2018

Metabolic Engineering of Serratia marcescens

Qiang Yan
Virginia Commonwealth University

Metabolic engineering of *Serratia marcescens*

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Phylosophy at Virginia Commonwealth University.

by

Qiang Yan

Master of Engineering in Fermentation Engineering, Jiangnan University, China

2011-2014

Bachelor of Engineering in Biotechnology, Dalian Polytechnic University, China

2007-2011

Director: Stephen S Fong, Ph.D.

Associate professor of Chemical and Life Science Engineering

Virginia Commonwealth University

May, 2018

Acknowledgements

The author wishes to thank several people. I would like to thank my wife, Zhou Ye, for her love, support and patience even though we were pursuing our Ph.D. degrees in two different contries. I would like to thank my parents for their selfless love and support. I also would like to thank my advisor Dr. Fong for his help, direction of this project and fundings. I would like to thank Dr. Allison A Johnson, Dr. Christina Tang, Dr. Fernando A Tenjo and Dr. Michael Peters for serving in my dissertation committee. I would like to thank Dr. J Paul Brooks for teaching me how to run the model algorithm. I would like to thank Dr. Seth Roberts for help construct the model *i*SR929. I would like to thank our lab aluminus Dr. Adam Fisher and Dr. George H McArthur IV for sharing their perspectives of the field of metabolic engineering and synthetic biology. I would like to thank my collegues: Dylan Rodenee, Andrew Harrison and Shani Levit.

Table of Contents

List of Figures

List of Table

Abstract

METABOLIC ENGINEERING OF *SERRATIA MARCESCENS*

By Qiang Yan

Virginia Commonwealth University, 2018

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2018

Major Director: Stephen S Fong, Ph.D.

Associate Professor, Chemical and Life Science Engineering

Chitin is the world's second most abundant polymer and contains in shrimp/crab shells as food waste. The potential value of the chitin biomass (e.g. food waste) is recently considered being ignored by landfill. Chitin can be a potential cheap carbon source for converting into value-added chemicals by microorganisms. Currently, one of the hurdles hindering utilizing chitin-based biomass feedstocks is the recalcitrant nature of chitin to hydrolyze. Wide ranges of chitinase enzymes are found natively in microorganisms that can potentially be used to effectively hydrolyze chitin to fermentable sugars. *Serratia marcescens* is a chitinolytic bacterium that harbors endogenous chitinase systems. Although *S. marcescens* chitinases are being characterized, studies regarding *S. marcescens* utilizing chitin as a substrate to produce chemical remain uncharacterized.

With goals of characterzing *S. marcescens* chitinolytic capabilities and applying *S. marcescens* to chemical production from chitin, my dissertation main content includes

five chapters: 1) Chapter 1 highlights background information of chitin source, *S. marcescens* and potential metabolic engineering targets using chitin as a substrate; 2) Chapter 2 demonstrates that ChiR is a key regulator in regulating 9 chitinase-related genes in *S. marcescens* Db11 and manipulation of *chiR* can be a useful and efficient genetic target to enhance chitin utilization; 3) Chapter 3 reports the production of *N*-acetylneuraminic acid (Neu5Ac) from chitin by a bottom-up approach of engineering the nonconventional chitinolytic bacterium, *Serratia marcescens*, including native constitutive promoter characterization and transcriptional and translational pathway balancing; 4) Chapter 4 describes improvement of *S. marcescens* chitinolytic capability by an adaptive evolution approach; 5) Chapter 5 elucidates *S. marcescens* intracellular metabolite profile using a constraint-based genome-scale metabolic model (*i*SR929) based on genomic annotation of *S. marcescens* Db11. Overall, the dissertation work is the first report of demonstrating the concept of chitin-based CBP using *S. marcescens* and the computational model and genetic molecular tools developed in this dissertation are valuable but not limited to design-build-test of *S. marcescens* for contributing to the field of biological science and metabolic engineering applications.

Abbreviations

atomic force microscopy; AFM; acylated homoserine lactones, AHL; α-acetolactate synthase, ALS; α-acetolactate decarboxylase, ALDC; chitin binding protein, CBP; consolidated bioprocessing, CBP; dry cell weight, DCW; flux balance analysis, FBA; fibronectin III domain, FnIII domain; green fluorescent protein, GFP; glycoside hydrolase, GH; gene-protein-reaction, GPR; genome-scale metabolic model, GSMM; hyaluronan, HA; LysR-type transcriptional regulator, LTTR; *N*-acetylglucosamine, GlcNAc; *N*-acetylneuraminic acid, Neu5Ac; *N*-acetylmannosamine, ManNAc; optical density; OD; open reading frame, ORF; phosphoenolpyruvate, PEP; quorum sensing, QS; RNA sequencing, RNA-seq; ribosome binding site, RBS; reads per kilobase per million mapped sequence reads, RPKM; *S. marcescens* ChiR overexpression strain, *S. marcescens* chiROE; *S. marcescens chiR* deletion strain, *S. marcescens ΔchiR*; *S. marcescens* end-point strain, *S. marcescens* EPS; *S. marcescens* middle-point strain, S. marcescens MPS; single-nucleotide variations, SNVs; simultaneous saccharidifcation, SSF; twin-arginine translocation, tat; 2-deoxystreptamine, 2-DOS; 2,3-butanediol, 2,3-BD; 2,3-butanediol dehydrogenase, BDH

Chapter 1 Introduction and Background

1.1 Introduction

Chemicals produced through bioprocesses range from bulk commodity chemicals such as organic acids and biofuels to high-value specialty chemicals such as isoprenoid drugs. For specialty chemicals with high monetary value per unit, the emphasis has been discovery and production feasibility. For other chemicals where economic pressures constrain industrial scale processes, one area of focus has been on utilizing cheap, abundant starting materials (Hasunuma et al. 2013; Mosier et al. 2005).

Bio-based degradation of lignocellulosic biomass (cellulose, hemicelluloses and lignin) was a logical focus of initial research efforts as cellulose is the most abundant available carbon substrate and potential mechanisms for breaking down lignocellulosic material should be identifiable from cellulolytic organisms. As the world's second most abundant carbon substrate, chitin has many similarities compared to cellulose: highly abundant in nature (10^{10}-10^{11} tons $year^{-1}$), homopolymer of a simple sugar (*N*-acetylglucosamine), requiring specialized enzymes for degradation (chitinase). Due to their similarity, implementation of chitin-based biomass as carbon source for biochemical production can enhance the nature of the biorefinery industry including broadening the range of practicable biomass conversion and lowering the environmental impact of our chemical-based

industry.

