knitted over the topology of bio-molecular interaction networks. *BMC Systems Biology* 2:17
100. Orjuela A, Yanez AJ, Peereboom L et al. (2011) A novel process for recovery of fermentation-derived succinic acid. *Separation and Purification Technology* 83:31-37
101. Osmani SA, Scrutton MC (1983) The Sub-Cellular Localisation of Pyruvate Carboxylase and of Some Other Enzymes in *Aspergillus nidulans*. *European Journal of Biochemistry* 133:551-560

102. Pachlinger R, Mitterbauer R, Adam G et al. (2005) Metabolically Independent and Accurately Adjustable *Aspergillus* sp. Expression System. *Appl. Environ. Microbiol.* 71:672-678
103. Palmieri F (1994) Mitochondrial carrier proteins. *FEBS Lett.* 346:48-54
104. Papagianni M (2007) Advances in citric acid fermentation by *Aspergillus niger*: Biochemical aspects, membrane transport and modeling. *Biotechnol. Adv.* 25:244-263
105. Papagianni M, Matthey M (2006) Morphological development of *Aspergillus niger* in submerged citric acid fermentation as a function of the spore inoculum level. Application of neural network and cluster analysis for characterization of mycelial morphology. *Microbial Cell Factories* 5:3
106. Papagianni M, Matthey M, Kristiansen B (1999) The influence of glucose concentration on citric acid production and morphology of *Aspergillus niger* in batch and culture. *Enzyme Microb. Technol.* 25:710-717
107. Papanikolaou S, Muniglia L, Chevalot I et al. (2002) *Yarrowia lipolytica* as a potential producer of citric acid from raw glycerol. *J. Appl. Microbiol.* 92:737-744
108. Patil KR, Nielsen J (2005) Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proceedings of the National Academy of Sciences of the United States of America* 102:2685-2689
109. Payne GA, Nierman WC, Wortman JR et al. (2006) Whole genome comparison of *Aspergillus flavus* and *A. oryzae*. *Med. Mycol.* 44:9-11
110. Pedersen H, Beyer M, Nielsen J (2000) Glucoamylase production in batch, chemostat and fed-batch cultivations by an industrial strain of *Aspergillus niger*. *Appl. Microbiol. Biotechnol.* 53:272-277
111. Peksel A, Torres N, Liu J et al. (2002) <sup>13</sup>C-NMR analysis of glucose metabolism during citric acid production by *Aspergillus niger*. *Appl. Microbiol. Biotechnol.* 58:157-163
112. Pel HJ, De Winde JH, Archer DB et al. (2007) Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nat Biotech* 25:221-231
113. Peleg Y, Barak A, Scrutton M et al. (1989) Malic acid accumulation by *Aspergillus flavus*. *Appl. Microbiol. Biotechnol.* 30:176-183
114. Peleg Y, Rahamim E, Kessel M et al. (1988) Malic acid accumulation by *Aspergillus flavus*. *Appl. Microbiol. Biotechnol.* 28:76-79
115. Peleg Y, Rokem JS, Goldberg I (1990) A simple plate-assay for the screening of L-malic acid producing microorganisms. *FEMS Microbiol. Lett.* 67:233-236
116. Peleg Y, Rokem JS, Goldberg I et al. (1990) Inducible overexpression of the FUM1 gene in *Saccharomyces cerevisiae*: localization of fumarase and efficient fumaric acid bioconversion to L-malic acid. *Appl. Environ. Microbiol.* 56:2777-2783
117. Peleg Y, Stieglitz B, Goldberg I (1988) Malic acid accumulation by *Aspergillus flavus*. *Appl. Microbiol. Biotechnol.* 28:69-75
118. Pines O, Shemesh S, Battat E et al. (1997) Overexpression of cytosolic malate dehydrogenase (MDH2) causes overproduction of specific organic acids in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 48:248-255

119. Prathumpai W, McIntyre M, Nielsen J (2004) The effect of CreA in glucose and xylose catabolism in *Aspergillus nidulans*. *Appl. Microbiol. Biotechnol.* 63:748-753
120. Prömper C, Schneider R, Weiss H (1993) The role of the proton-pumping and alternative respiratory chain NADH:ubiquinone oxidoreductases in overflow catabolism of *Aspergillus niger*. *Eur. J. Biochem.* 216:223-230
121. Punt PJ, Oliver RP, Dingemanse MA et al. (1987) Transformation of *Aspergillus* based on the hygromycin B resistance marker from *Escherichia coli*. *Gene* 56:117-124
122. Radivojac P, Vacic V, Haynes C et al. (2010) Identification, analysis, and prediction of protein ubiquitination sites. *Proteins: Structure, Function, and Bioinformatics* 78:365-380
123. Rashid M, Javed M, Kawaguchi T et al. (2008) Improvement of *Aspergillus oryzae* for hyperproduction of endoglucanase: expression cloning of cmc-1 gene of *Aspergillus aculeatus*. *Biotechnol. Lett.* 30:2165-2172
124. Rathnasingh C, Raj SM, Lee Y et al. (2012) Production of 3-hydroxypropionic acid via malonyl-CoA pathway using recombinant *Escherichia coli* strains. *J. Biotechnol.* 157:633-640
125. Rhodes RA, Lagoda AA, Misenheimer TJ et al. (1962) Production of Fumaric Acid in 20-Liter Fermentors. *Applied Microbiology* 10:9-15
126. Roa Engel C, Straathof AJ, Zijlmans T et al. (2008) Fumaric acid production by fermentation. *Appl. Microbiol. Biotechnol.* 78:379-389
127. Romano AH, Bright MM, Scott WE (1967) Mechanism of Fumaric Acid Accumulation in *Rhizopus nigricans*. *J. Bacteriol.* 93:600-604
128. Ruijter GJG, Visser J (1997) Carbon repression in aspergilli. *FEMS Microbiol. Lett.* 151:103-114
129. Röhr M, Kubicek C (1981) Regulatory aspects of citric acid fermentation by *Aspergillus niger*. *Process Biochem.* 16:34-37
130. Sakai K, Kinoshita H, Nihira T (2012) Heterologous expression system in *Aspergillus oryzae* for fungal biosynthetic gene clusters of secondary metabolites. *Appl. Microbiol. Biotechnol.* 93:2011-2022
131. Sandor A, Johnson JH, Srere PA (1994) Cooperation between enzyme and transporter in the inner mitochondrial membrane of yeast. Requirement for mitochondrial citrate synthase for citrate and malate transport in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 269:29609-29612
132. Sauer M, Porro D, Mattanovich D et al. (2008) Microbial production of organic acids: expanding the markets. *Trends Biotechnol.* 26:100-108
133. Schmidt K, Marx A, De Graaf AA et al. (1998) <sup>13</sup>C tracer experiments and metabolite balancing for metabolic flux analysis: Comparing two approaches. *Biotechnol. Bioeng.* 58:254-257
134. Schmitt AP, McEntee K (1996) Msn2p, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences* 93:5777-5782
135. Scholten E, Renz T, Thomas J (2009) Continuous cultivation approach for fermentative succinic acid production from crude glycerol by *Basfia succiniciproducens* DD1. *Biotechnol. Lett.* 31:1947-1951

136. Schuster E, Dunn-Coleman N, Frisvad J et al. (2002) On the safety of *Aspergillus niger* – a review. *Appl. Microbiol. Biotechnol.* 59:426-435
137. Shen Y, Chen X, Peng B et al. (2012) An efficient xylose-fermenting recombinant *Saccharomyces cerevisiae* strain obtained through adaptive evolution and its global transcription profile. *Appl. Microbiol. Biotechnol.* 96:1079-1091
138. Shu P, Johnson MJ (1948) The Interdependence of Medium Constituents in Citric Acid Production by Submerged Fermentation. *J. Bacteriol.* 56:577-585
139. Song CW, Kim DI, Choi S et al. (2013) Metabolic engineering of *Escherichia coli* for the production of fumaric acid. *Biotechnol. Bioeng.* 110:2025-2034
140. Stephanopoulos G, Vallino J (1991) Network rigidity and metabolic engineering in metabolite overproduction. *Science* 252:1675-1681
141. Taing O, Taing K (2007) Production of malic and succinic acids by sugar-tolerant yeast *Zygosaccharomyces rouxii*. *Eur. Food Res. Technol.* 224:343-347
142. Takao S, Yokota A, Tanida M (1983) L-Malic Acid Fermentation by a Mixed Culture of *Rhizopus arrhizus* and *Paecilomyces varioti*. *Journal of fermentation technology* 61:643-645
143. Todaka N, Lopez C, Inoue T et al. (2010) Heterologous Expression and Characterization of an Endoglucanase from a Symbiotic Protist of the Lower Termite, *Reticulitermes speratus*. *Appl. Biochem. Biotechnol.* 160:1168-1178
144. Torres NV (1994) Modeling approach to control of carbohydrate metabolism during citric acid accumulation by *Aspergillus niger*: I. Model definition and stability of the steady state. *Biotechnol. Bioeng.* 44:104-111
145. Torres NV (1994) Modeling approach to control of carbohydrate metabolism during citric acid accumulation by *Aspergillus niger*: II. Sensitivity analysis. *Biotechnol. Bioeng.* 44:112-118
146. Tsuchiya K, Tada S, Gomi K et al. (1992) High level expression of the synthetic human lysozyme gene in *Aspergillus oryzae*. *Appl. Microbiol. Biotechnol.* 38:109-114
147. Valdehuesa K, Liu H, Nisola G et al. (2013) Recent advances in the metabolic engineering of microorganisms for the production of 3-hydroxypropionic acid as C3 platform chemical. *Appl. Microbiol. Biotechnol.* 97:3309-3321
148. Wang J, Zhu J, Bennett GN et al. (2011) Succinate production from different carbon sources under anaerobic conditions by metabolic engineered *Escherichia coli* strains. *Metab. Eng.* 13:328-335
149. Vemuri GN, Eiteman MA, Altman E (2002) Succinate production in dual-phase *Escherichia coli* fermentations depends on the time of transition from aerobic to anaerobic conditions. *J Ind Microbiol Biotech* 28:325-332
150. Verdoes JC, Punt PJ, Burlingame R et al. (2007) A dedicated vector for efficient library construction and high throughput screening in the hyphal fungus *Chrysosporium lucknowense*. *Industrial Biotechnology* 3:48-57
151. Werpy T, Petersen G (2004) Top value added chemicals from biomass: I. Results of screening for potential candidates from sugars and synthesis gas. U.S. Department of Energy, Washington ,DC.

152. Vongsangnak W, Hansen K, Nielsen J (2011) Integrated analysis of the global transcriptional response to  $\alpha$ -amylase over-production by *Aspergillus oryzae*. *Biotechnol. Bioeng.* 108:1130-1139
153. Vongsangnak W, Olsen P, Hansen K et al. (2008) Improved annotation through genome-scale metabolic modeling of *Aspergillus oryzae*. *BMC Genomics* 9:245
154. Xu G, Liu L, Chen J (2012) Reconstruction of cytosolic fumaric acid biosynthetic pathways in *Saccharomyces cerevisiae*. *Microbial Cell Factories* 11:24
155. Yao Z, Wu H, Liu H et al. (2008) Method for separation succinic acid from anaerobic fermentation broth. Chinese Patent CN200610086003.7. In:
156. Yedur S, Berglund K, Dunuwila D (2001) Succinic acid production and purification. US Patent 6:265,190. In:
157. Yu C-S, Chen Y-C, Lu C-H et al. (2006) Prediction of protein subcellular localization. *Proteins: Structure, Function, and Bioinformatics* 64:643-651
158. Yuill E (1950) The numbers of nuclei in conidia of aspergilli. *Transactions of the British Mycological Society* 33:324-IN310
159. Zeikus JG, Jain MK, Elankovan P (1999) Biotechnology of succinic acid production and markets for derived industrial products. *Appl. Microbiol. Biotechnol.* 51:545-552
160. Zelle RM, De Hulster E, Kloezzen W et al. (2010) Key Process Conditions for Production of C4 Dicarboxylic Acids in Bioreactor Batch Cultures of an Engineered *Saccharomyces cerevisiae* Strain. *Appl. Environ. Microbiol.* 76:744-750
161. Zelle RM, De Hulster E, Van Winden WA et al. (2008) Malic Acid Production by *Saccharomyces cerevisiae*: Engineering of Pyruvate Carboxylation, Oxaloacetate Reduction, and Malate Export. *Appl. Environ. Microbiol.* 74:2766-2777
162. Zelle RM, Harrison JC, Pronk JT et al. (2011) Anaplerotic Role for Cytosolic Malic Enzyme in Engineered *Saccharomyces cerevisiae* Strains. *Appl. Environ. Microbiol.* 77:732-738
163. Zelle RM, Trueheart J, Harrison JC et al. (2010) Phosphoenolpyruvate Carboxykinase as the Sole Anaplerotic Enzyme in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* 76:5383-5389
164. Zhang B, Skory CD, Yang S-T (2012) Metabolic engineering of *Rhizopus oryzae*: Effects of overexpressing pyc and pepc genes on fumaric acid biosynthesis from glucose. *Metab. Eng.* 14:512-520
165. Zhang X, Wang X, Shanmugam KT et al. (2011) l-Malate Production by Metabolically Engineered *Escherichia coli*. *Appl. Environ. Microbiol.* 77:427-434
166. Zhu N, Xia H, Yang J et al. (2013) Improved succinate production in *Corynebacterium glutamicum* by engineering glyoxylate pathway and succinate export system. *Biotechnol. Lett.*:1-8

Paper I

Investigation of Malic Acid Production in *Aspergillus oryzae*  
under Nitrogen Starvation Conditions

**Knuf C, Nookaew I, Brown SH, McCulloch M, Berry A, Nielsen J.**

Appl Environ Microbiol (2013) 79(19):6050-6058



# Investigation of Malic Acid Production in *Aspergillus oryzae* under Nitrogen Starvation Conditions

Christoph Knuf,<sup>a</sup> Intawat Nookaew,<sup>a</sup> Stephen H. Brown,<sup>b</sup> Michael McCulloch,<sup>b</sup> Alan Berry,<sup>b</sup> Jens Nielsen<sup>a</sup>

Department of Chemical and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden<sup>a</sup>; Novozymes, Inc., Davis, California, USA<sup>b</sup>

**Malic acid has great potential for replacing petrochemical building blocks in the future. For this application, high yields, rates, and titers are essential in order to sustain a viable biotechnological production process. Natural high-capacity malic acid producers like the malic acid producer *Aspergillus flavus* have so far been disqualified because of special growth requirements or the production of mycotoxins. As *A. oryzae* is a very close relative or even an ecotype of *A. flavus*, it is likely that its high malic acid production capabilities with a generally regarded as safe (GRAS) status may be combined with already existing large-scale fermentation experience. In order to verify the malic acid production potential, two wild-type strains, NRRL3485 and NRRL3488, were compared in shake flasks. As NRRL3488 showed a volumetric production rate twice as high as that of NRRL3485, this strain was selected for further investigation of the influence of two different nitrogen sources on malic acid secretion. The cultivation in lab-scale fermentors resulted in a higher final titer,  $30.27 \pm 1.05 \text{ g liter}^{-1}$ , using peptone than the one of  $22.27 \pm 0.46 \text{ g liter}^{-1}$  obtained when ammonium was used. Through transcriptome analysis, a binding site similar to the one of the *Saccharomyces cerevisiae* yeast transcription factor Msn2/4 was identified in the upstream regions of glycolytic genes and the cytosolic malic acid production pathway from pyruvate via oxaloacetate to malate, which suggests that malic acid production is a stress response. Furthermore, the pyruvate carboxylase reaction was identified as a target for metabolic engineering, after it was confirmed to be transcriptionally regulated through the correlation of intracellular fluxes and transcriptional changes.**

Malic acid belongs to the group of C<sub>4</sub> dicarboxylic acids, which are structurally similar to maleic acid and maleic anhydride, which represent key building blocks in the chemical industry. The C<sub>4</sub> dicarboxylic acids may therefore replace petrochemically derived compounds in the future, when increased oil and gas prices favor the production of renewable chemicals from biomass. The C<sub>4</sub> acids of interest, malic, succinic, and fumaric acids, are intermediates of the tricarboxylic acid (TCA) cycle and are naturally produced by many organisms. The first patent on malic acid production was filed in 1960 (1). The inventors selected an *Aspergillus flavus* strain to be the best natural producer and optimized the fermentation process for this organism, resulting in final titers of 58.4 g liter<sup>-1</sup> after 9 days of fermentation from minimal medium identical to MAF3 medium (see below) containing 0.2% ammonium sulfate and 100 g liter<sup>-1</sup> glucose. This represents a yield of 0.78 mol mol<sup>-1</sup> on glucose and a productivity of 0.27 g liter<sup>-1</sup> h<sup>-1</sup>. Furthermore, they investigated the effect of the nitrogen source, including, among others, peptone and ammonium sulfate, and reported final titers of 32.6 g liter<sup>-1</sup> and 30.4 g liter<sup>-1</sup>, respectively, after 7 days of fermentation from 100 g liter<sup>-1</sup> glucose. The same strain was used in the late 1980s and early 1990s for further investigation of the metabolism toward malic acid production. It was reported that *de novo* enzyme synthesis during nitrogen starvation resulted in an increase of malate synthesis, as malate dehydrogenase activity increased and fumarylase activity changed only slightly (2). In shake flasks and fermentors, the molar yield on glucose was 0.68 mol mol<sup>-1</sup> after 8 days and 0.57 mol mol<sup>-1</sup> after 6 days (2). Further <sup>13</sup>C nuclear magnetic resonance analysis of the produced malic acid showed that the majority of the acid was produced via the reductive TCA cycle branch, from pyruvate via oxaloacetate to malate (3). For *Aspergillus flavus*, the localization of pyruvate carboxylase was shown to be cytosolic (4), whereas isoenzymes could be found in the cytosol and mitochondrion in the case of the closely related organism *A. oryzae*. Isoenzymes of malate dehydro-

genase were shown to be active in both the cytosol and mitochondrion compartments for *Aspergillus flavus* (3). After optimization of the fermentation process, yields of 1.26 mol mol<sup>-1</sup> on glucose and a productivity of 0.59 g liter<sup>-1</sup> h<sup>-1</sup> were achieved in fermentors after 190 h of fermentation (5).

Though high yields and titers could be achieved using *A. flavus*, aflatoxin production by this organism disqualifies it for industrial malic acid production. Consequently, several attempts have been made to produce malic acid using highly engineered model organisms, like *Saccharomyces cerevisiae* and *Escherichia coli*. An *S. cerevisiae* strain overexpressing pyruvate carboxylase, malate dehydrogenase, and a malate exporter and carrying a pyruvate decarboxylase deletion reached malate yields of 0.42 mol mol<sup>-1</sup> on glucose at a productivity of 0.19 g liter<sup>-1</sup> h<sup>-1</sup> (6). Engineered *E. coli* strains could also reach very high molar yields and high productivities, e.g., 1.42 mol mol<sup>-1</sup> and 0.47 g liter<sup>-1</sup> h<sup>-1</sup> (7), in an anaerobic two-stage fermentation or 0.74 mol mol<sup>-1</sup> and 0.74 g liter<sup>-1</sup> h<sup>-1</sup> (8). The first strain was developed from a strain already engineered for succinic acid production and carried 11 gene deletions in total. In the latter strain, the ATP generation during the malic acid production process was changed by overexpressing the *Mannheimia succiniciproducens* phosphoenolpyruvate carboxykinase. The yields and productivities obtained with these engineered strains are similar to those obtained with the *A. flavus* wild-type

Received 4 May 2013 Accepted 21 July 2013

Published ahead of print 26 July 2013

Address correspondence to Jens Nielsen, nielsenj@chalmers.se.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AEM.01445-13>.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AEM.01445-13

strain. Comparative genomics of *A. flavus* and *A. oryzae* suggest that these are very close relatives or even ecotypes of the same species (9), which suggests that they have similar malic acid production capabilities. This leads to the question of whether *A. oryzae* strains are suitable for malic acid production, as well as what impact the nitrogen source has on malic acid production capacity.

In this study, we present *A. oryzae* as a cell factory for the production of malic acid which combines high malic acid production capabilities and production security using a class 1 organism, which would be preferred for industrial production. With the introduction of systems biology tools (10, 11) and the availability of the whole-genome sequence (12), high-throughput analysis has become possible. By using the *A. oryzae* genome-scale metabolic model (GEM) in combination with microarrays for transcriptome analysis, we investigated the malic acid production mechanisms and predicted metabolic engineering targets to further increase malic acid production yields and productivities to industrial targets.

## MATERIALS AND METHODS

**Strains.** Wild-type *Aspergillus oryzae* strains NRRL3485 and NRRL3488 were initially compared with each other. Strains NRRL3485 (DSM1862) and NRRL3488 were obtained from the German Collection of Microorganisms and Cell Cultures (Deutsche Sammlung von Mikroorganismen und Zellkulturen [DSMZ]) and the Agricultural Research Service (ARS) culture collection (Northern Regional Research Laboratory [NRRL]), respectively. The physiological characterization and transcriptome analysis were performed using NRRL3488. Spores of both strains were stored at  $-80^{\circ}\text{C}$  in 20% glycerol solution.

**Media.** The compositions of the spore propagation medium (Cove-N-Gly) and the preculture medium (G2-GLY) were described before (13). In this study, the preculture medium also contained 100 g/liter  $\text{CaCO}_3$ . *A. oryzae* malic acid fermentation (MAF3) medium consisted of 50 g liter $^{-1}$  glucose monohydrate, 0.15 g liter $^{-1}$   $\text{KH}_2\text{PO}_4$ , 0.15 g liter $^{-1}$   $\text{K}_2\text{HPO}_4$ , 0.10 g liter $^{-1}$   $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 0.10 g liter $^{-1}$   $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , 0.005 Fe $\text{SO}_4 \cdot 7\text{H}_2\text{O}$ , 0.005 g liter $^{-1}$  NaCl, 100.0 g liter $^{-1}$   $\text{CaCO}_3$ , and 1 ml liter $^{-1}$  pluronic acid (PE6100; BASF). The nitrogen sources used were either 1.4 g liter $^{-1}$   $(\text{NH}_4)_2\text{SO}_4$  or 6 g liter $^{-1}$  Bacto peptone.