Aiming at developing chemical production bioprocess using chitin as a carbon source, this chapter mainly focuses on three aspects of background information: 1) chitin structure and its natural sources, 2) chitinases for degrading chitin into fermentable sugar, 3) potential chemical targets based on the chitin monomer structure similarities.

1.2 Chitin-based biomass resources

As the world's second most abundant polymer, chitin/chitosan occurring as a main component in seafood wastes (i.e., shrimp, crab, lobster shells) (Yan and Fong 2015). Annually, such organic marine waste pose a potential issue to the world and society: disposal has an associated high capital cost (e.g., $150/ton in Australia while dried shrimp cost $100–120/ton and estimated 1.5 million tons in Southeast Asia alone) (Dahiya et al. 2006). The potential value of the shell waste for the chemical industry is being ignored.

1.3 Chitin structure

Chitin, poly(β-(1→4)-*N*-acetyl-D-glucosamine(GlcNAc)), consists of *N*-acetylglucosamine by β-(1→4) glycosidic bond, shown in **Figure 1**. In nature, there are two different crystalline forms, α and β. α-chitin is by far the most abundant and the linear chains of GlcNAc unit are arranged in an antiparallel manner. On the other hand, β-chitin consists of parallel chains (Rinaudo 2006).

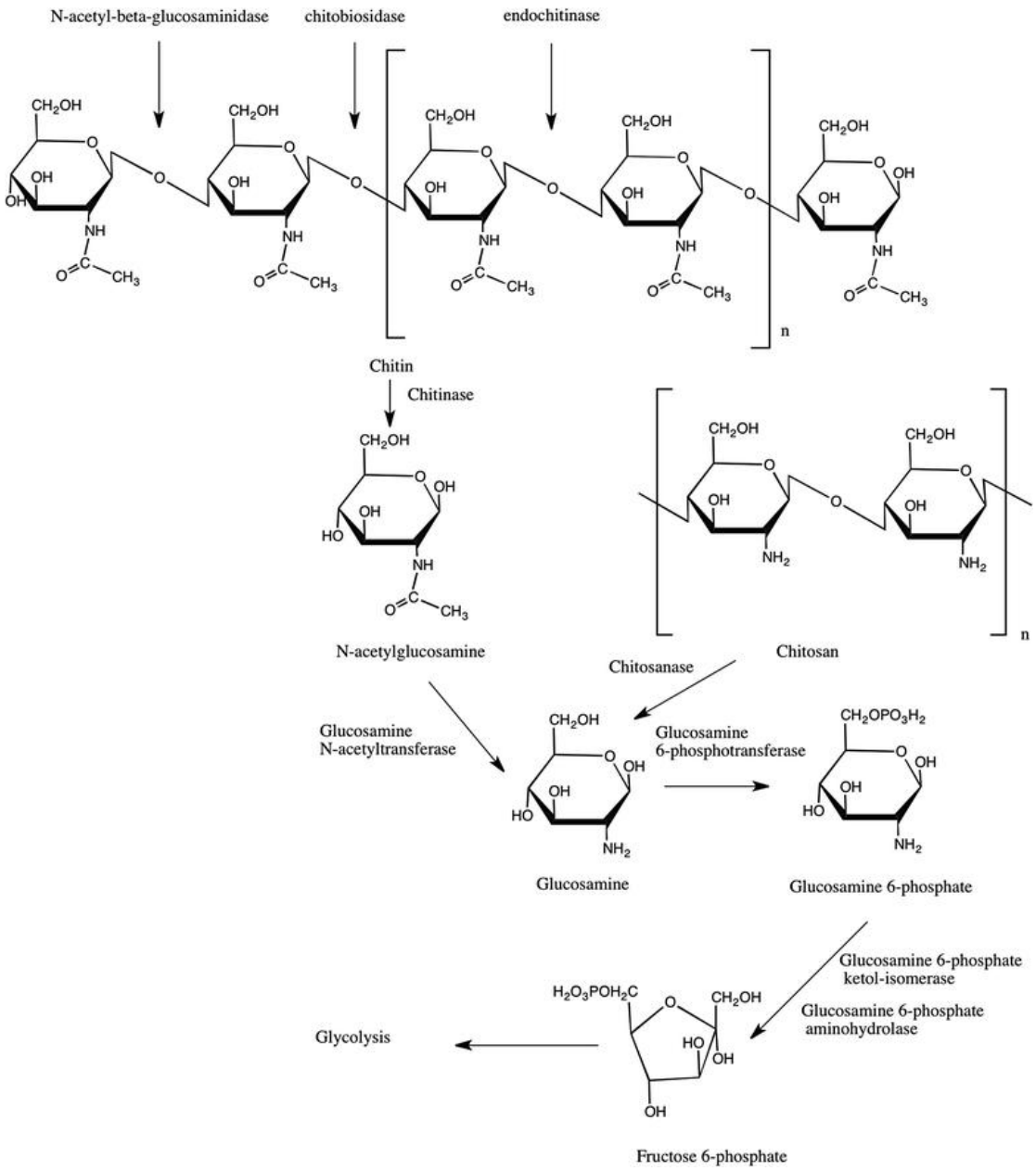


Figure 1 Structural of chitin, chitosan, and degradation products with specific enzymes and reaction.

1.4 Chitinase resources, structures, and categories

1.4.1 Chitinase resources

Chitinases are enzymes that can degrade/break down the long chain chitin into chitooligosaccharides or the monomer (e.g. *N*-acetylglucosamine). Chitinases are encoded in a wide range of organisms including viruses, bacteria, fungi, insects,

higher plants, and animals for different purposes such as nutrition, morphogenesis, and defense against chitin-containing pathogens (Aam et al. 2010; Adrangi and Faramarzi 2013; Harti et al. 2012).

1.4.2 Chitinase structures

Chitinolytic bacteria generally produce multiple chitinases derived from different genes. Many chitinolytic bacteria produce only family 18 chitinases, whereas other bacteria such as the *Streptomyces* species produce family 19 chitinases in addition to family 18 chitinases (Dahiya et al. 2006).

1.4.2.1 Catalytic domain

Chitinases catalytic domains have a $(\beta/\alpha)_8$ TIM-barrel fold with crucial catalytic residues being located on β-strand number 4 (Vaaje-Kolstad et al. 2013). High-resolution 3D structures of bacterial chitinases from *B. circulans* (ChiA1) have been revealed by X-ray crystallography (Toratani et al. 2006). The catalytic domain of ChiA1 consists of a deep substrate-binding cleft on the top of its $(\beta/\alpha)_8$-barrel structure. A mechanism has been suggested by a hydrophobic stacking interaction between the aromatic residues in this cleft and bound oligosaccharide (Watanabe et al. 2002; Watanabe et al. 2001).