**Preparation of inoculum.** Cove-N-Gly plates were inoculated with a suspension of spores from the  $-80^{\circ}\text{C}$  stock of each strain. The plates were incubated at  $30^{\circ}\text{C}$  for 7 days. The spores were subsequently harvested by addition of 10 ml of 0.01% Tween 80 solution. The spore suspension was used to inoculate the preculture medium (G2-GLY) with a final concentration of  $6 \times 10^9$  spores liter $^{-1}$ . The preculture was incubated with shaking at 250 rpm in a 500-ml shake flask without baffles for 24 h at  $30^{\circ}\text{C}$ . Thereafter, the main cultures for malic acid production in fermentors or shake flasks were inoculated with preculture broth using 10% of the final working volume.

**Batch cultivations.** For evaluation of the malic acid production capabilities, batch cultivations were carried out in shake flasks and fermentors. In the case of shake flask cultivations, the strains were incubated in 100 ml MAF3 medium supplied with peptone as the nitrogen source in 500-ml Erlenmeyer flasks with agitation at 250 rpm in an orbital shaker. The cultivations in bioreactors were performed in 2.7-liter Applikon bioreactors (Applikon, Schiedam, The Netherlands) with 2-liter working volumes. Reactors were equipped with two Rushton six-blade disc impellers, and the temperature was maintained at  $34^{\circ}\text{C}$ . The pH was buffered by the calcium carbonate in the medium. The temperature, agitation, gassing, pH, and composition of the off gas were monitored and/or controlled using a DasGip monitoring and control system (DasGip, Jülich, Germany). The stirrer speed was set to 950 rpm, and the aeration rate was 1 volume of gas per volume of fermentation broth per minute (vvm). The concentrations of oxygen and carbon dioxide in the exhaust gas

were analyzed with DasGip fedbatch-pro gas analysis systems with the off-gas analyzer GA4, based on zirconium dioxide and a two-beam infrared sensor.

**Sampling.** One fraction of the sample withdrawn from the fermentors was stored at  $-20^{\circ}\text{C}$  for subsequent analysis of extracellular metabolites. For quantification of cell mass, a known sample volume was treated with 2 N HCl in order to dissolve the calcium carbonate. The treated broth was centrifuged, and the pellet was washed once. After that, the wet biomass pellet was redissolved in distilled water and poured onto a preweighed aluminum dish. Aluminum dishes were kept at  $90^{\circ}\text{C}$  for 24 h in order to evaporate the water and subsequently stored in a desiccator until weighing.

**Metabolite analysis.** The concentrations of sugars and metabolites in the culture medium were determined by high-pressure liquid chromatography (HPLC). Organic acids were measured using a Synergi 4- $\mu\text{m}$  Hydro-RP 80-Å HPLC column (Phenomenex Ltd., Aschaffenburg, Germany), together with a Dionex Ultimate 3000 system and a photodiode array detector (Dionex, Sunnyvale, CA) at a wavelength of 210 nm. The eluent consisted of 145 mM phosphoric acid with 10% methanol at pH 3. Elution was performed isocratically at a flow rate of 0.7 ml/min and a temperature of  $20^{\circ}\text{C}$ .

Glucose was determined using an Aminex HPX-87H column from Bio-Rad (Bio-Rad, Sundbyberg, Sweden). The assay was performed at  $45^{\circ}\text{C}$  at a flow rate of 0.6 ml/min with a 5 mM sulfuric acid solution for isocratic elution.

**Enzyme assay.** Enzyme assays were performed with samples from shake flasks. A 50-ml sample of medium was poured into a strainer. The mycelia were washed with distilled water until most  $\text{CaCO}_3$  was removed. The washed mycelia were subsequently collected in a Falcon tube and stored at  $-20^{\circ}\text{C}$  until further use. For cell disruption, the frozen pellet was placed in a prechilled mortar and ground into an even, smooth paste. The paste was mixed into a smooth suspension after addition of 400  $\mu\text{l}$  of phosphate-buffered saline buffer (pH 7.4). The crude enzyme extract was separated from the debris by centrifugation at  $18,000 \times g$  and cleaned up using a Micro Bio spin column and SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) buffer (Bio-Rad). The clean flowthrough was immediately used for protein quantification and enzyme assay. The assay mixture for malate dehydrogenase contained 100 mM Tris, pH 8.0, 10 mM NaHCO<sub>3</sub>, 5 mM MgCl<sub>2</sub>, 0.667 mM oxaloacetic acid, 0.2 mM NADH. The reaction was started with the addition of the cell extract. The malate dehydrogenase activity was measured spectrophotometrically by monitoring NADH oxidation at 340 nm. The pyruvate carboxylase activity was measured indirectly by coupling the assay to the malate dehydrogenase reaction. The assay mixture for pyruvate carboxylase contained 100 mM Tris, pH 8.0, 10 mM NaHCO<sub>3</sub>, 5 mM MgCl<sub>2</sub>, 1 mM pyruvate, 0.2 mM NADH, concentrated malate dehydrogenase, and 1 mM ATP. The assays were performed at  $30^{\circ}\text{C}$  with freshly prepared extracts. One unit of enzyme converts 1  $\mu\text{mol}$  of substrate and coenzyme per minute, which, in this case, is 1  $\mu\text{mol}$  of NADH to NAD<sup>+</sup> per minute. The total protein concentration was measured with a NanoDrop 2000 apparatus (Thermo Scientific).

**RNA extraction.** Samples for RNA extraction were briefly treated with HCl in order to solubilize the  $\text{CaCO}_3$ . The samples were subsequently filtered through Miracloth (Calbiochem, San Diego, CA), and the filter cake was shock frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ . For RNA extraction, the frozen mycelium was ground to a powder in a prechilled mortar with a prechilled pestle. RNA was extracted from the mycelium powder using an RNeasy plant minikit (Qiagen) according to the protocol described by the manufacturer. The quality of the extracted RNA was checked with a 2100 Bioanalyzer apparatus (Agilent Technologies, Inc., Santa Clara, CA).

**Microarray processing.** Details concerning the microarray manufacturing and design, biotin-labeled cDNA and microarray processing, and microarray data acquisition can be found in a previously published report (13).

**Microarray analysis.** The microarrays were analyzed using the R program, version 2.15.1 (14), and the R package Piano (15) for microarray analysis. This package combines several other packages, as described below. The Affy package (16) was used to load the CEL files. Those were preprocessed using the Plier package (17) and applying cubic spline normalization. For further statistical pairwise analysis to determine significantly differentially expressed genes, the Limma package (18) was utilized for moderated Student's *t* test. The standard errors within each gene were moderated using empirical Bayesian statistics, and multiple testing was adjusted for by applying the Benjamini-Hochberg method (19). The *P* values and fold changes for each gene so obtained were further analyzed using the reporter algorithm (20) for gene ontology (GO) term analysis and reporter metabolites. For each gene set, distinct directional *P* values were calculated. The distinct directional *P* value is the outcome of two one-tailed *t* tests for either up- or downregulation, taking the transcriptional changes of all genes in the gene set into account. The case for the most significant distinct directional *P* value was chosen. Hence, the distinct directional *P* value, other than the directional *P* value, shows the overall significance of up- or downregulation of genes in a given gene set and not only the significance of the up- or downregulated genes in the gene set. For the reporter metabolite analysis, the topology was inferred from the *A. oryzae* GEM. The topology for the GO term analysis was downloaded from the AspGD *Aspergillus* genome database (21).

In order to identify transcriptionally regulated reactions, a correlation between expression and flux changes was searched for. For this, the method of random sampling (22) was applied to generate 3,292 possible flux distributions for each of the conditions and constraints considered according to the measured external fluxes. Loops were avoided by setting the default bounds for the metabolic fluxes to infinity and negative infinity instead of 1,000 and -1,000, as is usually done. In order to obtain significance scores for the flux change in each reaction, the first sample of condition A (exponential phase) was compared with all the samples of condition B (stationary phase). The number of times that condition A had a higher or lower value for each reaction was subsequently computed. Then, the same procedure was repeated for the second sample of condition A, and this procedure was continued until the last sample of condition A was analyzed. The probability score for a flux increase or decrease is reflected by the fraction of  $3,292 \times 3,292$  comparisons in which condition A had a higher or lower value than condition B. This probability score was subsequently compared with the probability scores of transcriptional changes. A correlation of the significance scores of both flux and transcript hints toward a transcriptional regulation of this reaction.

## RESULTS

**Physiological investigation.** An initial screening of wild-type *A. oryzae* strains showed broad malic acid production capabilities. To illustrate this, the two wild-type *A. oryzae* strains NRRL3485 and NRRL3488, which were located at the lower and upper ends of the malic acid production spectrum, respectively, were cultivated over 78 h in shake flasks containing MAF3 medium with 6 g liter<sup>-1</sup> peptone as the nitrogen source for initial biomass production. Strain NRRL3485 secreted malic acid at a volumetric rate of  $0.299 \pm 0.011$  g liter<sup>-1</sup> h<sup>-1</sup>, whereas NRRL3488 produced malic acid at almost twice that rate of  $0.563 \pm 0.020$  g liter<sup>-1</sup> h<sup>-1</sup>. The final titers were  $23.12 \pm 0$  g liter<sup>-1</sup> and  $38.86 \pm 2.80$  g liter<sup>-1</sup> for NRRL3485 and NRRL3488, respectively.

The faster-producing strain, NRRL3488, was selected for further evaluation of malic acid production during nitrogen starvation in well-controlled lab-scale bioreactors. The medium contained either the complex nitrogen source peptone or the nitrogen-containing salt ammonium sulfate, which is more suitable for use in industrial large-scale processes, as it is cheaper and simplifies downstream processing. The batch cultivations were performed in quadruplicate experiments for

each nitrogen source. The results are presented in Fig. 1, which illustrates the profiles of the averaged biomass, the concentrations of the extracellular metabolites of malate, succinate, and citrate, and glucose concentrations for the peptone (Fig. 1A) and ammonium (Fig. 1B) conditions. The maximum specific growth rates, the rates of malic acid volumetric production, the rate of glucose consumption, and the molar yields of the organic acids in the exponential and stationary phases for each nitrogen source are summarized in Table 1.

Comparison of the growth curves of NRRL3488 revealed similar behaviors on peptone and ammonium. The strain showed exponential growth for the first 12 to 15 h of fermentation with a slightly elevated maximum specific growth rate and prolonged exponential phase on peptone, resulting in a biomass concentration of about 5 g liter<sup>-1</sup> on peptone in comparison to one of 4 g liter<sup>-1</sup> on ammonium at the end of the exponential phase. At this point, nitrogen limitation was confirmed by measurement of ammonium concentrations in the medium in case of ammonium fermentation. Thereafter, an adaptation phase, during which the biomass increased under both conditions, but to a higher extent on peptone, bridged over to a stationary phase, in which the biomass stayed constant on peptone and ammonium at about 8.5 g liter<sup>-1</sup> and 6 g liter<sup>-1</sup>, respectively.

During the exponential growth phase, only malate secretion could be detected in significant amounts under both conditions. On peptone, small amounts of succinate and citrate were detected, and on ammonium, marginal succinate and citrate secretion was detected. During the stationary phase, malate, citrate, succinate, and pyruvate were detected in the fermentation broth under both conditions. The yields of malate on a glucose base were tripled during the stationary phase, resulting in final titers of malate of  $30.27 \pm 1.05$  g liter<sup>-1</sup> in the peptone case and  $22.27 \pm 0.46$  g liter<sup>-1</sup> for the ammonium cultivations. This substantial secretion of organic acids, mainly malate, raised the question about the underlying regulation and mechanisms of the switch between exponential growth, with biomass production as the objective function, and the stationary phase, in which half of the metabolized glucose is directly converted to malic acid and secreted into the fermentation broth. In order to further investigate this issue, comparative transcriptome analysis was performed.

**Comparative transcriptome analysis.** In order to understand the key processes in the shift toward malic acid production and the response of the cell to nitrogen deprivation, results of transcription analysis of samples taken in mid-exponential phase (6 h) and stationary phase (30 h) from triplicate cultivations were compared pairwise for each nitrogen source using Student's *t*-test statistics. For the statistical analysis of the transcriptional changes, Student's *t* test with an adjusted-*P*-value cutoff of 0.005 was applied. In total, the expression of 3,766 genes changed significantly under the peptone condition; of these, 946 genes had lower levels of expression and 2,820 had higher levels of expression in response to nitrogen starvation (30 h versus 6 h). Among the 4,540 differentially expressed genes under the ammonium condition, 28% (1,306 genes) showed a lower level of expression in response to nitrogen starvation. For further investigation, reporter metabolite analysis and gene ontology (GO) term analysis were performed using the reporter function of the R package Piano.

**Reporter feature analysis of the transcriptome data.** For the reporter metabolite analysis, the gene metabolite network was inferred from an updated version of the published *A. oryzae* model

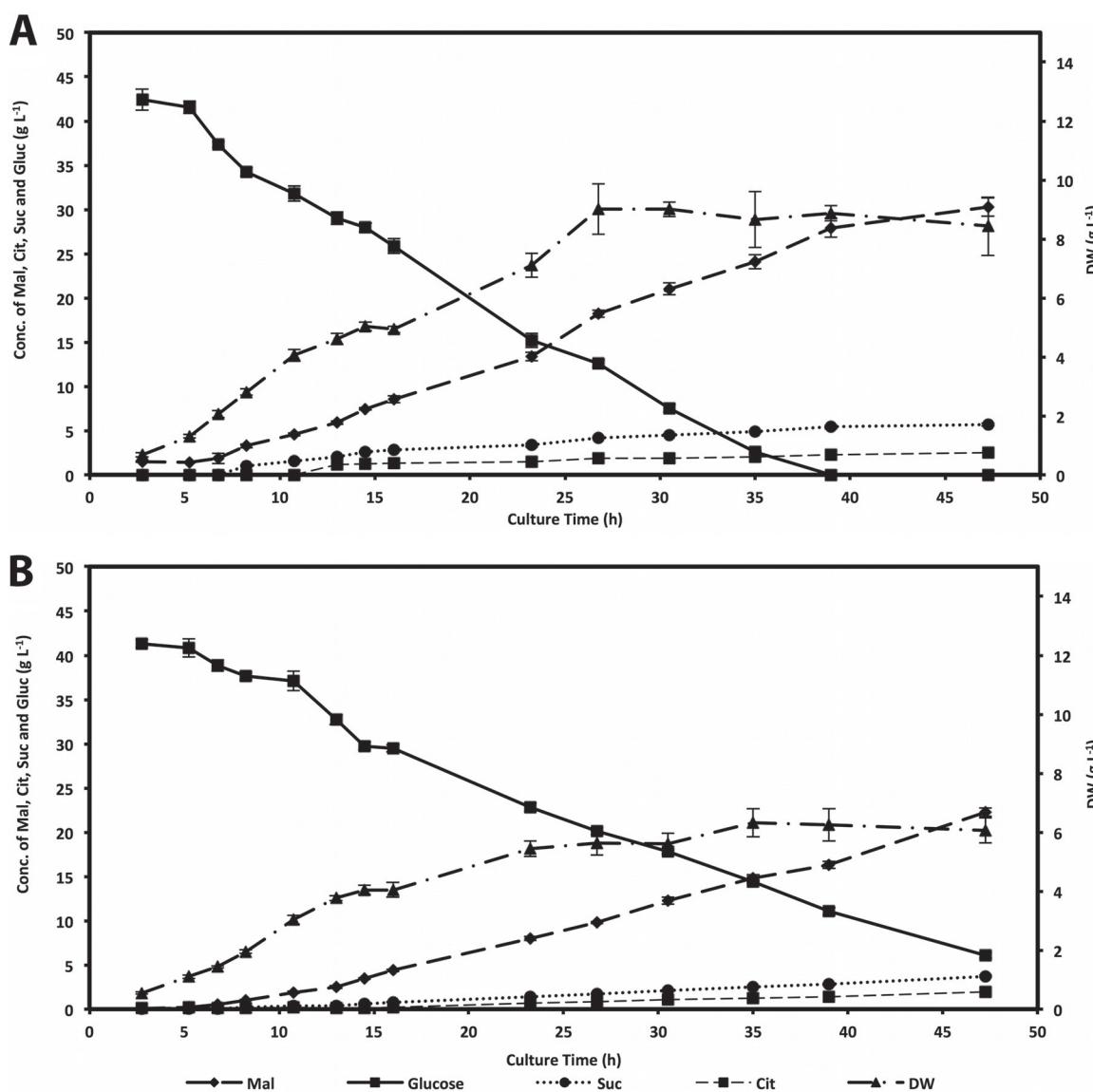


FIG 1 Extracellular metabolite concentrations during bioreactor cultivations of NRRL3488 in MAF medium containing  $6 \text{ g liter}^{-1}$  peptone (A) or  $1.4 \text{ g liter}^{-1}$  ammonium sulfate (B) in 2.7-liter Applikon bioreactors with a working volume of 2 liters controlled by a DasGip control system. The results shown are averages and standard errors of 4 reactors. Mal, malate; Cit, citrate; Suc, succinate; Gluc, glucose; DW, dry weight.

(10), which is available in the BioMet tool box (23). The reporter metabolites with a distinct directional *P* value of less than 0.001 in either of the two comparisons are shown in Fig. 2B. A full list of the results is provided in Data Set S2 in the supplemental material. Among those 59 metabolites around which the most significant transcriptional changes occurred, there were only 8 metabolites with a significant upregulation of the correlated genes. It was not surprising to find both intra- and extracellular ammonia among the upregulated reporter metabolites in nitrogen-starved cells. Allantoate and urate, which are intermediates or end products of purine metabolism, were found, as were two metabolites of glutathione metabolism and the end product of glutathione metabolism, glutathione. The last identified metabolite was UDP, which is involved in the polymerization of glycogen. Among the residual group of 51 metabolites associated with repressed expression, 8 metabolites that are involved in amino acid synthesis and 8 that

are involved in the TCA cycle were found, and 10 were energy or reduction equivalents.

Another method for correlating gene expression and biological processes is GO term analysis. For this analysis, only the ontology of biological processes was used as input for the reporter features algorithm (24). A heat map displaying the most significantly changed GO terms that were characterized by a distinct directional *P* value of less than 0.001 under at least one condition is represented in Fig. 2A. A full list of the results is provided in Data Set S3 in the supplemental material. Among the 15 induced processes, significant changes in piecemeal microautophagy of the nucleus, the purine base catabolic process, protein ubiquitination, or conidium formation were found. The list of GO terms associated with transcriptional repression reflects the same general trend of reduced cellular viability shown by the induction-associated GO terms. In association with the scarcity of available nitro-

TABLE 1 Physiological data FOR NRRL3488 grown in malic acid fermentation medium with either peptone or ammonium as the nitrogen source<sup>a</sup>

Nitrogen source	Final titer of malate (g liter <sup>-1</sup> )	Phase <sup>b</sup>	$\mu_{\max}^c$ (h <sup>-1</sup> )	$r_{\text{malate}}^d$ (mmol liter <sup>-1</sup> h <sup>-1</sup> )	$r_s^e$ (mmol liter <sup>-1</sup> h <sup>-1</sup> )	Yield on glucose (mol mol <sup>-1</sup> )			
						Citrate	Malate	Succinate	Pyruvate
Peptone	30.27 ± 1.05	Exp.	0.23 ± 0.01	4.22 ± 0.25	8.10 ± 0.81	ND <sup>f</sup>	0.33 ± 0.05	ND	ND
		Stat.		6.61 ± 0.57	6.13 ± 0.34	0.03 ± 0.01	0.98 ± 0.13	0.14 ± 0.03	0.02 ± 0.01
Ammonium	22.27 ± 0.46	Exp.	0.21 ± 0.05	1.59 ± 0.15	3.71 ± 0.44	0.01 ± 0.01	0.34 ± 0.06	0.07 ± 0.02	ND
		Stat.		4.36 ± 0.14	3.92 ± 0.10	0.07 ± 0.01	1.09 ± 0.05	0.20 ± 0.01	0.03 ± 0.00

<sup>a</sup> The numbers stated are the means of four individual bioreactors ± standard errors.

<sup>b</sup> Exp., exponential growth phase; Stat., stationary phase.

<sup>c</sup>  $\mu_{\max}$ , maximum specific growth rate.

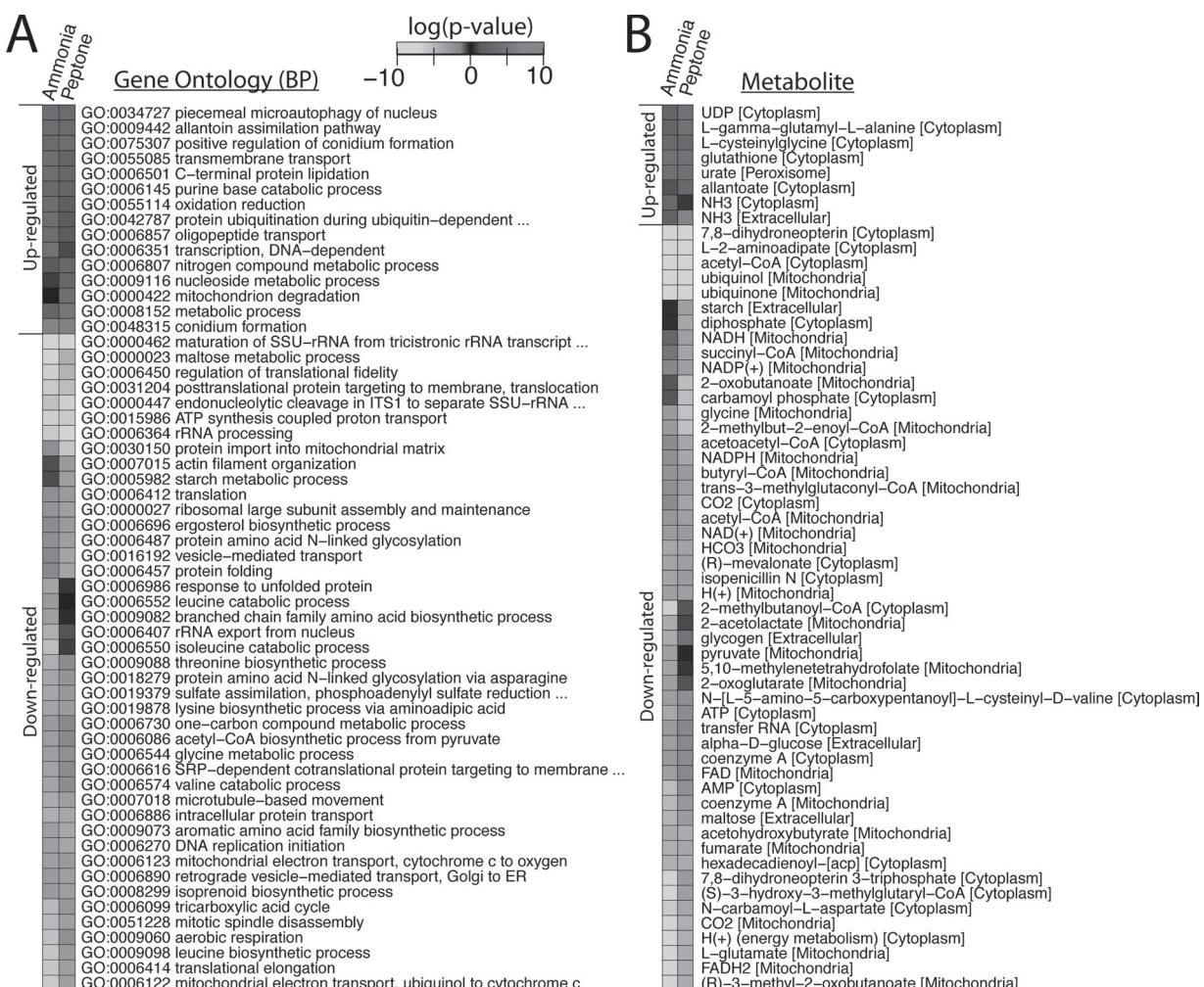
<sup>d</sup>  $r_{\text{malate}}$ , specific malate production rate.

<sup>e</sup>  $r_s$ , substrate consumption rate.

<sup>f</sup> ND, not determined.

gen, several biosynthetic processes for amino acid synthesis seemed to be repressed. Consistent with the low availability of amino acids, GO terms associated with protein synthesis, starting from translation/translational elongation and continuing to pro-

tein folding and intracellular protein transport, were correlated with transcriptional repression. As the availability of nitrogen limited cell growth, the largest sink of energy in the cell was absent. The cells seemed to respond to this with transcriptional repression



**FIG 2** Heat map of overrepresented biological process GO terms (A) and reporter metabolites (B) in the comparison of the nitrogen starvation phase to exponential growth phase. GO terms and reporter metabolites with *P* values of at most 0.005 under either of the conditions with peptone or ammonium are shown. BP, biological process; SSU, small subunit; CoA, coenzyme A; SRP, signal recognition particle; Golgi, Golgi apparatus; ER, endoplasmic reticulum; acp, acyl carrier protein.

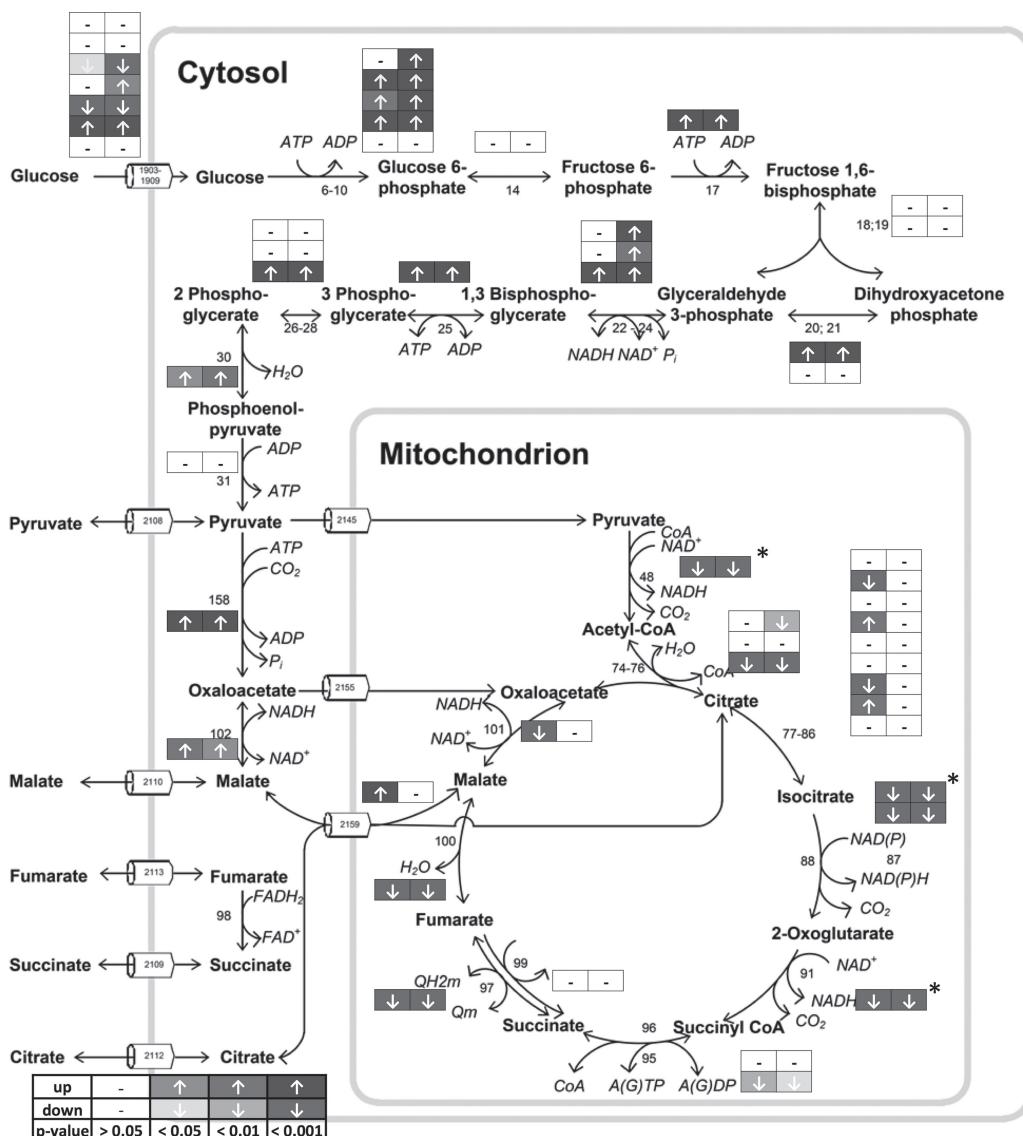


FIG 3 Schematic drawing of the reactions related to malic acid secretion, including the significance and direction of transcriptional changes of the genes encoding the enzymes catalyzing those steps. The reaction arrows show the direction of the reaction stated in the GEM of *A. oryzae*. Double-headed arrows indicate reactions assumed to be reversible. The numbers next to the reaction arrows correlate with the reaction numbers in the model. The arrowed fields indicate the significance of transcriptional change in the comparison of the stationary phase to the exponential phase, with the ammonium condition indicated in the left set of boxes and the peptone condition indicated in the right set of boxes. The darker that the shading is, the more significant that the change is. The direction of transcriptional change is indicated by the direction of the arrows: up arrows, upregulation in stationary phase; down arrows, downregulation in stationary phase. Asterisks, data for the most significantly changed genes of an enzyme complex.

of energy-supplying processes, such as aerobic respiration and processes connected to that, e.g., mitochondrial electron transport or ATP synthesis-coupled proton transport.

Taken together, the reporter metabolite and GO term analyses point toward a transcriptional response of the cell to the nitrogen starvation stress that leads to a degradation of cellular components, degradation of nitrogen-containing compounds in order to recycle the nitrogen, and reduced nitrogen consumption in protein production. However, the findings so far do not directly point toward an explanation for the high level of malic acid secretion. To shed light on the malic acid production pathway, the *P* values of the reactions related to it were plotted onto a metabolic map, including the pathway from glucose to malic acid, taking the di-

rection of fold change into account (Fig. 3). From this map it becomes obvious that glycolysis seems to be upregulated at the transcriptional level. Except for the glucose-6-phosphate isomerase, fructose bisphosphate aldolase, and pyruvate kinase steps, for the other reactions from glucose to malic acid, at least one gene involved in those reactions was more highly expressed in response to nitrogen starvation under both conditions. On the other hand, the general trend for the expression of genes involved in the TCA cycle was for these to be downregulated, which is in agreement with the results of the GO term analysis regarding the reduced ATP generation during oxidative phosphorylation. In order to find possible transcription factors that regulate the overexpression of the genes involved in the malic acid production pathway, the

upstream sequences of the genes that showed a significant induction during the starvation phase were investigated for conserved sequences. The upstream sequences of the genes of interest were retrieved from the web service regulatory sequence analysis tools (RSAT) using the retrieve sequence function, and subsequently, the oligonucleotide analysis tool (25) of the motif discovery functions. Two conserved motifs were discovered, kwCCCCCTCyy and kbbCACCGGTGvvm (where k is G or T, w is A or T, y is C or T, b is C, G, or T, and v is A, C, or G). The conserved 6-oligonucleotide sequence contained in the first motif (CCCCTC) has an occurrence *P* value of 6.6E-06. By passing this pattern on to the pattern-matching tool of the web service interface YEASTRACT (26), we identified this to be the binding site for the *Saccharomyces cerevisiae* yeast transcription factor Msn2/4, which is the transcriptional activator of the multistress response in *S. cerevisiae* (27).

**Identification of transcriptionally regulated fluxes.** In the next step of the analysis of malic acid production under conditions of nitrogen starvation, we identified transcriptionally regulated reactions by correlating the changes in flux and the transcription of genes encoding the enzymatic steps. Those transcriptionally regulated reactions are possible targets for metabolic engineering, as a simple overexpression of the gene should directly positively influence the flux of the corresponding reaction. By using a random sampling approach, the average and standard deviation for each flux were calculated for the reactions in the *A. oryzae* GEM, using the measured exchange fluxes as constraints (the GEM in XLSX format, including constraints, can be found in Data Set S4 in the supplemental material) (22). For the calculation of exchange fluxes, the physiological data for the ammonium fermentation were used. During the starvation phase, the biomass exceeded the theoretical biomass. As the biomass concentration has a big influence on the specific rates and it is inherently difficult to determine the biomass concentration exactly from filamentous fungal fermentations, a theoretical value was calculated from the amount of nitrogen supplied using elemental biomass compositions from previous studies. Very accurate measurements for *A. oryzae* have been obtained by Pedersen et al. (28), but the samples were taken from chemostats with a maximum growth rate of 0.17 h<sup>-1</sup>. In order to get a better estimate for unlimited growth, the elemental biomass composition for *A. niger* for unlimited growth was used, and a maximum biomass concentration of 4.19 g liter<sup>-1</sup> was calculated (29). As this value correlates well with the biomass concentration at the end of the exponential growth phase, the specific rates during the stationary phase were calculated using this biomass concentration for better comparison of the rates during exponential and stationary phases. After obtaining the flux distribution for each of the two conditions, the significance of the change in each reaction rate could be calculated. By correlating the *P* values of the transcriptional and flux changes for each reaction, genes that have direct transcriptional control over the flux of the reaction could be identified (22). In total, expression and theoretical flux were significantly positively correlated for only two genes, encoding the cytosolic pyruvate carboxylase (AspGD accession number AO090023000801 [[http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090023000801&organism=A\\_oryzae\\_RIB40](http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090023000801&organism=A_oryzae_RIB40)]) and lactate dehydrogenase (AspGD accession number AO090023000577 [[http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090023000577&organism=A\\_oryzae\\_RIB40](http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090023000577&organism=A_oryzae_RIB40)]).

A list with negatively correlated reactions can be found in Data Set S5 in the supplemental material. As those transcriptionally

regulated genes are prone to be regulated directly by transcription factors, the same procedure described above was followed to find regulatory sequences. The pattern matching resulted in the conserved consensus sequence vwTCAATTGAwb. The possible recognition pattern CAATTG had an occurrence *P* value of 4.9E-10. The search for the yeast transcription factor known to bind to this sequence did not return any hits.

In order to investigate the feasibility of the overexpression of pyruvate carboxylase, the enzyme activities of pyruvate carboxylase and malate dehydrogenase were measured in the exponential (6 h) and stationary (48 h) phases on MAF medium with ammonium and a glucose concentration of 100 g liter<sup>-1</sup> in shake flasks. The activity of pyruvate carboxylase increased slightly from 0.024 ± 0.004 to 0.033 ± 0.007 units mg<sup>-1</sup> total protein. In the malate dehydrogenase case, the activity decreased slightly from 4.848 ± 0.828 to 4.304 ± 0.358 units mg<sup>-1</sup> total protein.

The changes in activities were expected to differ more significantly. The reasoning for this expectation was based on the fold changes in expression and because the enzymes were identified to be transcriptionally regulated. Therefore, the possibility of ubiquitination and a resulting faster protein degradation were considered. The ubiquitination sites of the genes involved in the malic acid production pathway, *pyc* (AspGD accession number AO090023000801 [[http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090023000801&organism=A\\_oryzae\\_RIB40](http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090023000801&organism=A_oryzae_RIB40)]), *mdhA* (AspGD accession number AO090701000013 [[http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090701000013&organism=A\\_oryzae\\_RIB40](http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090701000013&organism=A_oryzae_RIB40)]), and *mae3* (AspGD accession number AO090023000318 [[http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090023000318&organism=A\\_oryzae\\_RIB40](http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090023000318&organism=A_oryzae_RIB40)]), were predicted using the UbPred program (30). For *mdhA* and *mae3*, ubiquitination sites were predicted with low confidence, whereas two sites were found in *pyc* with high confidence.

## DISCUSSION

An initial test of two wild-type *A. oryzae* strains led to the selection of NRRL3488 for further investigation. NRRL3488 was capable of producing malic acid in shake flasks with peptone as the nitrogen source at a volumetric rate of 0.563 ± 0.020 g liter<sup>-1</sup> h<sup>-1</sup>. This productivity was almost twice as high as the one of NRRL3485, indicating a large variation in malic acid production among *A. oryzae* wild-type strains. Therefore, further investigation of high malic acid production in lab-scale bioreactors using either peptone or ammonium as the nitrogen source was conducted with NRRL3488. The productivity during the stationary phase was 34% higher in the case of the peptone cultivations than in the case of the ammonium cultivations; nevertheless, the yields did not differ significantly. This might be due to the same metabolic efficiency but different biomasses built up from the amount of nitrogen available in the two cultivations.

NRRL3488 showed high final titers, productivities, and yields during cultivation for 47.5 h on ammonium salt medium. Though the transient yield of NRRL3488 (1.09 mol mol<sup>-1</sup>) was slightly lower than that previously reported for *A. flavus* (1.26 mol mol<sup>-1</sup>), the volumetric productivities were almost equal: 0.59 g liter<sup>-1</sup> h<sup>-1</sup> in the case of *A. flavus* and 0.58 g liter<sup>-1</sup> h<sup>-1</sup> in the case of NRRL3488 (5). The productivity and yield were more than twice as high in NRRL3488 than in a highly engineered *S. cerevisiae* strain (6) during the acid production phase. In comparison with engineered *E. coli* strains, strain *E. coli* XZ658 (7) showed a higher

molar yield, but the productivity of NRRL3488 was about 23.4% higher. On the other hand, *E. coli* WGS-10 (8) showed a lower yield but a 21.1% higher productivity than NRRL3488.

In order to investigate the regulatory mechanisms of this massive secretion of malic acid by a wild-type *A. oryzae* strain, transcription analysis using DNA microarrays was used. The general trend indicated by GO term and reporter metabolite analyses comparing the stationary phase with the exponential growth phase showed that the cells respond to nitrogen starvation by recycling nitrogen by degradation of proteins and other nitrogen-containing cellular compounds and reducing protein synthesis.

The metabolic function responsible for this might be an increased ubiquitination of proteins. Two strong ubiquitination sites were predicted for pyruvate carboxylase but not for other proteins directly involved in malic acid production. Nevertheless, this potential ubiquitination might explain the need for a more significant upregulation of *pyc* than malate dehydrogenase gene expression and still result in enzyme activity slightly higher than that during the exponential growth phase. As the *in vitro* specific activity of pyruvate carboxylase is 2 orders of magnitude lower than that of malate dehydrogenase, it seems to be the rate-limiting step in the malic acid production pathway.

As nitrogen depletion hampers cellular growth, the generation of ATP via oxidative phosphorylation and the production of reduction equivalents are not needed anymore. This could also be seen from the GO terms, which contained genes with significant repression, and reporter metabolites, around which a general repression of genes was detected.

The regulation of energy metabolism might be the reason for malic acid secretion. During the stationary phase, the cells are not able to respire glucose to carbon dioxide, as this would result in the production of large amounts of ATP that the cells cannot use due to a lack of growth. The advantage of malic acid production from glucose through the reductive cytosolic TCA branch from pyruvate via oxaloacetate is that it is completely balanced in terms of ATP and NADH production from glycolysis to pyruvate and consumption in the subsequent steps to malate. This pathway enables *A. oryzae* to produce large amounts of malic acid uncoupled from growth, which makes it a very interesting trait for future industrial applications. From an ecological and evolutionary point of view, direct glucose conversion to malic acid makes sense, as (i) glucose consumption can continue, reducing the availability for competing microorganisms, like many fast-growing bacteria; (ii) microbial growth is oppressed by low pH, whereas aspergilli are tolerant to acidic pHs (*A. oryzae*, for example, shows optimal growth over a broad pH range from 3 to 7 [31]); and (iii) sequence analysis of *A. oryzae* showed that it has the largest expansion of secretory hydrolytic enzymes that work at low pH, in comparison to the expansion for *A. nidulans* and *A. fumigatus* (12). This might help the cell to survive in an environment with only complex nitrogen sources available. In order to achieve acidification of the medium up to the lower end of the growth optimum, malic acid production is very efficient in its acidifying potential, with pK<sub>a</sub>s at 3.46 and 5.10.

The switch between ATP generation and malic acid secretion might be regulated by a transcription factor binding to the same recognition pattern as the yeast transcription factor Msn2/4. This transcriptional activator is known to positively affect transcription of genes under stressed conditions, such as conditions of osmotic, temperature, and nitrogen stress in budding yeast (32).

Taken together, we present a natural malate-producing microorganism that does not have to be disqualified for commercial large-scale production because of either the production of toxic by-products or a dependence on complex medium. The yield and productivity of the organism are comparable to those of highly engineered *E. coli* strains and exceed those of *S. cerevisiae*, leaving space for additional improvement by metabolic engineering. One target for this might be the overexpression of the pyruvate carboxylase, which was identified to be transcriptionally regulated and seems to be a rate-limiting step in malic acid biosynthesis.

## ACKNOWLEDGMENTS

We thank Sergio Bordel for assisting in identifying transcriptionally regulated reactions and Wanwipa Vongsangnak for valuable discussions.

We acknowledge funding for our research activities from the Chalmers Foundation, the Knut and Alice Wallenberg Foundation, the European Research Council (grant no. 247013), and Novozymes, Inc.

## REFERENCES

1. Abe S, Akira F, Ken-Ichiro T. November 1962. Method of producing L-malic acid by fermentation. US patent 3,063,910.
2. Peleg Y, Stieglitz B, Goldberg I. 1988. Malic acid accumulation by *Aspergillus flavus*. Appl. Microbiol. Biotechnol. 28:69–75.
3. Peleg Y, Barak A, Scrutton M, Goldberg I. 1989. Malic acid accumulation by *Aspergillus flavus*. Appl. Microbiol. Biotechnol. 30:176–183.
4. Bercovitz A, Peleg Y, Battat E, Rokem JS, Goldberg I. 1990. Localization of pyruvate carboxylase in organic acid-producing *Aspergillus* strains. Appl. Environ. Microbiol. 56:1594–1597.
5. Battat E, Peleg Y, Bercovitz A, Rokem JS, Goldberg I. 1991. Optimization of L-malic acid production by *Aspergillus flavus* in a stirred fermentor. Biotechnol. Bioeng. 37:1108–1116.
6. Zelle RM, de Hulster E, van Winden WA, de Waard P, Dijkema C, Winkler AA, Geertman J-M, van Dijken JP, Pronk JT, van Maris AJA. 2008. Malic acid production by *Saccharomyces cerevisiae*: engineering of pyruvate carboxylation, oxaloacetate reduction, and malate export. Appl. Environ. Microbiol. 74:2766–2777.
7. Zhang X, Wang X, Shanmugam KT, Ingram LO. 2011. L-Malate production by metabolically engineered *Escherichia coli*. Appl. Environ. Microbiol. 77:427–434.
8. Moon SY, Hong SH, Kim TY, Lee SY. 2008. Metabolic engineering of *Escherichia coli* for the production of malic acid. Biochem. Eng. J. 40:312–320.
9. Payne GA, Nierman WC, Wortman JR, Pritchard BL, Brown D, Dean RA, Bhatnagar D, Cleveland TE, Machida M, Yu J. 2006. Whole genome comparison of *Aspergillus flavus* and *A. oryzae*. Med. Mycol. 44:9–11.
10. Vongsangnak W, Olsen P, Hansen K, Krogsbaard S, Nielsen J. 2008. Improved annotation through genome-scale metabolic modeling of *Aspergillus oryzae*. BMC Genomics 9:245. doi:10.1186/1471-2164-9-245.
11. Andersen MR, Vongsangnak W, Panagiotou G, Salazar MP, Lehmann L, Nielsen J. 2008. A trispecies Aspergillus microarray: comparative transcriptomics of three Aspergillus species. Proc. Natl. Acad. Sci. U. S. A. 105:4387–4392.
12. Machida M, Asai K, Sano M, Tanaka T, Kumagai T, Terai G, Kusumoto K-I, Arima T, Akita O, Kashiwagi Y, Abe K, Gomi K, Horiuchi H, Kitamoto K, Kobayashi T, Takeuchi M, Denning DW, Galagan JE, Nierman WC, Yu J, Archer DB, Bennett JW, Bhatnagar D, Cleveland TE, Fedorova ND, Gotoh O, Horikawa H, Hosoyama A, Ichinomiya M, Igarashi R, Iwashita K, Juvvadi PR, Kato M, Kato Y, Kin T, Kokubun A, Maeda H, Maeyama N, Maruyama J-I, Nagasaki H, Nakajima T, Oda K, Okada K, Paulsen I, Sakamoto K, Sawano T, Takahashi M, Takase K, Terabayashi Y, Wortman JR, et al. 2005. Genome sequencing and analysis of *Aspergillus oryzae*. Nature 438:1157–1161.
13. Vongsangnak W, Hansen K, Nielsen J. 2011. Integrated analysis of the global transcriptional response to α-amylase over-production by *Aspergillus oryzae*. Biotechnol. Bioeng. 108:1130–1139.
14. The R Development Core Team. 2012. R: a language and environment for statistical computing. The R Foundation for Statistical Computing, Vienna, Austria.
15. Varemo L, Nielsen J, Nookaew I. 2013. Enriching the gene set analysis of

- genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic Acids Res.* 41: 4378–4391.
16. Gautier L, Cope L, Bolstad BM, Irizarry RA. 2004. Affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 20:307–315.
  17. Affymetrix Inc, Crispin JM, PICR. plier: implements the Affymetrix PLIER algorithm. R package, version 1.28.0. Affymetrix, Santa Clara, CA.
  18. Smyth GK. 2005. Limma: linear models for microarray data, p 397–420. In Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W (ed), *Bioinformatics and computational biology solutions using {R} and bioconductor*. Springer, New York, NY.
  19. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57:289–300.
  20. Patil KR, Nielsen J. 2005. Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proc. Natl. Acad. Sci. U. S. A.* 102:2685–2689.
  21. Arnaud MB, Cerqueira GC, Inglis DO, Skrzypek MS, Binkley J, Chibucus MC, Crabtree J, Howarth C, Orvis J, Shah P, Wymore F, Binkley G, Miyasato SR, Simison M, Sherlock G, Wortman JR. 2012. The Aspergillus Genome Database (AspGD): recent developments in comprehensive multispecies curation, comparative genomics and community resources. *Nucleic Acids Res.* 40:D653–D659.
  22. Bordel S, Agren R, Nielsen J. 2010. Sampling the solution space in genome-scale metabolic networks reveals transcriptional regulation in key enzymes. *PLoS Comput. Biol.* 6:e1000859. doi:[10.1371/journal.pcbi.1000859](https://doi.org/10.1371/journal.pcbi.1000859).
  23. Cvijovic M, Olivares-Hernández R, Agren R, Dahr N, Vongsangnak W, Nookae W, Patil KR, Nielsen J. 2010. BioMet toolbox: genome-wide analysis of metabolism. *Nucleic Acids Res.* 38:W144–W149. doi:[10.1093/nar/gkq404](https://doi.org/10.1093/nar/gkq404).
  24. Oliveira A, Patil K, Nielsen J. 2008. Architecture of transcriptional regulatory circuits is knitted over the topology of bio-molecular interaction networks. *BMC Syst. Biol.* 2:17. doi:[10.1186/1752-0509-2-17](https://doi.org/10.1186/1752-0509-2-17).
  25. van Helden J, André B, Collado-Vides J. 1998. Extracting regulatory sites from the upstream region of yeast genes by computational analysis of oligonucleotide frequencies. *J. Mol. Biol.* 281:827–842.
  26. Abdulrehman D, Monteiro PT, Teixeira MC, Mira NP, Lourenço AB, dos Santos SC, Cabrito TR, Francisco AP, Madeira SC, Aires RS, Oliveira AL, Sá-Correia I, Freitas AT. 2011. YEASTRACT: providing a programmatic access to curated transcriptional regulatory associations in *Saccharomyces cerevisiae* through a web services interface. *Nucleic Acids Res.* 39:D136–D140. doi:[10.1093/nar/gkq964](https://doi.org/10.1093/nar/gkq964).
  27. Schmitt AP, McEntee K. 1996. Msn2p, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 93:5777–5782.
  28. Pedersen H, Carlsen M, Nielsen J. 1999. Identification of enzymes and quantification of metabolic fluxes in the wild type and in a recombinant *Aspergillus oryzae* strain. *Appl. Environ. Microbiol.* 65:11–19.
  29. Stephanopoulos GN, Aristidou AA, Nielsen J. 1998. Metabolic engineering. Academic Press, New York, NY.
  30. Radivojac P, Vacic V, Haynes C, Cocklin RR, Mohan A, Heyen JW, Goebel MG, Iakoucheva LM. 2010. Identification, analysis, and prediction of protein ubiquitination sites. *Proteins* 78:365–380.
  31. Carlsen M, Spohr AB, Nielsen J, Villadsen J. 1996. Morphology and physiology of an  $\alpha$ -amylase producing strain of *Aspergillus oryzae* during batch cultivations. *Biotechnol. Bioeng.* 49:266–276.
  32. Santhanam A, Hartley A, Düvel K, Broach JR, Garrett S. 2004. PP2A phosphatase activity is required for stress and Tor kinase regulation of yeast stress response factor Msn2p. *Eukaryot. Cell* 3:1261–1271.



## Paper II

### Physiological characterization of the high malic acid-producing *Aspergillus oryzae* strain 2103a-68

**Knuf C, Nookaew I, Remmers I, Khoomrung S, Brown S, Berry A,  
Nielsen J.**

Appl. Microbiol. Biotechnol.:1-11 (2014)

(Epub ahead of print, doi: 10.1007/s00253-013-5465-x)



# Physiological characterization of the high malic acid-producing *Aspergillus oryzae* strain 2103a-68

Christoph Knuf · Intawat Nookaew · Ilse Remmers ·  
Sakda Khoomrung · Stephen Brown · Alan Berry ·  
Jens Nielsen

Received: 6 November 2013 / Revised: 9 December 2013 / Accepted: 10 December 2013  
© Springer-Verlag Berlin Heidelberg 2014

**Abstract** Malic acid is a C<sub>4</sub> dicarboxylic acid that is currently mainly used in the food and beverages industry as an acidulant. Because of the versatility of the group of C<sub>4</sub> dicarboxylic acids, the chemical industry has a growing interest in this chemical compound. As malic acid will be considered as a bulk chemical, microbial production requires organisms that sustain high rates, yields, and titers. *Aspergillus oryzae* is mainly known as an industrial enzyme producer, but it was also shown that it has a very competitive natural production capacity for malic acid. Recently, an engineered *A. oryzae* strain, 2103a-68, was presented which overexpressed pyruvate carboxylase, malate dehydrogenase, and a malic acid transporter. In this work, we report a detailed characterization of this strain including detailed rates and yields under malic acid production conditions. Furthermore, transcript levels of the genes of interest and corresponding enzyme activities were measured. On glucose as carbon source, 2103a-68 was able to secrete malic acid at a maximum specific production rate during stationary phase of 1.87 mmol(g dry weight (DW))<sup>-1</sup> h<sup>-1</sup> and with a yield of 1.49 mol mol<sup>-1</sup>. Intracellular fluxes were obtained using <sup>13</sup>C flux analysis during exponential

growth, supporting the success of the metabolic engineering strategy of increasing flux through the reductive cytosolic tricarboxylic acid (rTCA) branch. Additional cultivations using xylose and a glucose/xylose mixture demonstrated that *A. oryzae* is able to efficiently metabolize pentoses and hexoses to produce malic acid at high titers, rates, and yields.

**Keywords** Malate · Dicarboxylic acid · *Aspergillus oryzae* · Metabolic engineering · Xylose

## Introduction

Malic acid is a four-carbon dicarboxylic acid that was first discovered by Carl Wilhelm Scheele in 1785 from apple juice. It is abundant in fruits and vegetables and is currently produced by hydration of fumaric or malic acid which results in a mixture of the D and L forms. The present use lies in the food industry as an acidulant for beverages, candy, and other foods (Goldberg et al. 2006) and in the nonfood industry, e.g., metal cleaning, textile finishing, and pharmaceuticals, and in paints.

Concerns about global warming and increasing oil prices not only impact the transportation fuel sector but also the chemical industry which mostly relies on petroleum-based feedstocks. This is forcing a global industry to search for sustainable solutions to replace oil-based compounds by renewable building blocks. In 2004, the US Department of Energy (DoE) identified chemical compounds that can be used as building blocks for production of fuels and chemicals that can be derived from sugars via biological or chemical conversion. In this analysis, malic acid was grouped together with succinic and fumaric acid to form the group of 1,4-diacidic acids. This group was selected to be among the 12 “top value-added chemicals from biomass” that could serve as an economic driver for biorefineries in the post-oil age. As

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s00253-013-5465-x) contains supplementary material, which is available to authorized users.

C. Knuf · I. Nookaew · I. Remmers · S. Khoomrung ·  
J. Nielsen (✉)

Department of Chemical and Biological Engineering, Chalmers  
University of Technology, Kemivägen 10, 412 96 Gothenburg,  
Sweden  
e-mail: nielsenj@chalmers.se

S. Brown · A. Berry  
Novozymes AS, Davis, CA, USA

J. Nielsen  
Novo Nordisk Foundation Center for Biosustainability, Technical  
University of Denmark, 2970 Hørsholm, Denmark

malic acid is a commodity type chemical, it needs to be produced in very high yields and with high productivities.

The highest production of malic acid was found in an *Aspergillus flavus* strain, which was then further characterized, and for which, the fermentation process was optimized (Peleg et al. 1989; Battat et al. 1991). *A. flavus* is a very close relative to *Aspergillus oryzae* and the main difference is that *A. flavus* produces the mycotoxin aflatoxin. As the industrial application of malic acid was still in the food industry, the process has never been adopted for large-scale production.

Comparative genomics reveals a close relationship between *A. flavus* and *A. oryzae* as their genome sizes (*A. flavus* 36.8 and 12,197 genes; *A. oryzae* 36.7 Mb and 12,079 genes), and the number of orthologous genes is very high. Thus, each species has only about 350 unique genes. These facts led to the assumption that the two fungi might even be ecotypes of the same species and that the domestication of *A. flavus* resulted in *A. oryzae* (Payne et al. 2006). This domesticated form has been used for centuries in the Asian fermentation industries, e.g., soy sauce, sake, and vinegar.

Recent research on malic acid production in *Saccharomyces cerevisiae* involved metabolic engineering of this organism. In order to enable malic acid production, the reductive tricarboxylic acid (TCA) cycle branch (pyc and mdh) was targeted to the cytosol, and a heterologous malic acid transporter from *Schizosaccharomyces pombe* was expressed (Zelle et al. 2008).

It was recently shown that the abovementioned pathway is transcriptionally upregulated in *A. oryzae* NRRL3488 when cultivated under nitrogen starvation conditions, which leads to significant accumulation of organic acids, mainly malic acid, in the fermentation broth (Knuf et al. 2013). Furthermore, the flux through the carboxylation reaction of pyruvate to oxaloacetic acid was identified to be transcriptionally regulated, making this a promising target for metabolic engineering. As the *A. oryzae* wild-type strain NRRL3488 and *A. flavus* (ATCC13697) (Abe et al. 1962) have comparable malic acid secretion capacities, this strain is a suitable background strain for further metabolic engineering of the malic acid production pathway. A recent publication reported the successful construction of the strain 2103a-68 derived from NRRL3488. The engineered strains carry additional copies of the *A. oryzae* pyc, mdhA, and the malic acid transporter C4T318 (Brown et al. 2013). The strain was evaluated on glucose medium and has shown the highest ever reported yields and titers for malic acid production. A more thorough physiological characterization of this strain will likely provide a basis for further strain improvement strategies.

Moreover, a sustainable production of biochemicals should not depend on starch containing crops, and industrial strains should be able to metabolize other carbon sources as well, preferably lignocellulosic raw materials. As a step towards this, malic acid production from xylose, the second most

abundant sugar in biomass after glucose, was evaluated. This carbon source is difficult to metabolize by other industrially relevant organisms like *S. cerevisiae* and *Corynebacterium glutamicum* which have to be engineered and evolved in the lab in order to be able to metabolize pentoses (Kawaguchi et al. 2006; Scalcinati et al. 2012). Furthermore, *A. oryzae* has a high capacity of secreting hydrolytic enzymes to break down complex carbohydrates. This ability can be used in a consolidated bioprocess, cutting down the costs for an expensive enzymatic pretreatment of the biomass prior to fermentation.

In this study, recombinant malic acid-producing strain 2103a-68 has been further characterized by evaluating the transcript levels of the overexpressed genes and correlating the results to intracellular activity assays of key metabolic pathway enzymes. Furthermore, <sup>13</sup>C flux analysis has given new insights into the metabolism of the engineered strain. The performance of this strain was also evaluated on xylose and a glucose/xylose mixture.

## Materials and methods

### Strains

The *A. oryzae* strain NRRL3488 was used as a wild-type strain and served as a parental strain for strain 2103a-68.1 (further called triple) which overexpresses pyc, mdh, and the putative malic acid transporter C4T318 and was obtained from Novozymes, Inc. Spores of both strains were stored at −80 °C in 20 % glycerol.

### Medium

*A. oryzae* spore propagation medium (Cove-N-Gly) contained 218 g L<sup>−1</sup> sorbitol, 10 g L<sup>−1</sup> glycerol 87 %, 2.02 g L<sup>−1</sup> KNO<sub>3</sub>, 25 g L<sup>−1</sup> agar, and 50 mL L<sup>−1</sup> Cove-N-Gly salt solution. Cove-N-Gly salt solution composed of 26 g L<sup>−1</sup> KCl, 26 g L<sup>−1</sup> MgSO<sub>4</sub>·7H<sub>2</sub>O, 76 g L<sup>−1</sup> KH<sub>2</sub>PO<sub>4</sub>, and 50 mL L<sup>−1</sup> Cove-N-Gly trace elements solution. Cove-N-Gly trace elements solution composed of 40 mg L<sup>−1</sup> Na<sub>2</sub>B<sub>4</sub>O<sub>7</sub>·10H<sub>2</sub>O, 400 mg L<sup>−1</sup> CuSO<sub>4</sub>·5H<sub>2</sub>O, 1,200 mg L<sup>−1</sup> FeSO<sub>4</sub>·7H<sub>2</sub>O, 700 mg L<sup>−1</sup> MnSO<sub>4</sub>·H<sub>2</sub>O, 800 mg L<sup>−1</sup> Na<sub>2</sub>MoO<sub>4</sub>·2H<sub>2</sub>O, and 10 g L<sup>−1</sup> ZnSO<sub>4</sub>·7H<sub>2</sub>O. *A. oryzae* medium for precultures (G2-GLY) contained 18 g L<sup>−1</sup> yeast extract, 24 g L<sup>−1</sup> glycerol 87 %, 1 mL L<sup>−1</sup> pluronic PE-6100, and 100.0 g L<sup>−1</sup> CaCO<sub>3</sub>. *A. oryzae* malic acid fermentation (MAF) medium contained 1.4 g L<sup>−1</sup> (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.15 g L<sup>−1</sup> KH<sub>2</sub>PO<sub>4</sub>, 0.15 g L<sup>−1</sup> K<sub>2</sub>HPO<sub>4</sub>, 0.10 g L<sup>−1</sup> MgSO<sub>4</sub>·7H<sub>2</sub>O, 0.10 g L<sup>−1</sup> CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.005 FeSO<sub>4</sub>·7H<sub>2</sub>O, 0.005 g L<sup>−1</sup> NaCl, 100.0 g L<sup>−1</sup> CaCO<sub>3</sub>, 1 mL L<sup>−1</sup> pluronic acid (PE6100, BASF), glucose, and xylose, and a mixture of the two was applied as the carbon source; concentrations are stated in the “Results” section.

## Preparation of inoculum

Cove-N-Gly plates were inoculated with spore suspension from the  $-80^{\circ}\text{C}$  stock of each strain. The plates were incubated at  $30^{\circ}\text{C}$  for 7 days. Subsequently, the spores were harvested by addition of 10 mL of 0.01 % Tween 80 solution. The spore suspension was used to inoculate G2-GLY medium with a final concentration of  $6 \times 10^9$  spores  $\text{L}^{-1}$ . The preculture was incubated after shaking at 250 rpm in a 500-mL shake flask without baffles for 24 h at  $30^{\circ}\text{C}$ . Thereafter, the main cultures for malic acid production in fermentors or shake flasks were inoculated with 10 % of the final volume of the preculture broth.

## Batch cultivations

For evaluation of the malic acid production capabilities, batch cultivations in shake flasks and fermentors were carried out. The cultivations in bioreactors were performed in 2.7-L Applikon bioreactors (Applikon, Schiedam, The Netherlands) with 2 L of working volume. Reactors were equipped with two Rushton six-blade disc impellers, and the temperature was maintained at  $34^{\circ}\text{C}$ . The pH was buffered by the calcium carbonate in the medium. The temperature, agitation, gassing, pH, and composition of the off-gas were monitored and/or controlled using the DasGip monitoring and control system (DasGip, Jülich, Germany). The stirrer speed was set to 950 rpm and the aeration rate to 1 vvm (volume of gas per volume of fermentation broth per minute). Concentrations of oxygen and carbon dioxide in the exhaust gas were analyzed with the DasGip fedbatch-pro® gas analysis systems with the off-gas analyzer GA4 based on zirconium dioxide and a two-beam infrared sensor.

In case of shake flask cultivations for enzyme assays and RNA extraction, the strains were incubated in 100 mL MAF4AS medium in 500-mL Erlenmeyer flasks with agitation at 250 rpm in an orbital shaker. For  $^{13}\text{C}$  tracer experiments, the strains were cultivated in 250-mL shake flasks containing 50 mL MAF4AS with  $1\text{-}^{13}\text{C}$ -labelled glucose as the carbon source at a concentration of 25 g  $\text{L}^{-1}$ .

## Sampling

Known sample volumes were withdrawn from the fermentors, and one fraction was stored at  $-20^{\circ}\text{C}$  for subsequent analysis of extracellular metabolites. For quantification of cell mass, a known sample volume was treated with 2N HCl in order to solubilize undissolved  $\text{CaCO}_3$ . The treated broth was centrifuged, and the pellet was washed once. After that, the wet biomass pellet was redissolved in distilled water and poured onto a preweight aluminum dish. Aluminum dishes were kept at  $90^{\circ}\text{C}$  for 24 h in order to evaporate the water and subsequently stored in a desiccator until weighing.

## Metabolite analysis

The concentration of sugars and metabolites in the culture medium were determined by HPLC. Organic acids were measured using the Synergi 4  $\mu$  Hydro-RP80 Å HPLC column (Phenomenex Ltd., Aschaffenburg, Germany) together with the Dionex Ultimate 3000 system and a photodiode array detector (Dionex, Sunnyvale, USA) at a wavelength of 210 nm. The eluent consisted of 145 mM phosphoric acid with 10 % methanol at pH 3. Elution was performed isocratically at a flow rate of 0.7 mL/min and a temperature of  $20^{\circ}\text{C}$ .

Glucose was determined using the Aminex HPX-87H column from Bio-Rad (Bio-Rad, Sundbyberg, Sweden). The assay was performed at  $45^{\circ}\text{C}$  at a flow rate of 0.6 mL/min with a 5 mM sulfuric acid solution for an isocratic elution.

## Enzyme assay

The procedure for sample preparation and subsequent activity determination of malate dehydrogenase and pyruvate carboxylase in crude cell extract is explained in the study of Knuf et al. (2013).

## RNA extraction

Samples for RNA extraction were treated shortly with HCl in order to solubilize the  $\text{CaCO}_3$ . Subsequently, the samples were filtered through Miracloth (Calbiochem, San Diego, CA, USA), the filter cake was shock frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ . For RNA extraction, the frozen mycelium was ground to a powder in a prechilled mortar with a prechilled pestle. RNA was extracted from the mycelium powder using the RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's recommendations. In order to remove the remaining traces of genomic DNA, the samples were treated with DNase (Nolan et al. 2006; Udvardi et al. 2008). The successful removal of genomic DNA was proven by PCR using the primer pair mae3\_F and mae3\_R. The quality of the extracted RNA was assessed with the NanoDrop 2000 determining the abundance of nucleic acid (260 nm), proteins (280 nm), and other contaminations (230 nm). The 28S, 18S, and 5S ribosomal RNA was visualized by gel electrophoresis.

## Transcriptional analysis

cDNA was synthesized from purified RNA samples using random hexamer primers and the RevertAid™ H Minus First Strand cDNA Synthesis Kit (Fermentas) according to the manufacturer's protocol. Quantitative real-time PCR was performed in 20  $\mu\text{L}$  reaction volume using Brilliant III Ultra-Fast SYBR® Green qPCR mix and the corresponding primers

(Table 1) according to the manufacturer's protocol. Amplification and product detection were carried out using the Stratagene Mx3005P qPCR System (Agilent technologies) with the following temperature profile: initial incubation at 95 °C, 40 cycles of 95 °C for 15 s and 60 °C for 20 s, and 1 cycle for detection of the dissociation curve consisting of 95 °C for 1 min, 55 °C for 30 s, and 95 °C for 30 s. The MxPro software (Agilent Technologies) determined the threshold, baseline, and threshold cycle ( $C_t$ ) values. The fold changes were determined using the  $2^{-\Delta\Delta C_t}$  method (Livak and Schmittgen 2001) with  $C_t$  values from three biological replicates.

### $^{13}\text{C}$ flux analysis

Samples for the calculation of fractional labeling were taken from the mid-exponential phase of shake flask cultures with initial glucose concentrations of 25 g L<sup>-1</sup> of D-glucose-1-<sup>13</sup>C (<sup>13</sup>C of >99 %; Isotec, Sigma-Aldrich).

Samples for the determination of proteinogenic amino acids were shortly treated with 2 M HCl in order to dissolve the calcium carbonate and subsequently washed by centrifugation at 12,000 g for 5 min and resuspension of the pellet in 0.5 M HCl. After the second centrifugation step, the pellet was resuspended in 6 N HCl. The further procedure for analysis can be found in the study of Papini et al. (2012). The metabolic model of the central carbon metabolism that was used as a scaffold to estimate an intracellular flux was adopted from a compartmentalized *Aspergillus niger* flux model (Meijer et al. 2009), and the biomass fluxes were taken from a study on *Aspergillus nidulans* (Panagiotou et al. 2008). The labeling patterns of proteinogenic amino acids were corrected for their natural labeling abundance (van Winden et al. 2002; Wittmann 2007) and further used for calculation of summed fractional labeling (SFL) as proposed by Gombert et al. (2001). The calculations were performed through in-house MATLAB software available in the BioMet toolbox

(Cvijovic et al. 2010) in which the calculations followed the examples of Grotkjaer et al. (Grotkjaer et al. 2005) and Wiechert (2001).

## Results

The strain 2103a-68.1 is a result of the transformation of the wild-type *A. oryzae* strain NRRL3488, which is a known malic acid producer (Bercovitz et al. 1990), with DNA fragments harboring genes for *pyc*, *mdh3*, and the malate transporter (C4T318) and a subsequent screening process. The genes are under control of the strong constitutive phosphoglycerate kinase (*pgk*) promoter and the *glaA* terminator, ensuring a strong expression of the desired cytosolic reductive TCA branch towards malic acid. The screening process selected transformants with the highest malic acid production. Although the presence of the three transformation cassettes was verified by PCR, the exact number of copies was not determined (Brown et al. 2013).