1.4.2.2 Chitin-binding domain

Chitinases are usually multi-modular. For example, ChiA1 from *B. circulans* WL-12 contains a C-terminal chitin-binding domain and two fibronectin III-like domains (FnIII-D-1 and FnIII-D-2) (Watanabe et al. 1990); *S. marcescens* contains an N-terminal chitin-binding domain with a fibronectin III domain (ChiA), or a

C-terminal chitin-binding domain (ChiB), or a C-terminal FnIII domain coupled with a downstream chitin-binding domain (ChiC) (Vaaje-Kolstad et al. 2013). The properties and roles of the chitin-binding domain have been addressed in several studies but remain partly unresolved. The typical mechanisms of a chitin-binding domain function to facilitate correct positioning of the catalytic domain, to contribute to processive action, and to facilitate local decrystallization of the substrate. Several studies have confirmed that the presence of these domains increases substrate affinity as well as the efficiency of chitin hydrolysis, especially for more crystalline chitin.

1.4.2.3 Fibronectin III (FnIII) domains

The FnIII domain is one of the most commonly found motifs in animal proteins, and it has been proposed that bacterial FnIII domains were acquired from animals by horizontal gene transfer. This domain shares barely detectable sequence identity with FnIII modules between ChiA from *S. marcescens* and ChiA1 from *B. circulans* (Jee et al. 2002; Watanabe et al. 1994). For example, the FnIII domain from *S. marcescens* ChiA consists of aromatic residues that contribute to substrate binding and substrate hydrolysis (Uchihashi et al. 2001). In contrast to the N-terminal domain of ChiA from *S. marcescens*, the structurally characterized FnIII domain of ChiA1 from *B. circulans* does not have exposed aromatic residues on its surface (Jee et al. 2002). Therefore, this domain does not appear to be directly involved in chitin binding. However, it is proposed that FnIII domains are important for enzyme activity, probably

by affecting the overall enzyme structure and the spatial localization of the catalytic domain and the chitin-binding domain.

1.5 A model chitinolytic organism: *Serratia marcescens*

Many organisms have been isolated and reported with capability of producing/secreting chitinase. *Serratia marcescens* is a Gram-negative, chitinolytic bacterium in the *Enterobacteriaceae* order. *S. marcescens* is known as an efficient chitin degrader (Monreal and Reese 1969; Vaaje-Kolstad et al. 2013). The wild-type Db11 strain of *S. marcescens* has 10 chitinase-related genes, of which 5 different chitinases have been identified, purified and characterized. One of these chitinases is an endo-acting non-processive chitinase (*chiC*) (Horn et al. 2006; Suzuki et al. 1999; Synstad et al. 2008); two are processive exochitinases (*chiA* and *chiB*) that move along chitin in opposite directions (e.g. reducing and non-reducing ends) (Brurberg et al. 1995; Brurberg et al. 1994; Hult et al. 2005; Igarashi et al. 2014); one is a novel surface-active CBM33-type lytic polysaccharide monooxygenase (*cbp21*) that introduces chain breaks by oxidative cleavage (Suzuki et al. 1998; Vaaje-Kolstad et al. 2005; Vaaje-Kolstad et al. 2010; Watanabe et al. 1997); and one chitobiase (*chb*) (Kless et al. 1989; Tews et al. 1996) that converts the oligomeric products from the other enzymes to monomeric *N*-acetylglucosamine.

1.6 Chitin-based consolidated bioprocessing

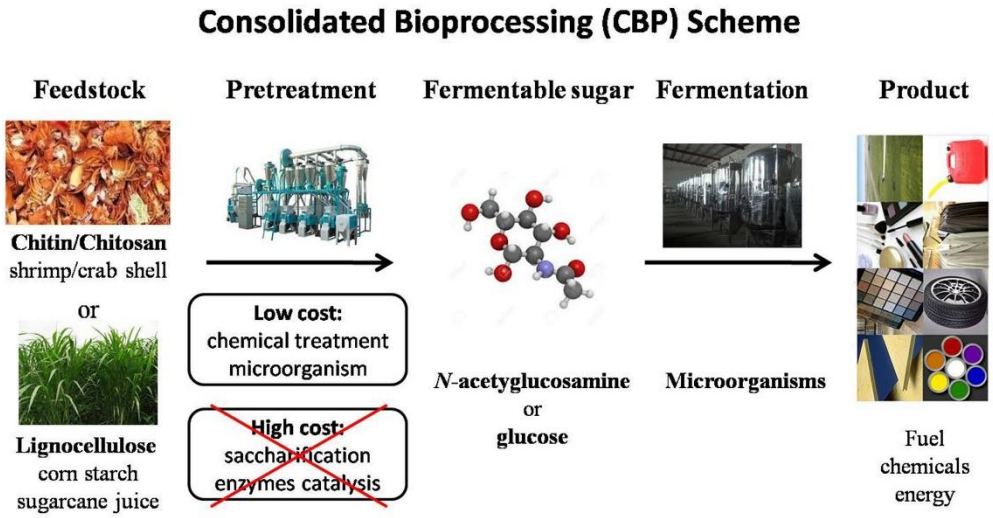


Figure 2 Scheme of consolidated bioprocessing (CBP) from biorenewable feedstock biomass.

Chitin-based consolidated bioprocessing (CBP) aims at one-step conversion of the naturally renewable biomass (i.e. shrimp/crab shells) into value-added chemical without protein purification processes. The CBP approach is mainly motivated by three benefits (**Figure 2**). First, it is a sustainable green approach that can significantly reduce greenhouse gas emissions (e.g. bioprocesses vs chemical processes). Second, many renewable feedstocks (shrimp/crab shells) are readily available, inexpensive resources that can lower material costs. Third, CBP is able to eliminate labor and capital cost of biomass processing by employing a single process step. Indeed, CBP is widely recognized as the ideal configuration for sustainable, low-cost hydrolysis and fermentation of cellulosic biomass. In principle, a CBP strategy can be applied to produce a broad range of chemicals from natural biomass. It requires degrading recalcitrant biomass substrates into solubilized sugars and metabolic intervention to direct metabolic flux toward desired products at high yield and titer (Yan and Fong 2015; Yan and Fong 2017a). With recent advances of

synthetic biology and systems biology, development of a non-model organism as a platform workhorse can be rationally designed and efficiently deployed (Yan and Fong 2016; Yan and Fong 2017a).