In order to evaluate the success of the metabolic engineering strategy towards establishing an industrially applicable malic acid producer, the transcript levels of the overexpressed genes were verified by quantitative PCR (qPCR). As the activity of the malate transporter is difficult to determine by enzyme assays, qPCR was used in order to check the transcript levels of the transporter gene in the production phase (48 h) in comparison to the *pyc* and *mdh* expression. TAF10 is a frequently used reference gene in *S. cerevisiae*, and the homolog in *A. oryzae* was used as a reference gene for this study. Optimal primers for the four genes were calculated using the Primer3Plus (Untergasser et al. 2012) web interface. The expression level of *pyc* and *mdh* in the stationary phase increased about 3.6 and 9.6 times, respectively, whereas the transcript level of the malate exporter was increased about seven times compared to the wild type (exact data and a column chart can be found in Online resource 2).

**Table 1** Primer used for the transcriptional analysis

Primer	Sequence, 5'-3'	Systematic name	Gene	Ortholog, non-AspGD organism	Reaction efficiency (%)
pyc_F	AGCATGTCCGGTATGACCTC	AO090023000801	<b><i>pyc</i></b>	<i>S. cerevisiae</i> (PYC2)	<b>95.45</b>
pyc_R	CAACTGACCACCAGGGATCT				
mae3_F	GGTCTCTTCTTCCCGACCTT	AO090023000318	<b><i>mae3</i></b>	<i>S. pombe</i> (mae1)	<b>107.4</b>
mae3_R	CATGGTTGAAGGCCGTACT				
mdh_3_1F	AGTCCCCTACCCTTCGCATT	AO090701000013	<b><i>mdh3</i></b>		<b>98.7</b>
mdh_3_1R	CGGGGTTTGAGATAACCAGA				
TAF10_F	CAGCTCTTCCGAAACCTGAC	AO090026000688	<b><i>taf10</i></b>	<i>S. cerevisiae</i> (TAF10)	<b>108.41</b>
TAF10_R	CTCGTCTCCTTCTCGATGG				

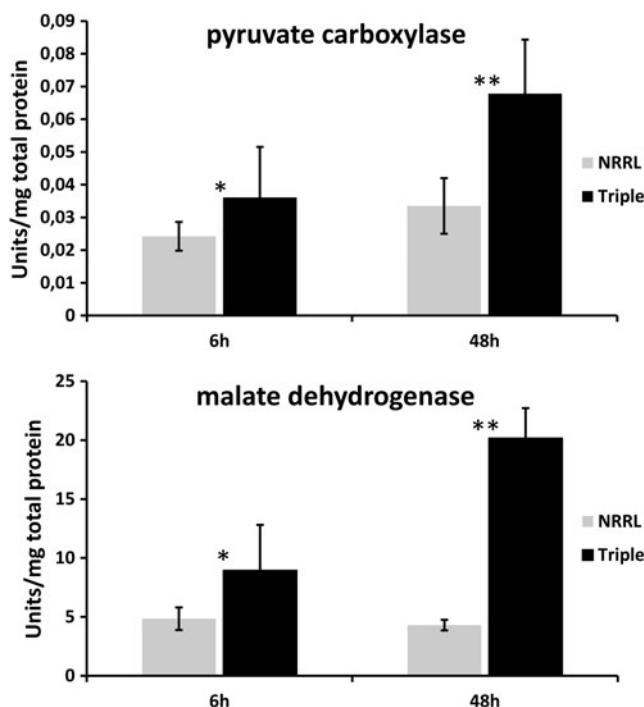
Gene names in bold are not yet annotated

*pyc* 1.87E-07; *mae3* 4.98E-07; *mdh3* 2.37E-08; *taf10* 2.03E-08

## Enzyme assay

In order to verify the enhancement of enzyme activity of pyc and mdh3 for increasing flux through the reductive TCA cycle, *in vitro* enzyme assays were performed as described above. Samples for enzymatic assays were taken in mid-exponential phase at 6 h and in the stationary phase at 48 h. In the sample of the exponential phase, the difference in enzyme activity between the triple transformant strain 2103a-68 and the wild type was not significant (*p* values of 0.299 and 0.220 for pyc and mdh, respectively), whereas the activity difference increased drastically in the stationary phase (Fig. 1). The pyc activity for 2103a-68 is about twice as high as the wild type and more than four times higher in the case of malate dehydrogenase (*p* values of 0.027 and 0.008 for pyc and mdh, respectively). Comparing the activities of the two enzymes, it becomes obvious that the mdh activity is roughly 2 orders of magnitude higher than the activity of pyruvate carboxylase, which is consistent with the expression analysis.

Assuming the activity of the pyruvate carboxylase to be the flux-controlling step in the pathway from pyruvate to malic acid, the flux through the rTCA was theoretically calculated. Using the activity at 48 h and the protein content of *A. oryzae* grown at a dilution rate of  $0.17 \text{ h}^{-1}$  of 42 % (Pedersen et al. 1999), the average flux through the pathway is 0.842 and



**Fig. 1** Enzyme activities of pyc (upper part) and mdh (lower part) for NRRL3488 (NRRL) and the triple transformant 2103a-68.1 (Triple) in exponential growth phase (6 h) and starvation phase (48 h). One unit describes the conversion of 1  $\mu\text{M}$  substrate per min. \**p*>0.05; \*\**p*<0.05, Student's *t* test

1.709  $\text{mmol g dry weight (DW)}^{-1} \text{h}^{-1}$  for the wild-type and 2103a-68.1 strain, respectively.

## Fermentation characteristics

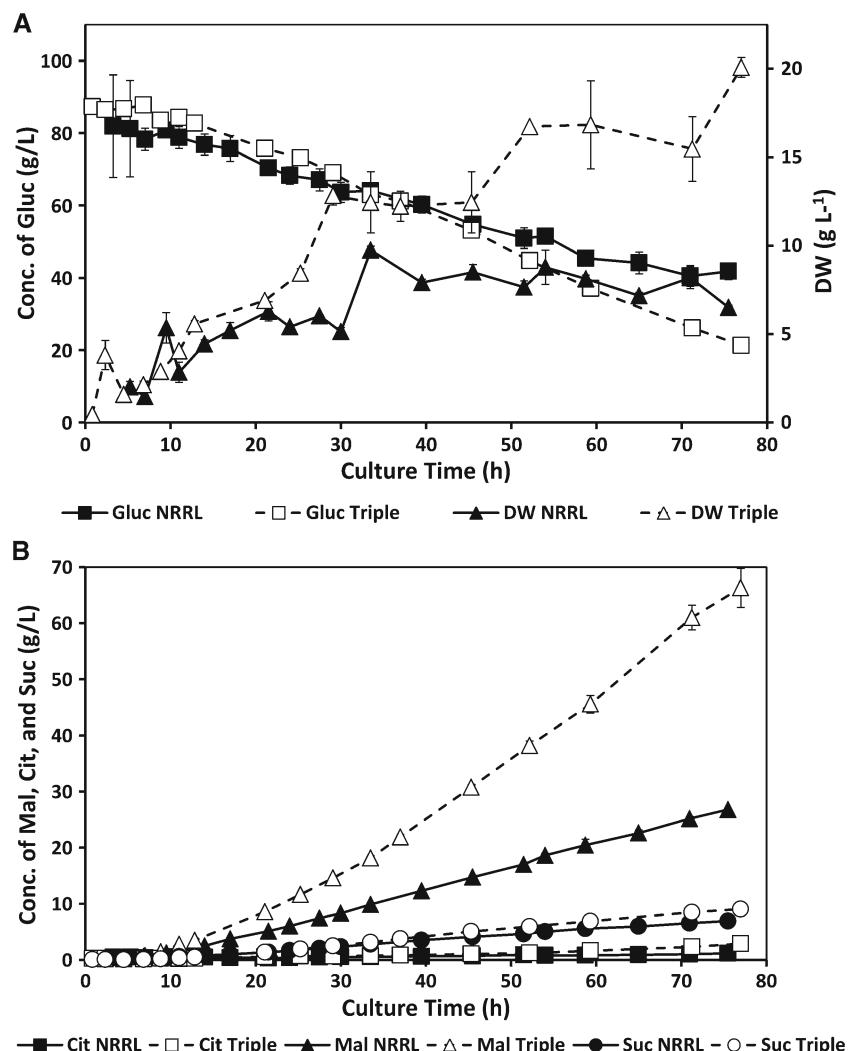
For a thorough determination of the malic acid production capabilities of the strain 2103a-68.1 compared to its parent strain NRRL3488, both were cultivated in MAF medium containing 100 g of glucose as the sole carbon source. The fermentation profiles of both strains are shown in Fig. 2. After an initial exponential growth phase, which is then limited by the availability of nitrogen, the NRRL3488 strain reached a constant biomass concentration of about  $7.5 \text{ g L}^{-1}$ . In the 2103a-68.1 case, the biomass increased during the production phase from around 12 to about  $20 \text{ g L}^{-1}$ . The final malic acid concentrations were  $26.77 \pm 0.197$  and  $66.3 \pm 2.36$  for the wild-type and improved strain, respectively.

As the industrial production process aims at a prolonged starvation phase in which the carbon can directly be converted to malic acid without the need of sacrificing it for cell growth, the improvement of production rates in this phase is crucial, and specific rates and yields were calculated for this phase. As previously reported, the biomass concentrations exceed the maximum theoretical biomass concentrations that could be reached with the amount of ammonium supplied (Knuf et al. 2013). Therefore, the specific rates were calculated taking a biomass concentration of  $4.19 \text{ g L}^{-1}$  into account. This value was calculated using the analytically determined elemental biomass composition of *A. niger* during exponential growth and corresponds well with biomass values of *A. oryzae* after ammonium was consumed from the medium. The rates, specific and volumetric, and yields, molar and C-mole basis, can be found in Table 2. The specific malic acid production rate increased about three times from  $0.61 \text{ mmol(g DW)}^{-1} \text{h}^{-1}$  in the parent strain to  $1.87 \text{ mmol(g DW)}^{-1} \text{h}^{-1}$  in the engineered strain, which correlated very well with the calculated fluxes from the enzyme assay data. The second largest flux was the succinate production rate with about  $0.28 \text{ mmol(g DW)}^{-1} \text{h}^{-1}$  in the 2103a-68.1 strain, a 1.8 times increase over the wild type. The citrate production in the engineered strain increased three times to reach  $0.06 \text{ mmol(g DW)}^{-1} \text{h}^{-1}$ . Along with the increased production rates of organic acids, the uptake rate for glucose increased by almost 70 %. This leads to a malic acid yield on a glucose basis of  $1.49 \text{ mol mol}^{-1}$  which is an improvement of 80 % over the wild type in the stationary phase.

## Metabolic flux analysis

In order to get a better understanding of the intracellular fluxes,  $^{13}\text{C}$  flux analysis was performed during the exponential growth phase. NRRL3488 and 2103a-68 were cultivated in shake flasks with MAF medium containing  $25 \text{ g L}^{-1} 1\text{-}^{13}\text{C}$ -

**Fig. 2** Dry weight and carbon source concentrations (a) and extracellular metabolite concentrations (b) during bioreactor cultivations of NRRL3488 and 2103a-68 in MAF medium in 2.7-L Applikon bioreactors with a working volume of 2 L, controlled by a DasGip control system. The results shown are averages and standard errors of three reactors. *Mal* malate, *Cit* citrate, *Suc* succinate, *Gluc* glucose, *DW* dry weight; 2103a-68 (*Triple*), NRRL3488 (*NRRL*)



labelled glucose. The biomass was harvested after 7.5 h of cultivation which is in the mid-exponential phase. The SFLs were calculated from the masses of hydrolyzed biomass samples and used as constraints for the flux calculation that was performed using the <sup>13</sup>C program from the BioMet Toolbox (Cvijovic et al. 2010). Considering Fig. 3, the main deviations in fluxes can be found in the split between the pentose phosphate pathway and glycolysis. The higher flux enters the pentose phosphate pathway in the engineered strain. After the carbon was fed back into the glycolysis, the fluxes are evened out until the pyruvate branch point was reached, from where the flux enters the mitochondrion to be oxidized in the TCA cycle. Another fraction stays in the cytosol and continues down the reductive TCA branch. In the engineered strain, about twice as much carbon is pulled down the reductive TCA branch compared to the wild type. About three quarters of the flux is then diverted out of the cell in the form of malate. The model also allowed further conversion of malate to succinate and subsequent excretion of this metabolite. A small fraction of the malate was shifted into the

mitochondrion in order to fuel the TCA cycle and export citrate for subsequent acetyl-CoA generation in the cytosol.

#### Xylose as an alternative carbon source

After obtaining the positive results demonstrating the production of malic acid from the standard industrial carbon source, glucose, by the engineered strain 2103a-68, the production from the alternative carbon source xylose was evaluated. The engineered strain was cultivated under the same conditions than before, but using 100 g L<sup>-1</sup> initial xylose as the sole carbon source. The average biomass concentration during the stationary phase was about 5 g L<sup>-1</sup> (Fig. 4). As mentioned above, the specific rates were calculated by assuming a biomass of 4.19 g L<sup>-1</sup>, and these can be found in Table 2.

The production rates for malic acid and citrate on xylose do not reach the same high values of the 2103a-68.1 strain on glucose but are still higher than the wild type grown on glucose. The specific xylose uptake rate stayed the same at around 1.27 mmol(g DW)<sup>-1</sup> h<sup>-1</sup>. The molar yield of malate on the

**Table 2** Physiological data for NRRL3488 and 2103a-68 (triple) grown in malic acid fermentation medium with varying carbon sources during acid production phase

Strain	Carbon source	Specific rates [mmol(g DW) <sup>-1</sup> h <sup>-1</sup> ]				Volumetric rates (g L <sup>-1</sup> h <sup>-1</sup> )				Yields (mol mol <sup>-1</sup> C source)				Yields (mmol C mol <sup>-1</sup> )	
		$r_{PMal}$	$r_{PSuc}$	$r_S$	$r_{PCit}$	$r_{PMal}$	$r_{PSuc}$	$r_S$	$r_{PCit}$	$Y_{SMal}$	$Y_{SCit}$	$Y_{SSuc}$	$Y_{SMal}$	$Y_{SCit}$	$Y_{SSuc}$
NRRL3488	Glucose	0.61±0.01	0.02±0	0.16±0.03	0.73±0.06	0.34±0.06	0.01±0	0.08±0.01	0.55±0.05	0.83±0.07	0.02±0	0.22±0.02	138.33±10.74	3.33±0.53	36.67±3.22
Triple	Glucose	1.87±0.23	0.06±0.01	0.28±0.04	1.26±0.11	1.05±0.13	0.05±0.01	0.14±0.02	0.95±0.08	1.49±0.05	0.05±0	0.22±0.01	247.67±8.27	8±0.79	36.49±1.74
Triple	Xylose	0.89±0.14	0.05±0.01	0.16±0.02	1.27±0.09	0.5±0.08	0.04±0.01	0.08±0.01	0.8±0.06	0.7±0.06	0.04±0.01	0.12±0.01	140.83±12.54	7.57±1.53	24.75±2.04
Triple	Glucose phase (16–35 h)	1.46±0.16	0.05±0.01	0.14±0.02	1.23±0.05	0.82±0.09	0.04±0.01	0.07±0.01	0.93±0.04	1.18±0.08	0.04±0	0.12±0.01	196.74±13.27	6.76±0.66	19.32±1.47
Triple	Xylose phase (43–66 h)	1.08±0.19	0.03±0.01	0.02±0.01	1.26±0.06	0.61±0.11	0.03±0.01	0.01±0.01	0.79±0.04	0.86±0.11	0.03±0	0.01±0.01	171.97±4.59	5.25±0.19	2.9±0.41

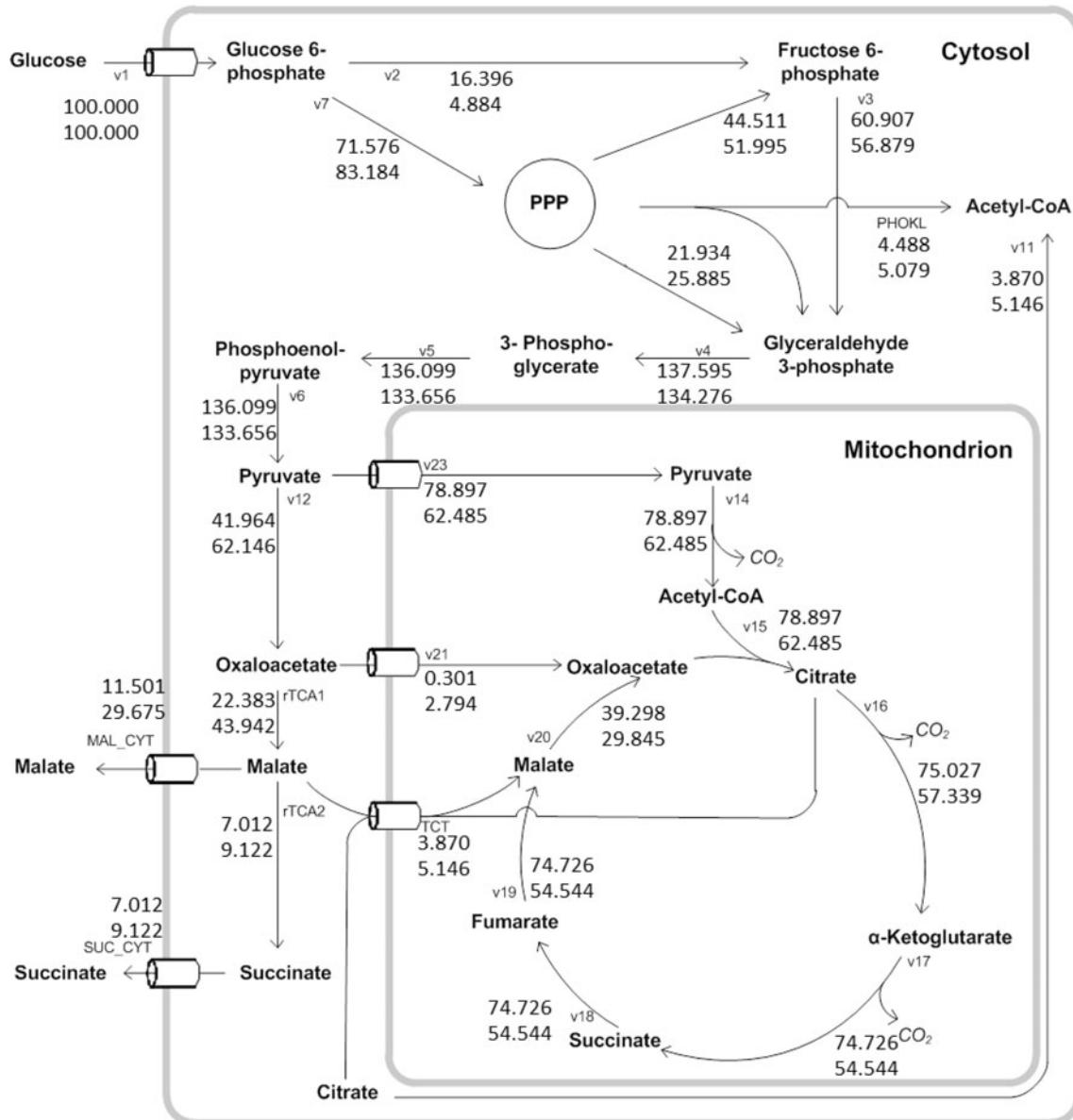
The numbers stated are means of four individual bioreactors ± standard errors  
 $r_{PMal}$  malic acid production rate,  $r_{PSuc}$  malic acid production rate,  $r_S$  substrate consumption rate,  $r_{PCit}$  yield of malate per substrate,  $Y_{SMal}$  yield of citrate per substrate,  $Y_{SCit}$  yield of succinate per substrate

carbon source dropped below the value of the wild type on glucose but was almost equal to the wild type on a C-mole basis.

As a first step towards characterizing strain performance on lignocellulosic feedstocks, a mix of carbon sources was evaluated as well. Batch fermentations were performed with initial concentrations of xylose and glucose of 30 g L<sup>-1</sup> each (Fig. 5). The biomass concentration during this cultivation stayed constant after about 16 h at 5.5 g L<sup>-1</sup>. In the initial phase of the cultivation, during the first 40 h, the major supply of carbon was from glucose, with marginal xylose consumption, until there was complete exhaustion of the hexose. Thereafter, xylose consumption continued until it was almost exhausted at the end of the cultivation at 70 h. The rates and yields for this cultivation were calculated in the stationary phase during the glucose consumption phase (16–35 h) and xylose consumption phase (43–66 h). The two phases reflect the same trend obtained with the single carbon source cultivations, whereas the results are not as divergent as before. Comparing the glucose consumption phase to the single carbon cultivation, the malic acid production rate dropped slightly to 1.46 mmol(g DW)<sup>-1</sup> h<sup>-1</sup>, whereas the carbon uptake rate almost stayed equal; this leads to a decrease in molar yield of 20 %. The malic acid production rate during the xylose phase was elevated compared to the cultivation with xylose as the single carbon source to reach about 1.08 mmol(g DW)<sup>-1</sup> h<sup>-1</sup>. As the carbon uptake rate was the same as in the single carbon source cultivation, the yield increased. The malic acid yields on a C-mole basis of the glucose and xylose phase are almost comparable with 196 and 171 mmol C-mol<sup>-1</sup>, respectively.

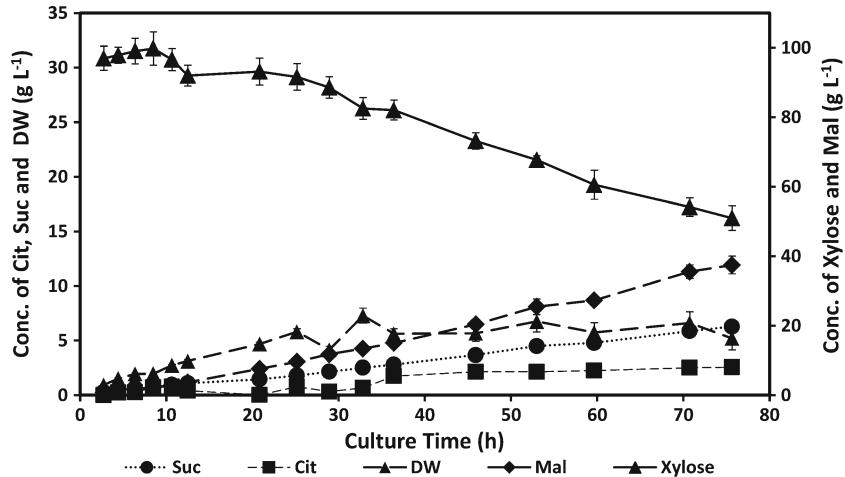
## Discussion

This detailed characterization of the previously reported strain 2103a-68 that has been engineered to overexpress the reductive TCA branch in the cytosol, together with evaluation of its performance on xylose, opens the possibility of future production of biochemicals from renewable feedstocks. It also potentially extends the utility of *A. oryzae* from the industrial production of enzymes to include organic acids. In a previous publication (Knuf et al. 2013), the transcriptional response of the parent strain NRRL3488 towards nitrogen depletion was shown, and an impressive malic acid production potential was reported for the wild-type strain. In the strain 2103a-68, the malic acid production rate was tripled by overexpression of three genes. The ectopic integration of transformed DNA molecules might lead to multiple integrations into the genome, which results in a variation of expression among transformants. In this case, the strain 2103a-68 was selected after a screening process as the highest malic producer. In order to evaluate the transcription ratios among the three genes that are responsible for the increase in malic acid production, the mRNA levels of pyruvate carboxylase, malate dehydrogenase, and the malic acid transporter were quantified by

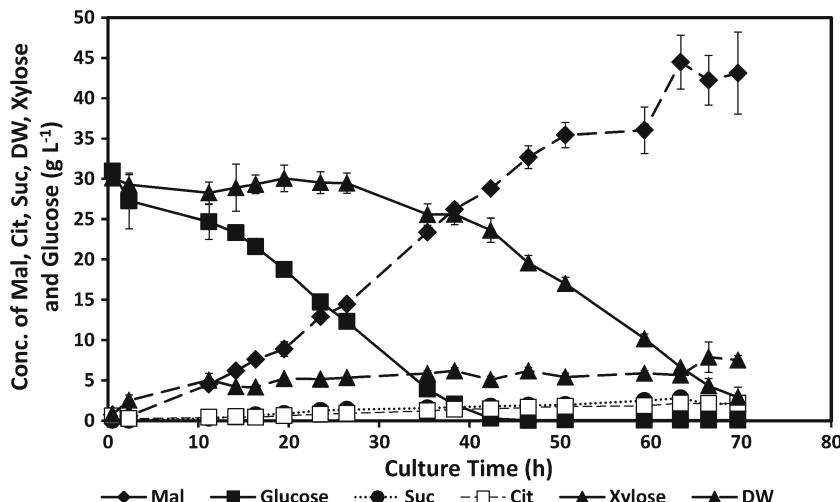


**Fig. 3** Intracellular flux distribution of NRRL3488 (*upper values*) and 2103a-68 (*lower values*). Samples for calculation were taken after 7.5 h (mid-exponential phase) of cultivation on MAF medium containing 25 g L<sup>-1</sup> glucose

**Fig. 4** Dry weight, carbon source concentrations, and extracellular metabolite concentrations during bioreactor cultivations of 2103a-68.1 in MAF medium in 2.7-L Applikon bioreactors with a working volume of 2 L, controlled by a DasGip control system. The results shown are averages and standard errors of three reactors. *Mal* malate, *Cit* citrate, *Suc* succinate, *Xyl* xylose, *DW* dry weight



**Fig. 5** Dry weight, carbon source concentrations, and extracellular metabolite concentrations during bioreactor cultivations of 2103a-68.1 in MAF medium in 2.7-L Applikon bioreactors with a working volume of 2 L, controlled by a DasGip control system. The results shown are averages and standard errors of three reactors. *Mal* malate, *Cit* citrate, *Suc* succinate, *Xyl* xylose, *DW* dry weight



qPCR. The lowest increase in transcript level over the wild type was seen for the *pyc* (AO090023000801) gene and might result from a single integration and the control of the strong constitutive promoter of *pgk* (Sakai et al. 2012). As all three fragments are under the same control, multiples of 3.6 can be assumed to be due to multiple integrations; therefore, the transporter fragment might be integrated twice (sevenfold change) and the *mdhA* (AO090701000013) fragment three times (ninefold change). The increased transcript levels are also reflected in the enzyme activities, which indicate that the enzyme activity is mainly transcriptionally regulated, supporting previous results integrating transcriptional and theoretical flux calculations with the identification of the *pyc* reaction as a metabolic engineering target, as increased fluxes were correlated with an increased transcription (Knuf et al. 2013). Thus, further overexpression of *pyc* may result in an additional increase in malic acid production.

Besides malate, some major by-products were produced, mainly succinic acid which was produced to final titers of up to  $9.03 \pm 0.49 \text{ g L}^{-1}$  in the 2103a-68 strain on glucose after 77 h of cultivation. The specific production rate for succinic acid was almost doubled (NRRL 0.16±0.03; 2103a-68 0.28±0.04) in the 2103a-68 strain. This suggests a direct correlation between the increase of the cytosolic malic acid pool and the secretion of succinic acid. In the current genome-scale metabolic model of *A. oryzae* iWV1314 (Vongsangnak et al. 2008), no reaction can be found that allows for a continuation of the rTCA towards succinic acid in the cytosol. The rTCA cycle from pyruvate via oxaloacetic acid to malic acid has been verified for *A. flavus* (Peleg et al. 1989) by  $^{13}\text{C}$ -NMR and isoenzyme analysis. The verification of the continuation of a reductive TCA branch from malic acid via fumaric acid to succinic acid has yet to be verified. The current *A. oryzae* genome-scale metabolic model lists the fumarase/fumarate hydratase reaction (reaction number 100, AO090120000133) as being localized to the mitochondrion. For the *S. cerevisiae*

fumarase, a dual localization has been found which is dependent on the metabolites in the compartments (Regev-Rudzki et al. 2009). In the *A. niger* model iMA871, the fumarate hydratase reaction (fumR, An12g07850) is localized in the mitochondrion as well as in the cytosol. Assuming that a continuation of the rTCA also exists in the cytosol of *A. oryzae*, the production of succinic acid indicates that the pool of intracellular malate increased and the flux continues further down to succinate.

Another by-product during malic acid fermentations is citrate with final titers of  $2.89 \pm 0.51 \text{ g L}^{-1}$  in the engineered strain, which is a specific production rate that is threefold increased over the wild type. This is another indication that an increased malic acid pool may trigger the tricarboxylate transporter which exchanges cytosolic malate for mitochondrial citrate, thereby replenishing the carbon in the TCA cycle for further citric acid production (Brown et al. 2013; de Jongh and Nielsen 2008; Karaffa and Kubicek 2003). Both side products indicate an accumulation of malate in the cytosol. This problem could be circumvented by increasing the export capacity of malate, especially lowering the  $K_m$  value, so that efficient transport occurs at low cytosolic concentrations of malate. An interesting transporter could be the extensively studied *S. pombe* malic acid transporter (SpMAE), for which an import flux of  $8.7 \text{ mmol g}^{-1} \text{ min}^{-1}$  was reported (Camarasa et al. 2001). As transporters can be promiscuous, transporters known for their large succinic acid transport capacities like the one from *Baetis succiniciproducens* which sustains fluxes of up to  $8.7 \pm 0.05 \text{ mmol g DW}^{-1} \text{ h}^{-1}$  (Becker et al. 2013) might serve this purpose. Another strategy for eliminating the citrate production would be the deletion of the abovementioned TCT gene, thereby blocking the drain of malate from the cytosol to the mitochondrion. The tricarboxylate transporter is a very interesting metabolic engineering target in connection with malic acid production. According to the theory for citric acid production in *A. niger* (Karaffa and Kubicek 2003), for each mole of citric

acid produced, one mole of malate was produced in the cytosol and then incorporated into the citrate molecule. This suggests very high fluxes through the rTCA in *A. niger*, and a block of the import of malate into the mitochondrion in collaboration with an efficient malic acid exporter might turn *A. niger* into a malic acid producer. A protein with possible malic acid transporter activity is An16g08330, which shows 87 % sequence similarity to the C4T318, and it therefore might have malic acid transporter activity, which is not enough to compete with the efficient tricarboxylate transporter.

This is only the second study using  $^{13}\text{C}$  labeling analysis and a subsequent balancing of metabolites to determine intracellular fluxes in the filamentous fungus *A. oryzae*. The first study of  $^{13}\text{C}$  flux analysis dealt with the elucidation of the relation of the pathway activity of glycolysis and PP pathway during cultivation on ammonium and nitrate as nitrogen sources (Schmidt et al. 1998). The drawback of that model is that it is not compartmentalized. As the comparison of the cytosolic and mitochondrial fluxes towards malic acid was the aim of this study, a compartmentalized model for *A. oryzae* was built, using an already compartmentalized network from a study on *A. niger* (Meijer et al. 2009) as a template. This network was extended by the reductive TCA branch down to succinate and the tricarboxylate transporter for exchange of citrate and malate between the cytosol and the mitochondrion. This improved model for the central carbon metabolism of *A. oryzae* including the carbon transitions and valid biomass fluxes can be found in Online resource 1 and contains all the necessary information for the use with the  $^{13}\text{C}$  function of the BioMet Toolbox, which is accessible at [www.biomet-toolbox.org](http://www.biomet-toolbox.org).

Using this compartmentalized network, intracellular fluxes in the wild type and 2103a-68 were calculated. Determining the intracellular fluxes during the actual production phase for malic acid was not possible due to technical limitations. GC-MS-based flux analysis measures the labeling enrichment in the amino acids, which allows deduction of labeling patterns of the amino acid precursors from the central carbon metabolism, in order to get a precise estimate of intracellular fluxes. As this production phase is characterized by nitrogen starvation, the de novo synthesis of amino acids is hampered; hence, the derived labeling patterns would not reflect the actual metabolic state of the cell. Nevertheless, the fluxes presented here, which were obtained during exponential growth, prove that the overexpression of the reductive TCA branch works to improve the malic acid production rates.

Batch cultivation is the so far described cultivation setup. One reason for that is that wild-type strains have been used, and natural regulation mechanisms as a result of nitrogen starvation were needed for the activation of the cytosolic pathway towards malic acid (Knuf et al. 2013). The results presented here clearly show an increased activity of the cytosolic pathway towards malic acid during exponential growth. One application for this trait would be the operation of a

continuous process using, e.g., sugar-containing waste streams from paper mills for malic acid production.

The cultivations of NRRL3488 and 2103a-68 on glucose revealed that the specific glucose uptake rate in the nitrogen starvation phase was almost doubled by means of genetic engineering around the rTCA pathway. This indicates that the pull of pyruvate into the rTCA has an influence on the carbon flux and the uptake of carbon source which might be facilitated by relieving allosteric inhibition.

The uptake of carbon sources is a major overall rate-controlling step in the production of high-volume chemical compounds that not only require high yields but also high conversion rates. Therefore, it was significant that the glucose uptake rate of the engineered strain was increased almost twofold over the wild-type strain. Even more interesting is the finding of constant specific glucose and xylose uptake rates in all cultivations of the engineered strain. Compared with *S. cerevisiae* (Shen et al. 2012), *C. glutamicum* (Buschke et al. 2011), or *Escherichia coli* (Liu et al. 2012), *A. oryzae* does not have to be extensively engineered in order to be able to metabolize sugars other than glucose, especially xylose, efficiently. *A. oryzae*'s natural uptake capacity of a variety of carbon sources allows for efficient conversions of biomass to the desired product and makes it a very interesting platform organism for the production of biochemicals.

Furthermore, the fact that this strain continues to metabolize carbon when the nitrogen source is depleted offers the possibility of production of biochemicals uncoupled from growth. This is an important factor for obtaining high-yielding production processes, as no carbon is sacrificed for the production of biomass, which inevitably reduces the yield of the desired product. In the case of *A. oryzae*, only a small fraction of the carbon needs to be shunted to the TCA cycle in order to produce ATP for cellular maintenance.

In conclusion, we demonstrate that the reconstruction of a reductive TCA cycle pathway in the cytosol allows for high-level production of malic acid not only from glucose but also xylose and glucose/xylose mixtures, and this paves the way for establishing a biorefinery-based process for the production of this important dicarboxylic acid.

**Acknowledgments** We acknowledge funding for our research activities from the Knut and Alice Wallenberg Foundation, the European Research Council (grant no. 247013), Novozymes, and the Novo Nordisk Foundation. Scientific discussions with Eugenio Meza and Sergio Bordel were much appreciated.

## References

- Abe S, Furuya A, Saito T, Takayama K (1962) Method of producing L-malic acid by fermentation. US Patent 3,063,910
- Battat E, Peleg Y, Bercovitz A, Rokem JS, Goldberg I (1991) Optimization of L-malic acid production by *Aspergillus flavus* in a

- stirred fermentor. *Biotechnol Bioeng* 37(11):1108–1116. doi:[10.1002/bit.260371117](https://doi.org/10.1002/bit.260371117)
- Becker J, Reinefeld J, Stellmacher R, Schäfer R, Lange A, Meyer H, Lalk M, Zelder O, von Abendroth G, Schröder H, Haefner S, Wittmann C (2013) Systems-wide analysis and engineering of metabolic pathway fluxes in bio-succinate producing *Basfia succiniciproducens*. *Biotechnol Bioeng* 110(11):3013–3023. doi:[10.1002/bit.24963](https://doi.org/10.1002/bit.24963)
- Bercovitz A, Peleg Y, Battat E, Rokem JS, Goldberg I (1990) Localization of pyruvate carboxylase in organic acid-producing *Aspergillus* strains. *Appl Environ Microbiol* 56(6):1594–1597
- Brown SH, Bashkirova L, Berka R, Chandler T, Doty T, McCall K, McCulloch M, McFarland S, Thompson S, Yaver D, Berry A (2013) Metabolic engineering of *Aspergillus oryzae* NRRL 3488 for increased production of L-malic acid. *Appl Microbiol Biotechnol* 97(20):8903–8912. doi:[10.1007/s00253-013-5132-2](https://doi.org/10.1007/s00253-013-5132-2)
- Buschke N, Schröder H, Wittmann C (2011) Metabolic engineering of *Corynebacterium glutamicum* for production of 1,5-diaminopentane from hemicellulose. *Biotechnol J* 6(3):306–317. doi:[10.1002/biot.201000304](https://doi.org/10.1002/biot.201000304)
- Camarasa C, Bidard F, Bony M, Barre P, Dequin S (2001) Characterization of *Schizosaccharomyces pombe* malate permease by expression in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 67(9):4144–4151. doi:[10.1128/aem.67.9.4144-4151.2001](https://doi.org/10.1128/aem.67.9.4144-4151.2001)
- Cvijovic M, Olivares-Hernández R, Agren R, Dahr N, Vongsangnak W, Nookaei I, Patil KR, Nielsen J (2010) BioMet Toolbox: genome-wide analysis of metabolism. *Nucleic Acids Res* 38(Suppl 2): W144–W149. doi:[10.1093/nar/gkq404](https://doi.org/10.1093/nar/gkq404)
- de Jongh WA, Nielsen J (2008) Enhanced citrate production through gene insertion in *Aspergillus niger*. *Metab Eng* 10(2):87–96. doi:[10.1016/j.ymben.2007.11.002](https://doi.org/10.1016/j.ymben.2007.11.002)
- Goldberg I, Rokem JS, Pines O (2006) Organic acids: old metabolites, new themes. *J Chem Technol Biotechnol* 81(10):1601–1611. doi:[10.1002/jctb.1590](https://doi.org/10.1002/jctb.1590)
- Gombert AK, Moreira dos Santos M, Christensen B, Nielsen J (2001) Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J Bacteriol* 183(4):1441–1451. doi:[10.1128/jb.183.4.1441-1451.2001](https://doi.org/10.1128/jb.183.4.1441-1451.2001)
- Grotkjaer T, Christakopoulos P, Nielsen J, Olsson L (2005) Comparative metabolic network analysis of two xylose fermenting recombinant *Saccharomyces cerevisiae* strains. *Metab Eng* 7(5–6):437–444. doi:[10.1016/j.ymen.2005.07.003](https://doi.org/10.1016/j.ymen.2005.07.003)
- Karaffa L, Kubicek C (2003) *Aspergillus niger* citric acid accumulation: do we understand this well working black box? *Appl Microbiol Biotechnol* 61(3):189–196. doi:[10.1007/s00253-002-1201-7](https://doi.org/10.1007/s00253-002-1201-7)
- Kawaguchi H, Vertès AA, Okino S, Inui M, Yukawa H (2006) Engineering of a xylose metabolic pathway in *Corynebacterium glutamicum*. *Appl Environ Microbiol* 72(5):3418–3428. doi:[10.1128/aem.72.5.3418-3428.2006](https://doi.org/10.1128/aem.72.5.3418-3428.2006)
- Knuf C, Nookaei I, Brown SH, McCulloch M, Berry A, Nielsen J (2013) Investigation of malic acid production in *Aspergillus oryzae* under nitrogen starvation conditions. *Appl Environ Microbiol* 79(19): 6050–6058. doi:[10.1128/aem.01445-13](https://doi.org/10.1128/aem.01445-13)
- Liu R, Liang L, Chen K, Ma J, Jiang M, Wei P, Ouyang P (2012) Fermentation of xylose to succinate by enhancement of ATP supply in metabolically engineered *Escherichia coli*. *Appl Microbiol Biotechnol* 94(4):959–968. doi:[10.1007/s00253-012-3896-4](https://doi.org/10.1007/s00253-012-3896-4)
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods* 25(4):402–408. doi:[10.1006/meth.2001.1262](https://doi.org/10.1006/meth.2001.1262)
- Meijer S, Otero J, Olivares R, Andersen MR, Olsson L, Nielsen J (2009) Overexpression of isocitrate lyase—glyoxylate bypass influence on metabolism in *Aspergillus niger*. *Metab Eng* 11(2):107–116. doi:[10.1016/j.ymen.2008.12.002](https://doi.org/10.1016/j.ymen.2008.12.002)
- Nolan T, Hands RE, Bustin SA (2006) Quantification of mRNA using real-time RT-PCR. *Nat Protoc* 1(3):1559–1582
- Panagiotou G, Andersen MR, Grotkjaer T, Regueira TB, Hofmann G, Nielsen J, Olsson L (2008) Systems analysis unfolds the relationship between the phosphoketolase pathway and growth in *Aspergillus nidulans*. *PLoS One* 3(12):e3847. doi:[10.1371/journal.pone.0003847](https://doi.org/10.1371/journal.pone.0003847)
- Papini M, Nookaei I, Siewers V, Nielsen J (2012) Physiological characterization of recombinant *Saccharomyces cerevisiae* expressing the *Aspergillus nidulans* phosphoketolase pathway: validation of activity through  $^{13}\text{C}$ -based metabolic flux analysis. *Appl Microbiol Biotechnol* 95(4):1001–1010. doi:[10.1007/s00253-012-3936-0](https://doi.org/10.1007/s00253-012-3936-0)
- Payne GA, Nierman WC, Wortman JR, Pritchard BL, Brown D, Dean RA, Bhatnagar D, Cleveland TE, Machida M, Yu J (2006) Whole genome comparison of *Aspergillus flavus* and *A. oryzae*. *Med Mycol* 44(s1):9–11. doi:[10.1080/13693780600835716](https://doi.org/10.1080/13693780600835716)
- Pedersen H, Carlsen M, Nielsen J (1999) Identification of enzymes and quantification of metabolic fluxes in the wild type and in a recombinant *Aspergillus oryzae* strain. *Appl Environ Microbiol* 65(1):11–19
- Peleg Y, Barak A, Scrutton M, Goldberg I (1989) Malic acid accumulation by *Aspergillus flavus*. *Appl Microbiol Biotechnol* 30(2):176–183. doi:[10.1007/bf00264008](https://doi.org/10.1007/bf00264008)
- Regev-Rudzki N, Battat E, Goldberg I, Pines O (2009) Dual localization of fumarase is dependent on the integrity of the glyoxylate shunt. *Mol Microbiol* 72(2):297–306. doi:[10.1111/j.1365-2958.2009.06659.x](https://doi.org/10.1111/j.1365-2958.2009.06659.x)
- Sakai K, Kinoshita H, Nihira T (2012) Heterologous expression system in *Aspergillus oryzae* for fungal biosynthetic gene clusters of secondary metabolites. *Appl Microbiol Biotechnol* 93(5):2011–2022. doi:[10.1007/s00253-011-3657-9](https://doi.org/10.1007/s00253-011-3657-9)
- Scalcani G, Otero JM, Van Vleet JRH, Jeffries TW, Olsson L, Nielsen J (2012) Evolutionary engineering of *Saccharomyces cerevisiae* for efficient aerobic xylose consumption. *FEMS Yeast Res* 12(5):582–597. doi:[10.1111/j.1567-1364.2012.00808.x](https://doi.org/10.1111/j.1567-1364.2012.00808.x)
- Schmidt K, Marx A, de Graaf AA, Wiechert W, Sahm H, Nielsen J, Villadsen J (1998)  $^{13}\text{C}$  Tracer experiments and metabolite balancing for metabolic flux analysis: comparing two approaches. *Biotechnol Bioeng* 58(2–3):254–257. doi:[10.1002/\(sici\)1097-0290\(19980420\)58:2/3<254::aid-bit19>3.0.co;2-c](https://doi.org/10.1002/(sici)1097-0290(19980420)58:2/3<254::aid-bit19>3.0.co;2-c)
- Shen Y, Chen X, Peng B, Chen L, Hou J, Bao X (2012) An efficient xylose-fermenting recombinant *Saccharomyces cerevisiae* strain obtained through adaptive evolution and its global transcription profile. *Appl Microbiol Biotechnol* 96(4):1079–1091. doi:[10.1007/s00253-012-4418-0](https://doi.org/10.1007/s00253-012-4418-0)
- Udvardi MK, Czechowski T, Scheible W-R (2008) Eleven golden rules of quantitative RT-PCR. *Plant Cell* 20(7):1736–1737. doi:[10.1105/tpc.108.061143](https://doi.org/10.1105/tpc.108.061143)
- Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. *Nucleic Acids Res* 40(15):e115. doi:[10.1093/nar/gks596](https://doi.org/10.1093/nar/gks596)
- van Winden W, Wittmann C, Heinze E, Heijnen J (2002) Correcting mass isotopomer distributions for naturally occurring isotopes. *Biotechnol Bioeng* 80:477–479
- Vongsangnak W, Olsen P, Hansen K, Krogsgaard S, Nielsen J (2008) Improved annotation through genome-scale metabolic modeling of *Aspergillus oryzae*. *BMC Genomics* 9(1):245
- Wiechert W (2001)  $^{13}\text{C}$  metabolic flux analysis. *Metab Eng* 3:195–206
- Wittmann C (2007) Fluxome analysis using GC-MS. *Microb Cell Fact* 6(1):6
- Zelle RM, de Hulster E, van Winden WA, de Waard P, Dijkema C, Winkler AA, Geertman J-M, van Dijken JP, Pronk JT, van Maris AJA (2008) Malic acid production by *Saccharomyces cerevisiae*: engineering of pyruvate carboxylation, oxaloacetate reduction, and malate export. *Appl Environ Microbiol* 74(9):2766–2777. doi:[10.1128/aem.02591-07](https://doi.org/10.1128/aem.02591-07)



## Paper IV

Aspergilli: Systems biology and industrial applications

**Knuf C, Nielsen J.** 2012

Biotechnology Journal (2012) 7:1147-1155



## Review

# Aspergilli: Systems biology and industrial applications

Christoph Knuf<sup>1</sup> and Jens Nielsen<sup>1,2</sup>

<sup>1</sup> Department of Chemical and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden

<sup>2</sup> Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark

Aspergilli are widely used as cell factories for the production of food ingredients, enzymes and antibiotics. Traditionally, improvement of these cell factories has been done using classical methods, that is, random mutagenesis and screening; however, advances in methods for performing directed genetic modifications has enabled the use of metabolic engineering strategies. Genome sequencing of Aspergilli was originally trailing behind developments in the field of bacteria and yeasts, but with the recent availability of genome sequences for several industrially relevant Aspergilli, it has become possible to implement systems biology tools to advance metabolic engineering. These tools include genome-wide transcription analysis and genome-scale metabolic models. Herein, we review achievements in the field and highlight the impact of *Aspergillus* systems biology on industrial biotechnology.