1.7 Potential metabolic engineering targets

Just as cellulose is viewed as a renewable starting carbon source for a variety of biochemical carbon sources, a variety of potential biochemical products can be made based on the monomers of chitin degradation, shown in **Figure 3**.

1.7.1 Glycan biosynthesis

In addition to being a structural component of homogeneous polysaccharides like chitin, GlcNAc is also a constituent of heterogeneous polysaccharides, such as murein and hyaluronic acid (also called hyaluronan or hyaluronate, HA). Murein is the basic component of the bacterial cell wall and consists of crosslinked peptide chains with repeating GlcNAc and muramic acid residues (Ashry and Aly 2007; Chen et al. 2010).

HA is a linear heteropolysaccharide that is composed of repeating D-glucuronic acid and GlcNAc residues. HA is a major component of extracellular matrix and is extensively distributed in connective, epithelial and neutral tissues (Chien and Lee 2007). HA performs numerous roles in cell motility, inflammation and cancer metastasis (Danishefsky et al. 1969).

Heparin, which is produced by basophils and mast cells, consists of a variably-sulfated repeating disaccharide unit and acts as an anticoagulant. The major

repeating disaccharide unit of keratin sulfate, also called keratosulfate, is composed of galactose and sulfated GlcNAc. Keratin sulfate is distributed in the cornea, cartilage and bone and usually acts as a cushion in joints to absorb mechanical shock (Tsai et al. 2001).

Sialyl-lewis is a tetrasaccharide carbohydrate that is usually attached to O-glycans on the surface of cells. It is known to play a vital role in cell-to-cell recognition processes. It is also the means by which an egg attracts sperm-first to stick to it, then bond with it and eventually form a fetus. There are two types of sialyl-lewis, X and A. Sialyl-lewis X is [NeuAcα2-3Galβ1-4(Fuca1-3)GlcNAc]. Defective synthesis of the sialyl-lewis X antigen results in immunodeficiency (leukocyte adhesion deficiency type 2 (Polley et al. 1991).

1.7.2 Acetamido sugars

GlcNAc is the precursor of modified acetamido sugars, such as *N*-acetylgalactosamine, sialic acid, *N*-acetylmuramic acid, *N*-acetylmannosaminuronate and several 6-deoxy-2-acetamido sugars, including bacillosamine and pseudaminic acid, which are core components of bacterial *N*-linked and *O*-linked glycans, respectively (Piacente et al. 2013).

N-acetylneuraminic acid (Neu5Ac) is the most common sialic acid and exists as > 40 structural derivatives in mammalian and avian species. Sialic acids are found at the end of sugar chains connected to the surfaces of cells and proteins, serving as

receptors for influenza viruses mediated endocytosis. The guanidyl derivative of Neu5Ac, Zanamivir, has been chemically synthesized as novel pharmaceutical agents to produce neuraminidase inhibitors are used to treat influenza infections.

Because the Neu5Ac content in natural products is too low for the isolation of Neu5Ac with sufficient recovery and purity, novel processes had to be developed for the industrial production of this compound (Vimr and Lichtenteiger 2002). Enzymatic synthesis of Neu5Ac from *N*-acetylmannosamine (ManNAc) and pyruvate using Neu5Ac aldolase as a catalyst has been reported (Kang et al. 2012).

1.7.3 Antibiotics biosynthesis

Aminoglycoside antibiotics have been widely used clinically since the first use of streptomycin as an effective antibiotic in the treatment of tuberculosis. These antibiotics consist of a central aminocyclitol ring, such as streptamine, streptidine, or 2-deoxystreptamine (2-DOS) (Song et al. 2013).

Neomycin is a 4,5-disubstituted aminoglucoside antibiotic. It is typically used as a topical preparation, such as Neosporin. Because neomycin is not absorbed from the gastrointestinal tract and has been used as a preventive measure for hepatic encephalopathy. By killing bacteria in the intestinal tract, it keeps ammonia levels low and prevents hepatic encephalopathy, especially prior to GI surgery. The first glycosylation step of neomycin involved in the formation of pseudodisaccharides is the addition of *N*-acetylglucosamine to 2-DOS. NeoM from the neomycin gene cluster

was first characterized as the UDP-*N*-acetylglucosamineglycosyltransferase.

Butirosin is a relatively new aminoglycosidic antibiotic complex active *in vitro* and *in vivo* against various pathogenic gram-positive and gram-negative bacteria. Butirosin includes two isomers, butirosin A and butirosin B, which differ only in the configuration at one carbon atom in the pentose moiety.

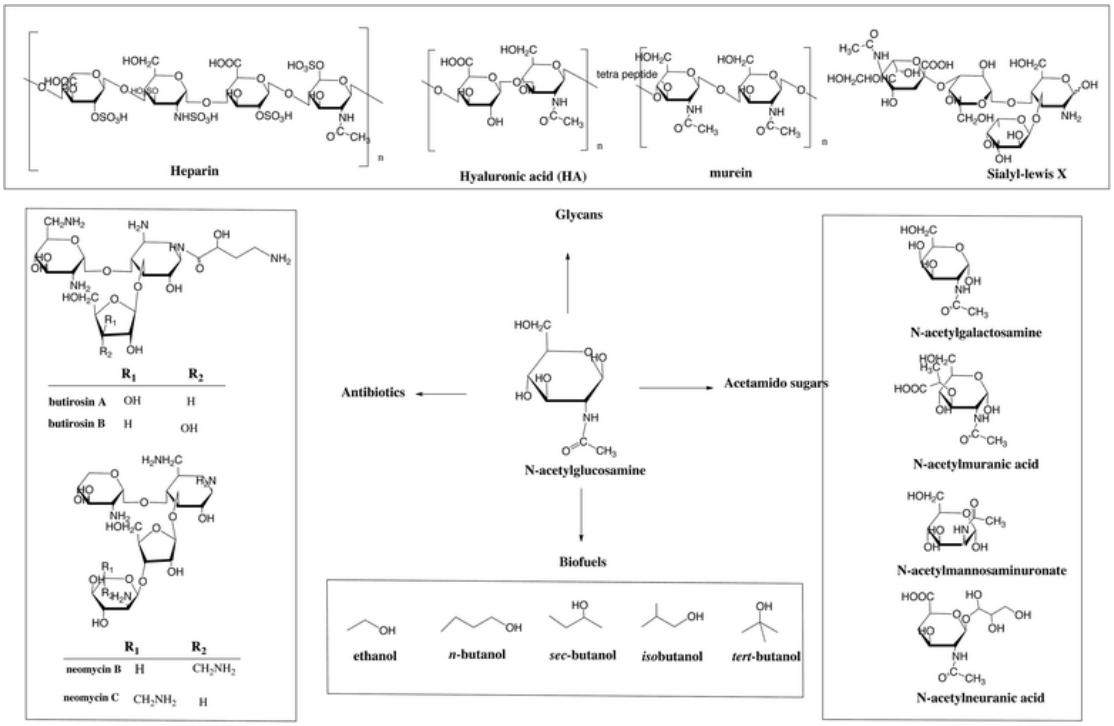


Figure 3 Potential metabolic targets from *N*-acetylglucosamine as a carbon source.