Received 24 APR 2012

Revised 25 JUN 2012

Accepted 10 JUL 2012

**Keywords:** *Aspergillus* · Industrial biotechnology · Metabolic engineering · Systems biology

## 1 Introduction

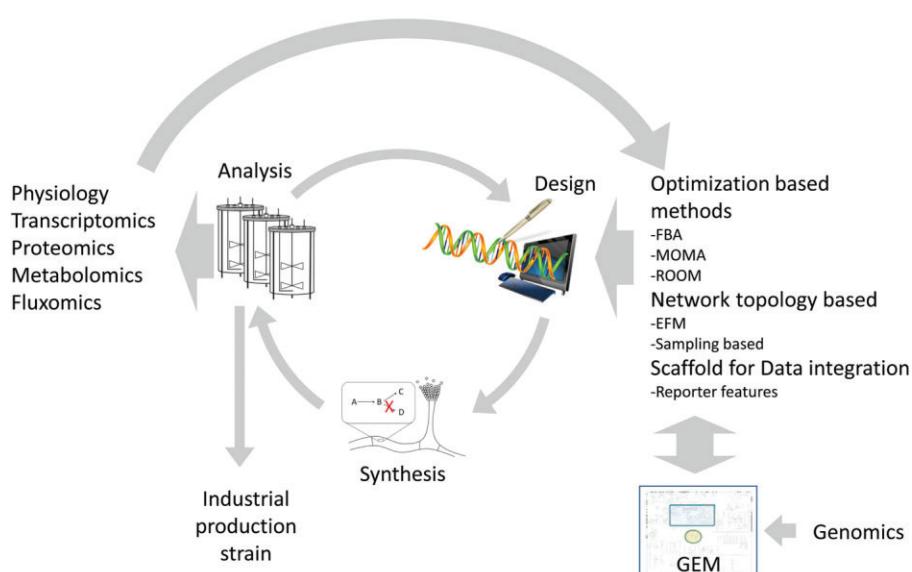
Systems biology started to evolve rapidly after the publication of microbial genomes in the mid-1990s (e.g. *Haemophilus influenza* [1], *Saccharomyces cerevisiae* [2], *Escherichia coli* [3]). Most progress in the field of systems biology has been seen for the platform and model organisms *S. cerevisiae* and *E. coli* [4]. With the availability of the genome sequence of the Aspergilli (*A. nidulans*, *A. niger*, and *A. oryzae*), systems biology for Aspergilli have become possible. Aspergilli are among the top organisms for industrial large-scale production of acids and enzymes for the chemical industry. In general, Aspergilli show good acid tolerance and a broad substrate spectrum because they are saprophytic organisms. In that biological niche they evolved a natural ability for high enzyme secretion, including a broad spectrum of hydrolytic enzymes. *A. niger*, in particular, shows a natural toler-

ance towards acidic media and several fermentation products are considered generally recognized as safe (GRAS) [5]. Therefore, it was the organism of choice for an annual production of 1.6 million tons of citric acid in 2007 [6]. The second industrial workhorse is *A. oryzae*, which combines a GRAS status and an extremely high enzyme secretion capacity. Because both organisms have GRAS status, new fermentation processes and the products thereof are more easily approved by the authorities. Furthermore, large-scale production processes have been developed and the fermentation industry, especially DSM (*A. niger*), Novozymes (*A. oryzae*), and Danisco-Genencor (*A. niger* and *A. oryzae*), have decades of experience with these filamentous fungi. These attributes make them the perfect platform for further improvement of existing processes and also for the development of future applications to expand the *Aspergillus* production portfolio.

Since visionary papers by Bailey [7] and Stephanopoulos and Valliano [8], in which they stressed the term metabolic engineering, this field has really taken off and new industrial production strains are being created with the help of directed genetic modifications. The limitation of practical tools for genetic modifications has long been overcome, so that the synthesis step of the metabolic engineering cycle (Fig. 1) is no longer a bottleneck, leaving the design and analysis steps open to improvement. Here

**Correspondence:** Prof. Jens Nielsen, Department of Chemical and Biological Engineering, Chalmers University of Technology, Kemivägen 10, SE-412 96 Göteborg, Sweden  
**E-mail:** nielsenj@chalmers.se

**Abbreviations:** GEM, genome-scale metabolic model; GRAS, generally recognized as safe; 6-MSA, 6-methylsalicylic acid



**Figure 1.** Extended metabolic engineering cycle. Metabolic engineering targets are identified and designed in the design part of the cycle. Subsequently the organism of choice is manipulated according to the previous design. The organism carrying the manipulation is then analyzed. The results of the analysis are again used in the design part and the cycle starts over again until the analysis of a new strain shows results matching the needs of the industry, regarding yield or final titer, for example. This strain leaves the cycle and enters production. Systems biology supports the analysis part with the fields transcriptomics, proteomics, metabolomics, and fluxomics and supplies genome-scale metabolic models (GEMs) as analysis frameworks that originate from genomics. Different computational biology approaches allow a better design of metabolic engineering targets and further refinement of the GEM. EFM, elementary flux mode; FBA, flux balance analysis; MOMA, minimization of metabolic adjustments; ROOM, regulatory on-off minimization.

industrial systems biology comes into play [9]. Genomics and the resulting genome-scale metabolic models (GEMs) are in symbiosis with huge datasets created by high-throughput analysis tools, such as transcriptomics, proteomics, metabolomics, and fluxomics; this allows multi-level pictures to be drawn with a hitherto unseen resolution of metabolism and its regulation. This picture is used for the prediction of future metabolic engineering targets using different computational biology approaches. Taken together, systems biology combines tools for the analysis and design part, and thereby, acts as a catalyst for the classic metabolic engineering cycle, accelerating the turnover time and leading towards the only exit, industrial production strains, more rapidly.

## 2 Systems biology toolbox of *Aspergillus*

The scaffold of the systems biology toolbox comprises mathematical models. The best examples are GEMs [10–12] that allow a mathematical description of the complex systems of interest: living cells with multilevel organization and regulation. To capture the different levels, omics studies are applied, namely, genomics, transcriptomics, proteomics, and metabolomics [13–15]. In the following sections, these are briefly summarized; for a more detailed review, we recommend the recent review by Andersen and Nielsen [16].

### 2.1 Genomics

The concept of genomics was born in 1977 when the first whole genome sequence was published, namely, the 5375 base pair (bp), large viral genome of bacteriophage  $\Phi$ X174, which could be completely deciphered [17]. Several genomes followed and a large number of microbial genomes were published in the mid-1990s. The first *Aspergillus* genome sequencing project was initiated by Cereon Genomics (Monsanto) in 1998, but the sequence of *A. nidulans* with a three-fold coverage was first released to the public some years later. In 2003, the Cereon sequence was combined with the results of the Whitehead Institute/MIT Center for Genome Research (now the BROAD institute) to give 13-fold coverage, which was released in March 2003. Table 1 gives an overview of the sequenced *Aspergillus* genomes.

Over three decades of genomics, the price for sequencing has dropped dramatically. The shift from the electrophoretic Sanger sequencing method towards the recently refined and more automated 'second-generation DNA sequencing' techniques [18] had a major impact on the sequencing price. The significant decrease in sequencing cost of over two orders of magnitude allows this technique to even become a routine technique in systems biology and metabolic engineering. Next-generation sequencing techniques are not only used for sequencing genomes, but also for transcriptome analysis [19].

**Table 1.** Overview of sequenced *Aspergillus* genomes, including the sequenced strain, the genome size, the calculated number of genes

	Strain	Genome size (Mb)	Genes	Reference
<i>A. nidulans</i>	FGSC A4	30.1	9 396	[75]
<i>A. niger</i>	CBS 513.88	33.9	14 165	[76]
<i>A. niger</i>	ATCC 1015	34.9	11 200	[77]
<i>A. oryzae</i>	RIB40	37.2	12 074	[78]
<i>A. fumigatus</i>	AF293	29.4	9 926	[79]
<i>A. terreus</i>	NIH 2624	29.3	10 406	–

The availability of the whole-genome sequence of an organism does not reveal many secrets. For further analysis, the more specific fields of functional genomics and comparative genomics need to be applied. The aim of functional genomics is to establish the link between predicted genes and their function. In the case of Aspergilli this is a difficult undertaking. In the case of *A. oryzae*, for example, more than 40% of the 12 074 predicted genes could not be assigned a function [20]. Comparative genomics tries to reveal evolutionary differences between strains or species. By comparing two *A. niger* strains, characteristic key enzymes and pathways for their industrial fermentation application could be found. The enzyme producer CBS 513.88 showed an enrichment of glucoamylase A, tRNA synthase, and transporter genes involved in amino acid metabolism and protein synthesis [21]. On the other hand, a high number of genes and pathways for organic acid production were identified in the organic acid producer ATCC1015 [21].

## 2.2 Mathematical modeling

Mathematical modeling tries to formulate cellular processes in a mathematical framework to obtain conclusive predictions of cellular responses towards perturbations. Models serve as frameworks for data integration from transcriptomics, proteomics, or metabolomics studies, too. In general, two kinds of models can be distinguished [22]: stoichiometric and kinetic models. The framework for kinetic modeling is metabolic control analysis [23], which focuses on the quantification of the influence of each enzyme on the overall flux of the pathway. The coefficients for quantification are elasticity coefficients, which are "local properties, in that they define the response of a particular velocity of a reaction to a change in an effector of that reaction" and flux control coefficients (FCC), which are "properties of the intact metabolic system and define the scaled sensitivity of a system variable (e.g. metabolite concentration, flux) to a system parameter (e.g. enzyme activity or concentration)" [24]. The drawback is that the FCC are difficult to obtain because they require detailed information on the kinetics of the enzymes involved and solid data on the intracellular metabolite levels [16]. The S-system has been proposed to circumvent that problem. Here a power-law function of all

metabolites in the network describes the kinetics of each reaction in the network [25]. Detailed kinetic models for *A. niger* have been developed by Torres [26, 27] and subsequently improved to finally include the main citric acid pathways, including mitochondrial reactions [28]. Stoichiometric modeling involves metabolic flux analysis or FBA. These methods use a list of biochemical reactions available in the cell to calculate carbon fluxes inside the cell from external fluxes (metabolite secretion and substrate uptake) and biomass changes by applying mass balances of each metabolite. Stoichiometric models at the genome scale, so-called GEMs, are available for three *Aspergillus* species: *A. nidulans* [29], *A. niger* [30–32], and *A. oryzae* [20].

Although GEMs contain a lot of knowledge, they lack information about regulation, for example, on the transcript or enzyme level. Because this influences the activities of pathways, this information has to be integrated into GEMs in the future to make them truly predictive. The question that arises at this point is how detailed a model has to be to still be computable and helpful. Furthermore, will it ever be possible to gain such a detailed insight to have the information needed at hand?

## 2.3 Transcriptomics

For analysis of the transcriptome, several platforms from varying manufacturers have been developed, ranging from custom-made spotted arrays to prefabricated oligonucleotide chips (Affymetrix GeneChip technology). For the three most important Aspergilli, *A. niger*, *A. nidulans*, and *A. oryzae*, a chip is available that enables a convenient comparison of the transcriptome of the three species [33]. By using this array, a study of carbon-source utilization in *A. oryzae* and *A. niger* revealed two different regulatory systems for maltose utilization [34] and a conserved regulatory system was identified for glycerol metabolism in the three Aspergilli [35]. This example highlights the importance of microarrays for the investigation of complex regulatory systems by looking at the snapshot of the expression of thousands of genes in a certain metabolic state and is therefore the method of choice for untangling regulatory networks. Because of the poor correlation of transcription and translation, it is difficult to use

transcriptome data directly to find metabolic engineering targets [36].

## 2.4 Proteomics

Another field of systems biology is proteomics. Because it can directly contribute to the optimization of fermentation processes and screening of biotechnologically relevant enzymes, it has a direct influence on the field of biotechnology [37]. A 2007 review of proteomics in filamentous fungi gave the following definition of the proteome as being the “global set of proteins expressed in a cell at a given time and biological state” [38]. To detect the global set of proteins, different methods can be applied, including isotope-coded affinity tags, protein arrays [39], yeast two hybrid systems [40], two-dimensional polyacrylamide gel electrophoresis combined with MS of excised proteins [41], and HPLC followed by MS [42]. The latest advances in MS led to the development of proteomics as a field and the almost exclusive use of MS-based techniques for modern proteomics [43]. The sample preparation methods are mostly optimized for *E. coli* and *S. cerevisiae*, the cell walls of which are easy to disrupt. The opposite is true for Aspergilli; therefore, cell lysis is a limiting step. Another obstacle for *Aspergillus* proteomics is the strong glycosylation of proteins; this affects the masses of the proteins and the isoelectric point significantly. The above-mentioned review by Kim et al. [38] summarized proteomics methods specially developed for Aspergilli and a review by de Oliveira and de Graaf [44] summarized trends and insights for proteomics of industrial fungi. In a recent case study, the effect of the isoprenoid alcohol farnesol was investigated in *A. nidulans* [45]. Forty-three proteins with significantly altered abundance were identified, confirming the targeting of mitochondria and induction of the unfolded protein response and oxidative stress by farnesol. Furthermore, the data was used to identify a dehydrin-like protein (designated DlpA), which was strongly induced. Deletion of this protein led to the sensitivity of conidia to oxidative and temperature stress. This shows that proteomics can help the assignment of biological functions to hitherto uncharacterized genes and proteins.

## 2.5 Metabolomics

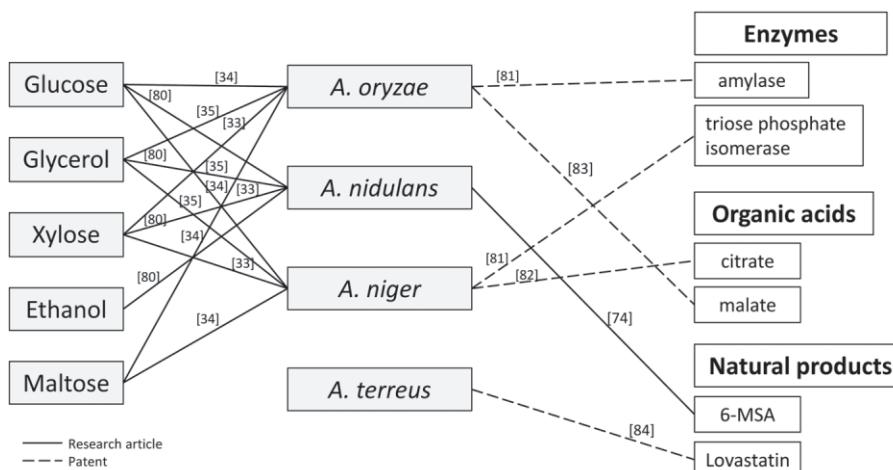
The term metabolome was first used in the literature in 1998 [46] and it is defined as “the quantitative complement of low-molecular-weight metabolites present in a cell under a given set of physiological conditions” [47]. Metabolome studies can be classified into three types: metabolic profiling, metabolic fingerprinting, and metabolic footprinting. According to the definitions given by Ellis et al. [48], metabolic profiling can be understood as the targeted analysis of metabolites known to be involved in a certain pathway of interest, whereas finger- and foot-

printing tries to identify all metabolites. Metabolic fingerprinting is defined as a comparative screening method using global, high-throughput, rapid analysis for sample classification, whereas metabolic footprinting is aimed at the analysis of exo-metabolites secreted by cells. Similar to the proteomics approach, metabolomics also employs sensitive spectrometry techniques, such as TOF-MS and normal MS, usually coupled to a chromatographic method, such as GC and LC [49]. A method specially developed for fungi uses direct infusion MS [50]. Because there are less metabolites than genes and proteins in a cell, metabolome studies have to deal with less data, which accelerates high-throughput studies, additionally, in some cases, genetic or environmental perturbations do not affect transcription or translation in a significant way, whereas enzyme activities and subsequently metabolite concentrations can be greatly affected [46]. These reasons lead to the consideration that metabolomics is “more discriminatory than transcriptomics and proteomics” [51].

To verify this assumption, Kouskoumvekaki et al. [51] used changes in metabolite concentrations to find biomarkers in recombinant 6-methylsalicylic acid (6-MSA) producers for metabolic engineering applications. Out of more than 450 detected metabolites, 10 biomarkers were identified. The downside was that six of these could not be identified with the authors’ in-house library. This underlines the need for common databases for metabolome analysis.

## 2.6 Fluxomics

The omics discipline that gives the most essential information for industrial biotechnology is the area of fluxomics. The aim is to describe the entire intracellular flux distribution from the carbon source towards the product and byproducts. With this knowledge, it is possible to design rational metabolic engineering targets. Fluxomics is the essence of metabolism because it represents the final outcome of regulation at different levels [52] measured separately with the previously described omics methods. For the calculation of intracellular fluxes, a pseudo-steady state is assumed. Through linear programming, it is possible to calculate intracellular fluxes from measured fluxes such as specific growth rate, substrate consumption rates, and production rates. However, for GEMs the number of degrees of freedom is too large, so calculations from external fluxes become problematic. To solve this problem, the technique of feeding specific <sup>13</sup>C-labeled substrates can be applied. By measuring the enriched labeling patterns of metabolites by NMR spectroscopy [53, 54] or MS [55, 56], it is possible to obtain information about the distribution of carbon fluxes to the different pathways. Because the intracellular metabolite pools are usually small, labeling patterns of amino acids in proteins are measured and from that the flux distribution is estimated



**Figure 2.** Connectivity graph of carbon sources, *Aspergillus* family members, and industrial products. Numbers above the connecting lines correlate to articles (solid line) or patents (dashed line) [33–35, 74, 80–84].

using mathematical models. For more comprehensive information, see [57].

### 3 Industrial biotechnology

Harnessing Aspergilli for human interests has a long tradition. *A. oryzae*, for example, has been used in the Japanese food industry for the production of Sake, soy sauce, and miso for several centuries. One of the first industrial products that originated from *A. oryzae* was the diastatic enzyme Taka-Diastase. This enzyme mix was the first microbial enzyme marketed in the United States. A patent (US Patent 525,823) was issued in 1895 to Jokichi Takamine, the father of American biotechnology. Since then Aspergilli are becoming the industrial workhorses for large-scale industrial production. The genus of *Aspergillus* comprises more than 180 species, with the most well known being *A. awamori*, *A. flavus*, *A. fumigatus*, *A. nidulans*, *A. niger*, and *A. terreus*. Aspergilli affect humans in many ways. *A. fumigatus* is an example of a human pathogen [58], whereas others produce secondary metabolites that are used as pharmaceutical drugs, for example, the production of the cholesterol-lowering drug lovastatin by *A. terreus* [59]. The most recognized Aspergilli in biotechnology are *A. oryzae* and *A. niger*, which are used in enzyme [60] and organic acid [61] production. In the following examples, how systems biology affected industrial biotechnology of Aspergilli is highlighted using investigations concerning substrate utilization, organic acid production, enzyme production, and natural products as case studies (Fig. 2).

#### 3.1 Substrate utilization

The substrate utilization of microorganisms has a significant impact on fermentation. For understanding the underlying control mechanism of carbon-source uptake, mi-

croarrays have been the method of choice. In recent years, several different carbon sources have been under investigation, among those glycerol [35], maltose [34], xylose, and glucose [33].

The maltose study compared the transcriptional profile of *A. oryzae* and *A. niger* with glucose and maltose as the carbon source. For *A. oryzae*, 16 significantly upregulated genes were found in the transcriptome among which the transcription of the *MAL* cluster, which is used in *S. cerevisiae* as well, could be confirmed. On the other hand, no significantly up-regulated genes could be found for *A. niger* and the authors concluded that the utilization of maltose would happen via extracellular glucoamylase regulated by *AmyR* [62].

#### 3.2 Organic acids

The dominant organic acid produced by Aspergilli is citric acid, but Aspergilli are also capable of producing fumaric and malic acid [63]. Industrial citric acid production started out with surface cultures in the 1930s, but was then converted to submerged cultures in the late 1940s [64]. Although this process has been applied industrially for over half a century, the mechanisms leading to citric acid production are still nebulous. Investigations on a systems level, including modeling and transcriptome analysis, tried to shed some light into the black box of citric acid production by investigating the influence of ambient pH [65]. In this study 109 genes directly influenced by varying pH levels were found. The outcome of this investigation was that Aspergilli evolved mechanisms for the production of organic acids to outgrow other microorganisms. The production of gluconic acid, which is used as the carbon source in later stages of fermentations, takes place to convert the easily accessible carbon source glucose into a carbon source that is not used preferentially by other microorganisms. The secretion of oxalic and citric acid aims to acidify the medium and is not a cause of

overflow metabolism. The authors concluded that “an aggressive acidification of the microenvironment combined with storing the available glucose as gluconic acid” [65] were *A. niger*’s strategies for outcompeting other microorganisms.

### 3.3 Enzyme production

Aspergilli are saprophytic fungi that degrade biomass extracellularly through excreted enzymes. This evolved mechanism leads to their overall optimized ability for large-scale enzyme production. *A. oryzae*, for example, is used for the industrial production of  $\alpha$ -amylase [66]. Although the industrial process has been running for years, it was unclear which regulation mechanisms led to the enormous secretion potential. Therefore, an industrial strain overexpressing  $\alpha$ -amylase and the wild-type A1560 were compared for their physiology and transcription profiles. The mutant strain showed a significant up-regulation of genes of nucleotide metabolism and several amino acid biosynthetic enzymes. Furthermore, using yeast as a scaffold, a protein–protein interaction network was generated and subsequent analysis of the transcription data led to the identification of 33 proteins that were possible key regulators. The abundance of Gcn2/CpcC and Hac1 among the 33 possible key regulators led to the conclusion that protein production was possibly limited by protein folding, since the unfolded protein response was activated [60].

In another study, integration of in silico and in vivo metabolic fluxes through application of elementary flux mode analysis and the correlation of fluxes to desired network properties [67] helped to identify new metabolic engineering strategies to improve fructofuranosidase production by *A. niger* [68]. By comparing a computed optimal flux distribution for fructofuranosidase production to the in vivo fluxes measured by  $^{13}\text{C}$  flux analysis, mitochondrial reactions catalyzed by malic enzyme and malate dehydrogenase and gluconate production were identified. Elimination of these undesired reactions has not been proven yet, but the approach seems to be promising and might facilitate the quest for metabolic engineering targets in the future.

### 3.4 Natural products

The production ability of a wide range of natural products, bioactive secondary metabolites[69], put the spotlight on the genus *Aspergillus*. On one hand, the positive effects of natural products, such as antibiotic, immunosuppressant, lipid-/cholesterol-lowering, or antifungal properties, are applied in the pharma and agricultural industry for example, the cholesterol-lowering polyketide lovastatin [59]. On the other hand, the secondary metabolites produced by Aspergilli can exhibit phyto- and mycotoxic activities. Ex-

amples are the toxic and hepatocarcinogenic class of aflatoxins produced by *A. flavus* [70] and fumonisin B2 [71].

The first example of industrially relevant research on natural products in Aspergilli dealt with fumonisin B<sub>2</sub> production in *A. niger* [72]. Although most industrial applications of *A. niger* have GRAS status, this organism bears the potential for fumonisin production. Therefore, the media composition of industrial processes was investigated by proteome analysis with emphasis on the conditions triggering the production of the mycotoxin. The authors concluded that increased flux through acetyl-CoA and high NADPH regeneration capacities lead to fumonisin B2 production in nitrate and starch medium with added lactate.

Other examples aimed to increase polyketide production, especially 6-methylsalicylate, which is a model compound for the evaluation of polyketide production capabilities [73]. The first study employed metabolite profiling and non-linear methods for clustering. They compared 6-methylsalicylate overproducing strains and wild-type *A. nidulans* strains to identify biomarkers for polyketide production that could aid metabolic engineering [51]. Another systems biology study investigated the same topic by combining transcriptome, flux, and physiology data with an interaction model that describes the competition between biomass formation and 6-MSA production for available acetyl CoA [74].

## 4 Perspectives

Systems biology is a well-established field in life sciences and has been shown to improve the basic understanding of cellular mechanisms and led to the identification of metabolic engineering targets. The sub-disciplines genomics, transcriptomics, proteomics, metabolomics, and fluxomics are well established for model organisms like *S. cerevisiae* and *E. coli*. The systems biology developments in the *Aspergillus* field trail behind; however, basic systems biology are available, but could still be improved.

Although it is difficult to put numbers on the involvement of systems biology in the latest improvements in strain design for industrial applications, the number of research articles published that involve systems biology approaches to identify new metabolic engineering targets is constantly increasing. Furthermore, most large biotech companies have a systems biology group to help with strain design and analysis.

A product class with growing interest and great future potential is the polyketides. With falling sequencing costs, even less popular members of the *Aspergillus* family will be sequenced and systems biology will have a huge impact on unraveling the strictly regulated pathways for secondary metabolite production.

With the evolution of a new hot topic in life sciences, namely, synthetic biology, the importance of systems bi-

ology will further grow, since the evaluation of heterologous pathway expression and interactions with the host metabolic network are prime examples for which a systems biology approach is beneficial.

We acknowledge funding for this research work from European Research Council (grant no. 247013), the Knut and Alice Wallenberg Foundation and the Chalmers Foundation.

The authors declare no conflict of interest

## 5 References

- [1] Smith, H. O., Tomb, J. F., Dougherty, B. A., Fleischmann, R. D., Venter, J. C., Frequency and distribution of DNA uptake signal sequences in the *Haemophilus influenzae* Rd genome. *Science* 1995, 269, 538–540.
- [2] Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W. et al., Life with 6000 Genes. *Science* 1996, 274, 546–567.
- [3] Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T. et al., The complete genome sequence of *Escherichia coli* K-12. *Science* 1997, 277, 1453–1462.
- [4] Nielsen, J., Keasling, J. D., Synergies between synthetic biology and metabolic engineering. *Nat. Biotechnol.* 2011, 29, 693–695.
- [5] Schuster, E. S., Dunn-Coleman, N. D.-C., Frisvad, J. F., van Dijck, P. v. D., On the safety of *Aspergillus niger*—a review. *Appl. Microbiol. Biotechnol.* 2002, 59, 426–435.
- [6] Berovic, M., Legisa, M., El-Gewely, M. R., Citric acid production. *Biotechnol. Annu. Rev.* 2007, 13, 303–343.
- [7] Bailey, J. E., Toward a science of metabolic engineering. *Science* 1991, 252, 1668–1675.
- [8] Stephanopoulos, G., Vallino, J. J., Network rigidity and metabolic engineering in metabolite overproduction. *Science* 1991, 252, 1675–1681.
- [9] Otero, J. M., Nielsen, J., Industrial systems biology. *Biotechnol. Bioeng.* 2010, 105, 439–460.
- [10] Price, N. D., Reed, J. L., Palsson, B. O., Genome-scale models of microbial cells: Evaluating the consequences of constraints. *Nat. Rev. Microbiol.* 2004, 2, 886–897.
- [11] Kim, T. Y., Sohn, S. B., Kim, Y. B., Kim, W. J., Lee, S. Y., Recent advances in reconstruction and applications of genome-scale metabolic models. *Curr. Opin. Biotechnol.* 2011. In press, DOI: 10.1016/j.copbio.2011.10.007.
- [12] Cvijovic, M., Bordel, S., Nielsen, J., Mathematical models of cell factories: Moving towards the core of industrial biotechnology. *Microb. Biotechnol.* 2011, 4, 572–584.
- [13] Patil, K. R., Nielsen, J., Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proc. Natl. Acad. Sci. USA* 2005, 102, 2685–2689.
- [14] Oliveira, A., Patil, K., Nielsen, J., Architecture of transcriptional regulatory circuits is knitted over the topology of bio-molecular interaction networks. *BMC Syst. Biol.* 2008, 2, 17.
- [15] Cakir, T., Patil, K. R., Onsan, Z. I., Ulgen, K. O. et al., Integration of metabolome data with metabolic networks reveals reporter reactions. *Mol. Syst. Biol.* 2006, 2.
- [16] Andersen, M. R., Nielsen, J., Current status of systems biology in Aspergilli. *Fungal Genet. Biol.* 2009, 46, S180–S190.
- [17] Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L. et al., Nucleotide sequence of bacteriophage [phi]X174 DNA. *Nature* 1977, 265, 687–695.
- [18] Shendure, J., Ji, H., Next-generation DNA sequencing. *Nat. Biotechnol.* 2008, 26, 1135–1145.
- [19] Ozsolak, F., Milos, P. M., RNA sequencing: Advances, challenges and opportunities. *Nat. Rev. Genet.* 2011, 12, 87–98.
- [20] Vongsangnak, W., Olsen, P., Hansen, K., Krogsbaard, S., Nielsen, J., Improved annotation through genome-scale metabolic modeling of *Aspergillus oryzae*. *BMC Genomics* 2008, 9, 245.
- [21] Andersen, M. R., Salazar, M. P., Schaap, P. J., van de Vondervoort, P. J. I. et al., Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. *Genome Res.* 2011, 21, 885–897.
- [22] Nielsen, J., Metabolic engineering: Techniques for analysis of targets for genetic manipulations. *Biotechnol. Bioeng.* 1998, 58, 125–132.
- [23] Fell, D. A., Sauro, H. M., Metabolic control analysis. *Eur. J. Biochem.* 1990, 192, 183–187.



**Jens Nielsen** has an M.Sc. degree in Chemical Engineering and a Ph.D. degree (1989) in Biochemical Engineering from the Danish Technical University (DTU). He then established his independent research group and was appointed full Professor in 1998. He was Fulbright visiting professor at MIT in 1995–1996. At DTU he founded and directed the Center for Microbial Biotechnology. In 2008 he was recruited as Professor and Director at the Chalmers University of Technology, Sweden, where he is currently directing a research group of more than 50 people and the Life Science Area of Advance, which coordinates over 200 researchers from 5 departments. He has published more than 350 research papers that have been cited more than 10 300 times (current H-index 52), co-authored more than 40 books, and is an inventor of more than 50 patents. He has founded several companies that have raised more than €25 million in venture capital.



**Christoph Knuf** is a PhD student in the Systems and Synthetic Biology Group. He received his B.Sc. and M.Sc. degrees from the University of Münster, Germany. His bachelor's thesis dealt with the development of an in vivo activity assay for plant *cis*-prenyltransferases at the Institute for Plant Biology and Biotechnology. For his master's thesis, he worked on metabolic engineering of the FPP branch point in *Saccharomyces cerevisiae* for improved isoprenoid production. For his Ph.D. studies, he has chosen the more challenging organism *Aspergillus oryzae* and combines systems biology and metabolic engineering to understand and improve organic acid production.

- [24] Small, J. R., Fell, D. A., Metabolic control analysis. *Eur. J. Biochem.* 1990, **191**, 413–420.
- [25] Voit, E. O., Smooth bistable s-systems. *IEE Proc.: Syst. Biol.* 2005, **152**, 207–213.
- [26] Torres, N. V., Modeling approach to control of carbohydrate metabolism during citric acid accumulation by *Aspergillus niger*: I. Model definition and stability of the steady state. *Biotechnol. Bioeng.* 1994, **44**, 104–111.
- [27] Torres, N. V., Modeling approach to control of carbohydrate metabolism during citric acid accumulation by *Aspergillus niger*: II. Sensitivity analysis. *Biotechnol. Bioeng.* 1994, **44**, 112–118.
- [28] Alvarez-Vasquez, F., González-Alcón, C., Torres, N. V., Metabolism of citric acid production by *Aspergillus niger*: Model definition, steady-state analysis and constrained optimization of citric acid production rate. *Biotechnol. Bioeng.* 2000, **70**, 82–108.
- [29] David, H., Hofmann, G., Oliveira, A., Jarmer, H., Nielsen, J., Metabolic network driven analysis of genome-wide transcription data from *Aspergillus nidulans*. *Genome Biol.* 2006, **7**, R108.
- [30] Andersen, M. R., Nielsen, M. L., Nielsen, J., Metabolic model integration of the bioRxome, genome, metabolome and reactome of *Aspergillus niger*. *Mol. Syst. Biol.* 2008, **4**.
- [31] David, H., Åkesson, M., Nielsen, J., Reconstruction of the central carbon metabolism of *Aspergillus niger*. *Eur. J. Biochem.* 2003, **270**, 4243–4253.
- [32] Melzer, G., Dalpiaz, A., Grote, A., Kucklick, M. et al., Metabolic flux analysis using stoichiometric models for *Aspergillus niger*: Comparison under glucoamylase-producing and non-producing conditions. *J. Biotechnol.* 2007, **132**, 405–417.
- [33] Andersen, M. R., Vongsangnak, W., Panagiotou, G., Salazar, M. P. et al., A trispecies *Aspergillus* microarray: Comparative transcriptomics of three *Aspergillus* species. *Proc. Natl. Acad. Sci. USA* 2008, **105**, 4387–4392.
- [34] Vongsangnak, W., Salazar, M., Hansen, K., Nielsen, J., Genome-wide analysis of maltose utilization and regulation in aspergilli. *Microbiology* 2009, **155**, 3893–3902.
- [35] Salazar, M., Vongsangnak, W., Panagiotou, G., Andersen, M., Nielsen, J., Uncovering transcriptional regulation of glycerol metabolism in Aspergilli through genome-wide gene expression data analysis. *Mol. Genet. Genomics* 2009, **282**, 571–586.
- [36] Jewett, M., Oliveira, A., Patil, K., Nielsen, J., The role of high-throughput transcriptome analysis in metabolic engineering. *Biotechnol. Bioprocess Eng.* 2005, **10**, 385–399.
- [37] Kniemeyer, O., Proteomics of eukaryotic microorganisms: The medically and biotechnologically important fungal genus *Aspergillus*. *PROTEOMICS* 2011, **11**, 3232–3243.
- [38] Kim, Y., Nandakumar, M. P., Marten, M. R., Proteomics of filamentous fungi. *Trends Biotechnol.* 2007, **25**, 395–400.
- [39] Zhu, H., Bilgin, M., Bangham, R., Hall, D. et al., Global analysis of protein activities using proteome chips. *Science* 2001, **293**, 2101–2105.
- [40] Auerbach, D., Thaminy, S., Hottiger, M. O., Stagljar, I., The post-genomic era of interactive proteomics: Facts and perspectives. *Proteomics* 2002, **2**, 611–623.
- [41] Bader, G. D., Hogue, C. W. V., Analyzing yeast protein–protein interaction data obtained from different sources. *Nat. Biotechnol.* 2002, **20**, 991–997.
- [42] Ouyang, H., Luo, Y., Zhang, L., Li, Y., Jin, C., Proteome analysis of *Aspergillus fumigatus*: Total membrane proteins identifies proteins associated with the glycoconjugates and cell wall biosynthesis Using 2D LC-MS/MS. *Mol. Biotechnol.* 2010, **44**, 177–189.
- [43] Moore, J. B., Weeks, M. E., Proteomics and systems biology: Current and future applications in the nutritional sciences. *Adv. Nutr.* 2011, **2**, 355–364.
- [44] de Oliveira, J., de Graaff, L., Proteomics of industrial fungi: Trends and insights for biotechnology. *Appl. Microbiol. Biotechnol.* 2011, **89**, 225–237.
- [45] Wartenberg, D., Vödisch, M., Kniemeyer, O., Albrecht-Eckardt, D. et al., Proteome analysis of the farnesol-induced stress response in *Aspergillus nidulans*—the role of a putative dehydrin. *J. Proteomics* 2012, **13**, 4038–4049.
- [46] Oliver, S. G., Winson, M. K., Kell, D. B., Baganz, F., Systematic functional analysis of the yeast genome. *Trends Biotechnol.* 1998, **16**, 373–378.
- [47] Kell, D. B., Brown, M., Davey, H. M., Dunn, W. B. et al., Metabolic footprinting and systems biology: The medium is the message. *Nat. Rev. Microbiol.* 2005, **3**, 557–565.
- [48] Ellis, D. I., Dunn, W. B., Griffin, J. L., Allwood, J. W., Goodacre, R., Metabolic fingerprinting as a diagnostic tool. *Pharmacogenomics* 2007, **8**, 1243–1266.
- [49] Villas-Bôas, S. G., Mas, S., Åkesson, M., Smedsgaard, J., Nielsen, J., Mass spectrometry in metabolome analysis. *Mass Spectrom. Rev.* 2005, **24**, 613–646.
- [50] Smedsgaard, J., Nielsen, J., Metabolite profiling of fungi and yeast: From phenotype to metabolome by MS and informatics. *J. Exp. Bot.* 2005, **56**, 273–286.
- [51] Kouskoumvekaki, I., Yang, Z., Jonsdottir, S., Olsson, L., Panagiotou, G., Identification of biomarkers for genotyping Aspergilli using non-linear methods for clustering and classification. *BMC Bioinf.* 2008, **9**, 59.
- [52] Nielsen, J., It is all about metabolic fluxes. *J. Bacteriol.* 2003, **185**, 7031–7035.
- [53] Sauer, U., Hatzimanikatis, V., Bailey, J. E., Hochuli, M. et al., Metabolic fluxes in riboflavin-producing *Bacillus subtilis*. *Nat. Biotechnol.* 1997, **15**, 448–452.
- [54] Marx, A., de Graaf, A. A., Wiechert, W., Eggeling, L., Sahm, H., Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by nuclear magnetic resonance spectroscopy combined with metabolite balancing. *Biotechnol. Bioeng.* 1996, **49**, 111–129.
- [55] Gombert, A. K., Moreira dos Santos, M., Christensen, B., Nielsen, J., Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J. Bacteriol.* 2001, **183**, 1441–1451.
- [56] Klapa, M. I., Quackenbush, J., The quest for the mechanisms of life. *Biotechnol. Bioeng.* 2003, **84**, 739–742.
- [57] Sauer, U., Metabolic networks in motion: <sup>13</sup>C-based flux analysis. *Mol. Syst. Biol.* 2006, **2**.
- [58] Dinamarco, T. M., Almeida, R. S., de Castro, P. A., Brown, N. A. et al., Molecular characterization of the putative transcription factor SebA involved in virulence in *Aspergillus fumigatus*. *Eukaryot. Cell* 2012, **4**, 518–531.
- [59] Pecyna, M., Bizukojc, M., Lovastatin biosynthesis by *Aspergillus terreus* with the simultaneous use of lactose and glycerol in a discontinuous fed-batch culture. *J. Biotechnol.* 2010, **151**, 77–86.
- [60] Vongsangnak, W., Hansen, K., Nielsen, J., Integrated analysis of the global transcriptional response to α-amylase over-production by *Aspergillus oryzae*. *Biotechnol. Bioeng.* 2011, **108**, 1130–1139.
- [61] de Jongh, W. A., Nielsen, J., Enhanced citrate production through gene insertion in *Aspergillus niger*. *Metab. Eng.* 2008, **10**, 87–96.
- [62] Yuan, X.-L., van der Kaaij, R., van den Hondel, C., Punt, P. et al., *Aspergillus niger* genome-wide analysis reveals a large number of novel alpha-glucan acting enzymes with unexpected expression profiles. *Mol. Genet. Genomics* 2008, **279**, 545–561.
- [63] Battat, E., Peleg, Y., Bercovitz, A., Rokem, J. S., Goldberg, I., Optimization of L-malic acid production by *Aspergillus flavus* in a stirred fermentor. *Biotechnol. Bioeng.* 1991, **37**, 1108–1116.
- [64] Shu, P., Johnson, M. J., Citric acid production by submerged fermentation with *Aspergillus niger*. *Ind. Eng. Chem.* 1948, **40**, 1202–1205.

- [65] Andersen, M., Lehmann, L., Nielsen, J., Systemic analysis of the response of *Aspergillus niger* to ambient pH. *Genome Biol.* 2009, **10**, R47.
- [66] Carlsen, M. C., Nielsen, J. N., Influence of carbon source on α-amylase production by *Aspergillus oryzae*. *Appl. Microbiol. Biotechnol.* 2001, **57**, 346–349.
- [67] Melzer, G., Esfandabadi, M., Franco-Lara, E., Wittmann, C., Flux design: In silico design of cell factories based on correlation of pathway fluxes to desired properties. *BMC Syst. Biol.* 2009, **3**, 120.
- [68] Driouch, H., Melzer, G., Wittmann, C., Integration of in vivo and in silico metabolic fluxes for improvement of recombinant protein production. *Metab. Eng.* 2012, **14**, 47–58.
- [69] Bok, J. W., Hoffmeister, D., Maggio-Hall, L. A., Murillo, R., et al., Genomic mining for *Aspergillus* natural products. *Chem. Biol.* 2006, **13**, 31–37.
- [70] Olarte, R. A., Horn, B. W., Dorner, J. W., Monacell, J. T., et al., Effect of sexual recombination on population diversity in aflatoxin production by *Aspergillus flavus* and evidence for cryptic heterokaryosis. *Mol. Ecol.* 2012, **21**, 1453–1476.
- [71] Måansson, M., Klejnstrup, M. L., Phipps, R. K., Nielsen, K. F. et al., Isolation and NMR characterization of fumonisin B2 and a new fumonisin B6 from *Aspergillus niger*. *J. Agric. Food Chem.* 2009, **58**, 949–953.
- [72] Sorensen, L., Lametsch, R., Andersen, M., Nielsen, P., Frisvad, J., Proteome analysis of *Aspergillus niger*: Lactate added in starch-containing medium can increase production of the mycotoxin fumonisin B2 by modifying acetyl-CoA metabolism. *BMC Microbiol.* 2009, **9**, 255.
- [73] Wattanachaisaereekul, S., Lantz, A. E., Nielsen, M. L., Nielsen, J., Production of the polyketide 6-MSA in yeast engineered for increased malonyl-CoA supply. *Metab. Eng.* 2008, **10**, 246–254.
- [74] Panagiotou, G., Andersen, M. R., Grotkjaer, T., Regueira, T. B., et al., Studies of the production of fungal polyketides in *Aspergillus nidulans* by using systems biology tools. *Appl. Environ. Microbiol.* 2009, **75**, 2212–2220.
- [75] Galagan, J. E., Calvo, S. E., Cuomo, C., Ma, L.-J. et al., Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* 2005, **438**, 1105–1115.
- [76] Pel, H. J., de Winde, J. H., Archer, D. B., Dyer, P. S. et al., Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nat. Biotechnol.* 2007, **25**, 221–231.
- [77] Baker, S. E., *Aspergillus niger* genomics: Past, present and into the future. *Med. Mycol.* 2006, **44**, 17–21.
- [78] Machida, M., Asai, K., Sano, M., Tanaka, T. et al., Genome sequencing and analysis of *Aspergillus oryzae*. *Nature* 2005, **438**, 1157–1161.
- [79] Nierman, W. C., Pain, A., Anderson, M. J., Wortman, J. R. et al., Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature* 2005, **438**, 1151–1156.
- [80] Panagiotou, G., Andersen, M. R., Grotkjaer, T., Regueira, T. B., et al., Systems analysis unfolds the relationship between the phosphotolase pathway and growth in *Aspergillus nidulans*. *PLoS ONE* 2008, **3**, e3847.
- [81] Udagawa, H., Taira, R., Takagi, S., Novozymes AS 2011, Polynucleotides having leader sequence function. Patent WO2011161206.
- [82] Nielsen, J., De Jongh, W., Univ Danmarks Tekniske 2007, Enhanced citrate production. Patent WO2007143999.
- [83] Brown, S., Luttringer, S., Yaver, D., Berry, A., Novozymes Inc 2011, Methods for improving malic acid production in filamentous fungi. Patent US20110053233.
- [84] Borosova, G., Gajdosikova, J., Vajcikova, A., Vollek, V., BIOTIKA AS 1999, A method of fermentation preparation of lovastatin by an *Aspergillus terreus* strain and the strain for performing the method. Patent EP1047768.