Chapter 2 Study of ChiR function in *Serratia marcescens* and its application for improving 2,3-butanediol from crystal chitin

Importance of this part of work:

Since *S. marcescens* chitinases are highly active, they can be potentially used in the chemical production industry. In *S. marcescens*, ChiR is the only chitinase regulatory protein among 10-chitinase related proteins. Although one study partially reported that ChiR is essential for chitinase production, it is still unclear how ChiR regulates the other 9-chitinase genes transcription. Before my dissertation study there was no studies reporting about *S. marcescens* producing 2,3-butanediol from chitin.

The conclusions and fundamental results derived from this part of work:

First, a *chiR* overexpression (*chiROE*) strain and a *chiR* deletion (*ΔchiR*) strain were generated and characterized in terms of cellular growth, chitinase activity, and total secreted protein. Compared to the wild-type Db11 strain, the *S. marcescens chiROE* strain showed an increase in chitinase activity (2.14- to 6.31-fold increase). Increased transcriptional expression of chitinase-related genes was measured using real-time PCR, showing 2.12- to 10.93-fold increases. The *S. marcescens ΔchiR* strain showed decreases in chitinase activity (4.5- to 25-fold decrease), confirming ChiR's role as a positive regulator of chitinase expression. Finally, *chiR* overexpression was investigated as a means of increasing biochemical production (2,3-butanediol) from crystal chitin. The *chiROE* strain produced 1.13 ± 0.08 g/L 2,3-butanediol from 2% crystal chitin, a 2.83-fold improvement from the wild-type strain, indicating ChiR is

an important and useful genetic engineering target for enhancing chitin utilization in *S. marcescens*.

2.1 Introduction

S. marcescens is recognized as an efficient and promising native producer of 2,3-butanediol (2,3-BD) (Ji et al. 2011). Recently, a high industry-level titer of meso-2,3-BD (152.0 g/L) was produced from *S. marcescens* by a fed-batch fermentation using sucrose as a substrate (Zhang et al. 2010a; Zhang et al. 2010b). In *S. marcescens*, the 2,3-BD synthetic pathway mainly involves three key enzymes, α-acetolactate synthase (ALS), α-acetolactate decarboxylase (ALDC), and 2,3-butanediol dehydrogenase (BDH) (Zhang et al. 2016). Two molecules of pyruvate condense to yield one molecule of α-acetolactate and release one molecule of CO_2 by ALS (BudA, SMDB11_2829), then α-acetolactate is decarboxylated into (3*R*)-acetoin by ALDC (BudB, SMDB11_2830). Under aerobic conditions, α-acetolactate readily undergoes non-enzymatic oxidative decarboxylation and forms diacetyl. Finally, (3*R*)-acetoin and diacetyl are reduced into 2,3-BD by BDH (SMDB11_1336).

ChiR (SMDB11_2876) has previously been identified as a regulator of several chitinase genes and is a member of LysR-type transcriptional regulators (LTTRs) (Suzuki et al. 2001). LTTRs are thought to contact the α-subunit of RNA polymerase by binding to the characteristic sequence T-N_{11}-A conserved upstream of the regulated promoter, which is usually an imperfect palindromic sequence (Burn et al. 1989; Gibson and Silhavy 1999; Maddocks and Oyston 2008). In an initial study of ChiR, *S.*

marcescens mutant strains (N1-N5) were generated by Tn5 insertion mutagenesis and did not form any clearing zone on YEM agar containing colloidal chitin, while other extracellular enzymes besides chitinases (e.g. proteases, lipases, nucleases) appeared to be produced normally, indicating the loss of the ability to form clearing zones on colloidal chitin is due to defects for chitinase production (Watanabe et al. 1997). Further genetic analysis of the Tn5 insertion mutants identified disruption of the ChiR open reading frame (ORF). Furthermore, the authors generated a Tn5 transposon mutant leading to an inactive *chiR* gene, and the mutant did not observe any chitinase activity. *In vitro* protein binding studies were conducted using gel mobility shift assays for 4 chitinase genes (*chiA*, *chiB*, *chiC* and *cbp21*) resulting in the determination that ChiR specifically binds to the intergenic region between *chiR* and *cbp21* (Suzuki et al. 2001). Despite these results, several questions remain unclear related to ChiR function and chitin degradation in *S. marcescens*: 1) the function and relationship of ChiR to four potential chitinase-related genes (*SMDB11_1083*, *SMDB11_1994*, *SMDB11_4602* and *SMDB11_1190*) is still unknown; 2) the transcriptional effects of ChiR on chitinase-related gene expression have not been studied; 3) the effect of modifying ChiR expression on cellular function and chitin utilization have not been studied.

In this study, we developed an efficient genetic engineering strategy for genetic manipulation of *S. marcescens* with a goal of improving chitinolytic activity based on ChiR function. Our approach was to construct a *S. marcescens chiR* deletion strain

and a *S. marcescens chiR* overexpression strain to first explicitly study ChiR function in terms of growth phenotypes, chitinase activity, and chitinase gene expression. To evaluate ChiR function on chitin utilization and biochemical production, direct bioconversion of untreated chitin to 2,3-BD is demonstrated with highest production titers achieved in a *chiR* overexpression strain of *S. marcescens*.

2.2 Materials and methods

2.2.1 Strains and culture conditions

The *Serratia marcescens* Db11 was purchased from the Caenorhabditis Genetics Center (Twin City, USA http://www.cbs.umn.edu/CGC) (Flyg et al. 1980). All strains used in this study can be found with detailed information in **Table 1**. The *S. marcescens chiROE* strain harbors plasmid pQY38 containing pUC19 backbone with an insertion of *chiR* gene under P_{ampR} promoter and a synthetic ribosome binding site (RBS). The *S. marcescens ΔchiR* strain was constructed by replacement of *chiR* gene with the *nptII* gene in its chromosomal DNA. *E. coli* NEB10β (New England Biolabs, Ipswich, USA) was used as the host to propagate pUC19 and pUN plasmid, construct pQY45 plasmid and store plasmids. All *E. coli* strains were grown in LB or SOC medium at 37℃ and 250 rpm supplemented with 200 mg/L ampicillin or 40 mg/L kanamycin. In order to monitor growth phenotype changes, all the *S. marcescens* strains were grown in M9 medium with 0.1% yeast extract and 2% various carbon sources (glucose, *N*-acetylglucosamine, and colloidal chitin) at 30℃ and 250 rpm. Chitinase assay experiments were analyzed by growing *S. marcescens* strains in LB or LB with crystal chitin or colloidal chitin.

Table 1 Strains and plasmids used in this study.

Strain or plasmid	Relevant genotype or description	Source or reference(s)
Strains		
E. coli NEB10β	Δ*(ara-leu) 7697 araD139* (Str^R)	New England Biolab
S. marcescens Db11	Wild-type	Caenorhabditis Genetics Center
S. marcescens chiROE	*S. marcescens* Db11 harbors plasmid pQY38	This study
S. marcescens ΔchiR	*chiR::nptII*	This study
plasmids		
pUC19	amp^r,	New England Biolab
pQY38	pUC19 vector cloned with *chiR* gene under the promoter of ampR, amp^r	This study
pUN	*nptII* gene under promoter of ampR, kan^r	This study
pQY45	pUC19 vector cloned with *nptII* gene under the promoter of ampR flanking around 1 kb homologous arms at upstream and downstream, amp^r and kan^r	This study

2.2.2 Plasmids

Primers and DNA fragments sequence can be found in **Table S1** of Supplementary Materials. For construction *chiR* overexpression vector, pUC19 backbone was amplified by PCR using primers pUC19bb-f and pUC19bb-r. The *chiR* gene insertion sequence was designed, synthesized and purchased as a double-stranded DNA fragment from IDT (Coralville, USA). A synthetic RBS (5′-TCGGAATTAAAATAAATATTAAGGAAAATATAAAC-3′) was designed at an arbitrary translation initiation rate (TIR) number of 60,000 using RBS calculator (https://www.denovodna.com/software). The pQY38 plasmid was assembled from two above-mentioned DNA fragments using isothermal assembly (Gibson et al. 2009). For construction of *chiR* deletion plasmid (pQY45), pUC19 was used as backbone; the kanamycin resistance marker (*nptII*) of the pUN plasmid was amplified by PCR

using primers kanoverlap-f and kanoverlap-r; about 1 kb fragments upstream and downstream of *chiR* gene of *S. marcescens* genome was amplified by PCR using the two appropriate sets of primers (primer ha1-f/ha1-r and primer ha2-f/ha2-r). The four DNA fragments were assembled using isothermal assembly reaction, and then pQY45 was generated. Briefly, the isothermal assembly reaction was conducted in total volume 20 μL with 0.05 pmol pUC19 backbones, 0.05 pmol insertion gblock, and 10 μL Gibson assembly master mix. The reaction was incubated at 50 ℃ for 1 h. After the reaction, 4 μL of the mixture was transformed into *E. coli* NEB10β by electroporation, and the recombinant strains were screened by LB medium agar plate containing 200 μg/ml ampicillin.

2.2.3 Transformation of *S. marcescens* and allelic exchange

The method for chromosomal deletion in *S. marcescens* relies on a successful delivery of double-stranded DNA in *S. marcescens* and endogenous recombinases can facilitate homologous recombination, similar to a previous study (Deng and Fong, 2010). To delete *chiR* gene of *S. marcescens*, the *nptII* gene with about 1 kb flanking regions was amplified by PCR from pQY45 plasmid and introduced into *S. marcescens* by electroporation. Up to 5 μL of the double-stranded DNA fragment (about 1 μg) amplified after PCR was added to 50 μL of *S. marcescens* competent cell and mixed immediately by tapping the electroporation cuvette (Biorad, USA). The cuvette was electroporated at 25 kV/cm using gene pulser (Biorad, Hercules, USA). After electroporation, 1 mL SOC medium was added and the cell recovered by incubating at 30 ℃ for 1 h. The 100 μL of recovered *S. marcescens* suspension was plated onto LB

plates supplemented with 320 μg/mL kanamycin and incubated at 30℃ for two days. Kanamycin-resistant candidates were selected. Each candidate genomic DNA was isolated using a QIAamp® DNA mini kit (QIAGEN, USA) according to the manual instruction. Colony PCR was performed with primers kan-f and kan-r. The chromosomal disruption was validated by Sanger sequencing (Eurofins Genomics, Louisville, USA).

For *chiR* gene overexpression, pQY38 vector was transformed into *S. marcescens* by electroporation. The electroporation procedure was the same as mentioned above and ampicillin (> 800 μg/mL) resistant colonies were selected. Plasmids of positive transformants were extracted, and the *chiR* gene insertion was validated by sequencing using UNS7-f and UNS5-f primer (Eurofin genomics, Louisville, USA).

2.2.4 Preparation of colloidal chitin

Crystal chitin from shrimp shells was obtained from Sigma-Aldrich (C7170, USA). Colloidal chitin was prepared using the modified method of Roberts and Selitrennikoff as follows (Roberts and Selitrennikoff 1988). Initially, 5 g of chitin from shrimp shells (Sigma, C7170) was added slowly to 90 mL of 37% HCl (w/v). The mixture was vigorously stirred for 2 h. Then, 500 mL of 95% ethanol was added to this suspension and centrifuged at 6000 rpm for 20 min at 4℃. The pellet was washed with distilled water until chitin achieved a neutral pH and was stored frozen until use.

2.2.5 Cell density measurement

The culture density of *S. marcescens* strains was generally determined by measuring optical density at 600 nm on a Biomate3 UV/VIS spectrophotometer (Thermo, USA). The growth of *S. marcescens* on crystal chitin or colloidal chitin was determined by measuring cytoplasmic protein content by centrifuging 1 mL culture at 10,000 ×*g* for 5 min. Pellets were re-suspended in fresh media and centrifuged at 10,000 ×*g* for 5 min again. Sediments were dissolved and lysis in 200 μL 1 ×Bugbuster protein extraction reagent (EMD Millipore, USA). After 20 min incubation, the lysate was centrifuged at 10,000 ×*g* for 5 min, and the proteins in the supernatant were measured by the Bradford protein assay (Bradford 1976). Using a calibration curve to correlate the overall protein content with dry cell weight (DCW), the measurements were converted to dry weight; the correlation factor was DCW (g/L) = 1.20 ± 0.03 × Protein (g/L)

2.2.6 Chitinase activity assay

Chitinase activity assays were measured according to the manual instruction of a chitinase assay kit (Sigma-Aldrich, CS0980) (Tronsmo and Harman 1993). Endochitinase activity was detected using 0.2 mg/mL 4-Nitrophenyl β-D-*N*,*N*′,*N*″-triacetylchitotriose as a substrate; β-*N*-acetylglucosaminidase activity was determined using 0.5 mg/mL 4-Nitrophenyl *N*-acetyl-β-D-glucosaminide as a substrate; chitobiosidase activity was measured using 1 mg/ml 4-Nitrophenyl *N*,*N*′-diacetyl-β-D-chitobioside as a substrate. Specifically, 1 mL cell culture was centrifuged at 10,000 × *g* for 5 min. Each reaction contained 10 μL supernatant

samples and 90 μL substrate in chitinase assay buffer (Sigma-Aldrich, USA). The reaction was incubated in a 96-well plate at 37℃ for 30 min and was stopped by adding 200 μL 0.4 M Na_2CO_3 solution. The adsorption at 405 nm was conducted immediately to measure the release of 4-Nitrophenol. Chitinase activity was defined as one unit of activity will release 1.0 μmole of 4-Nitrophenol from the appropriate substrate per minute at pH 4.8 at 37℃.

2.2.7 RNA preparation and real-time PCR

In order to study transcriptional level changes in the culture of the three *S. marcescens* strains after overexpression of *chiR* gene and after *chiR* gene deletion, gene expression was measured using real-time PCR. All the *S. marcescens* strains were grown in M9 medium supplemented with 0.1% yeast extract and 2% chitin. All cells were harvested at the mid-log growth phase (OD_{600} 0.4 around 2.5 h). Primer sequences for real-time PCR experiments are available in **Table S2** of the Supplementary Material. The expression levels of 10 different chitinase-related genes were measured along with the level of expression of one housekeeping gene (*luxS, SMDB11_0167*) that was used as a control, and all transcript levels were normalized to the level of this housekeeping gene.

2.2.8 Prediction of chitinase-related gene promoter and its potential binding sites

Prediction of each chitinase-related gene promoter region was conducted based on each upstream intergenic region using online software BPROM (http://www.softberry.com/berry.phtml). The promoter region was predicted based upon bacterial sigma 70 RNAP recognition sites (Solovyev and Salamov 2011). Each

potential binding site region was manually found based on a part of palindromic T-N_{11}-A motif in each gene upstream intergenic region.

2.2.9 2,3-BD fermentation

S. marcescens pre-culture was performed in LB medium at 30 ℃ and 250 rpm for overnight. Then, 10% (v/v) seed culture was inoculated at 150 mL fermentation medium with a 250 mL Erlenmeyer flask at 30 ℃ with different initial pH value and shaking speed. The fermentation medium contains M9 minimum medium with 0.1% yeast extract and 20 g/L different carbon sources (glucose, *N*-acetyglucosamine or crystal chitin). Glucose and *N*-acetylglucosamine were used to investigate 2,3-BD production in *S. marcescens* as two carbon sources. In order to evaluate *S. marcescens* production of 2,3-BD from chitin, the fermentation was conducted under M9 medium with 1% yeast extract, 5 g/L *N*-acetylglucosamine and 20 g/L crystal chitin at pH 7.5, 100 rpm and 30 ℃.

2.2.10 Detection of 2,3-BD, sugar and byproducts

Sugar, 2,3-BD and other byproducts were detected using an HPLC system (Dionex Ultimate3000) equipped with Bio-Rad HPX-87H ion exclusion column. The mobile phase was 5 mM H_2SO_4 at the rate of 0.6 mL/min and RI, and UV detectors (wavelength at 199 nm) were used.

2.3 Results

2.3.1 Construction *S. marcescens chiR* deletion strain and *chiR* overexpression strain

In order to generate a *S. marcescens* chromosomal *chiR* deletion strain, a

PCR-targeted DNA fragment was introduced in *S. marcescens* competent cells to replace the chromosomal *chiR* gene with a kanamycin resistance gene. Native recombinases in *S. marcescens* were used to facilitate recombination. To address the problem of degradation of extracellular DNA by nucleases in *S. marcescens* (Benedik and Strych 1998; Friedholff et al. 1994; Miller et al. 1994), long nucleotide regions (> 1 kb) homologous to the target *chiR* gene were used. About 1000-bp fragments with homology upstream and downstream of the *chiR* gene were amplified from the *S. marcescens* Db11 genome. The kanamycin resistance gene under P_{ampR} promoter was amplified from pUN plasmid by high-fidelity PCR and used for replacement of the native gene and selection of transformants. The DNA fragments were assembled with the pUC19 backbone, generating pQY45 plasmid (**Figure S1A**), which was transformed into *E. coli* NEB10β. To disrupt *chiR*, the insertion fragment was amplified by PCR and transformed into *S. marcescens*. After transformation, cells were incubated on plates containing 320 μg/mL kanamycin for two days. There were no colonies found on the wild-type plates, however, 7 colonies were found on the *chiR* knockout plates with a transformation efficiency of ~5 transformants/μg DNA. Subsequently, 2 transformants were selected and grown on kanamycin LB liquid medium with both remaining resistant for 21 generations of serial passage, indicating a stable gene insertion/disruption occurred. These *chiR* deletion strains were confirmed by both PCR and DNA sequencing. PCR using genomic DNA of *S. marcescens chiR* deletion strain yielded a single 0.8-kb band, which corresponds to the expected size of a DNA fragment containing the *nptII* gene. In contrast, PCR

using genomic DNA of *S. marcescens* Db11 did not show a clear band using the same pair of primer (**Figure S1B**). These results indicated that *chiR* gene was successfully inactivated and kanamycin-resistance gene was swapped with the *chiR* region. DNA sequencing data (primer kan-f and kan-r) also confirmed that *chiR* gene of *S. marcescens* Db11 was successfully deleted.

To create a *S. marcescens chiR* overexpression strain, a shuttle vector, pQY38, was constructed by inserting a constitutive ampicillin promoter (P_{ampR}) and *chiR* gene into the pUC19 plasmid backbone (**Figure S1C**). After transformation with pQY38 (harbors ampicillin resistance), 105 colonies were found on an ampicillin plate yielding an efficiency of ~1×10^5 transformants/μg plasmid DNA. We grew one transformant on ampicillin for 100 generations by serial passage and it remained resistant to ampicillin, indicating a stable replication during cell propagation. Overall, we successfully generated a *S. marcescens chiR* overexpression strain (designated as *S. marcescens chiROE*) and a *S. marcescens chiR* deletion strain (designated as *S. marcescens ΔchiR*).

2.3.2 Growth phenotypes

In order to characterize growth phenotypic changes among the three *S. marcescens* strains (wild-type Db11, *ΔchiR*, and *chiROE*), cell growth was monitored using the minimum medium (M9) supplemented with various carbon sources such as *N*-acetylglucosamine, glucose, and colloidal chitin. *N*-acetylglucosamine and glucose were chosen as a carbon source because *S. marcescens* exhibits catabolite repression

on different carbon sources (Monreal and Reese 1969; Watanabe et al. 1997). For cell culture with *N*-acetylglucosamine and glucose, the *S. marcescens chiROE* exhibited a decreased cell yield, while the *S. marcescens ΔchiR* had an increased cell yield compared to *S. marcescens* Db11 (**Figure 4A**). For cell growth on colloidal chitin (**Figure 4B**), the *S. marcescens ΔchiR* strain exhibited severely inhibited growth and a maximum cell lysate protein of 0.21 g/L, while both the *S. marcescens* Db11 and the *S. marcescens chiROE* were able to grow on colloidal chitin with maximum overall cell lysate protein concentrations of 1.42 g/L and 1.25 g/L, respectively. Colloidal chitin creates an opaque suspension so protein concentration is used as a proxy for cell biomass since optical density cannot be used.

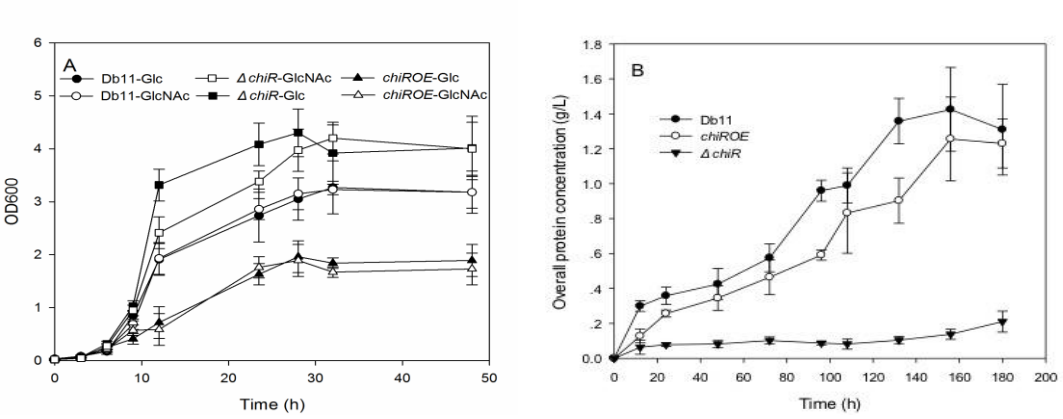


Figure 4 Time profile of cell growth of three *S. marcescens* strains. (A) Optical density OD_{600} of *S. marcescens* Db11 (circle), *S. marcescens* chiROE (triangle), and *S. marcescens ΔchiR* (square) on glucose (glc, filled) and *N*-acetylglucosamine (GlcNAc, unfilled) in a flask culture at 30℃ for 48 h; (B) Overall protein concentration of cell lysate of *S. marcecens* Db11 (filled circle), *S. marcescens* chiROE (open circle), and *S. marcescens ΔchiR* (inverted triangle) using colloidal chitin as sole carbon in a flask culture at 30℃ for 180 h.

2.3.3 Chitinase activity and overall secreted protein changes

In order to determine chitinolytic capability changes among the three *S. marcescens* strains, we measured secreted chitinase activity corresponding to the three different

chitinase mechanisms: endochitinase activity, chitobiosidase activity, and β-*N*-acetylglucosaminidase activity. Since *S. marcescens* chitinases are induced by the presence of crystal chitin or colloidal chitin, the three *S. marcescens* strains secreted chitinase activity was tested under three different culture conditions, plain LB medium and LB medium supplemented with either colloidal chitin or crystal chitin. For endochitinase activity, the *S. marcescens chiROE* strain showed increased activity compared to the wild-type Db11 strain (1.93-, 1.84-, and 2.23-fold increases when grown on crystal chitin, colloidal chitin or plain LB, respectively). The *S. marcescens ΔchiR* strain showed a significant reduction in endochitinase activity (22.7-, 14.3-, and 10.0-fold reduction on crystal chitin, colloidal chitin, and LB, respectively) compared to the wild-type Db11 (**Figure 5A**). For chitobiosidase activity, the *S. marcescens chiROE* showed increased activity compared to the *S. marcescens* Db11 with 1.40-, 1.39-, and 2.14-fold increases on crystal chitin, colloidal chitin or plain LB, respectively, while the *ΔchiR* strain exhibited decreased activity by 25.0-, 25.0- and 10.0-fold from the wild-type Db11 for growth on crystal chitin, colloidal chitin, and LB, respectively (**Figure 5B**). For β-*N*-acetylglucosaminidase activity, the *chiROE* strain showed increased activity (2.46-, 3.88-, and 6.31-fold increases on crystal chitin, colloidal chitin, or LB, respectively) compared to the wild-type Db11. The *S. marcescens ΔchiR* strain showed decreased β-*N*-acetylglucosaminidase activity (4.54-, 2.63-, and 3.45-fold decreases on crystal chitin, colloidal chitin, or LB, respectively) compared to the wild-type Db11 (**Figure 5C**). In addition, we observed an increase in the overall amount of secreted protein in the *S. marcescens chiROE* strain and a

decrease in overall secreted protein in the *S. marcescens ΔchiR* strain as compared to *S. marcescens* Db11 (**Figure 5D**). The total secreted protein concentration for the *S. marcescens chiROE* strain was 1.46-, 2.03- and 2.41-fold higher than the Db11 strain for growth on crystal chitin, colloidal chitin, or LB, respectively. Decreases of 5.88-, 3.13-, and 2.38-fold were observed in overall secreted protein under crystal chitin, colloidal chitin or LB, respectively for the *ΔchiR* strain when compared to the wild-type Db11. Overall, the *S. marcescens chiROE* strain showed increased chitinase activity while the *S. marcescens ΔchiR* strain showed decreased chitinase activity compared to *S. marcescens* Db11.

2.3.4 Regulation of chitinase gene expression by ChiR

In order to understand the role of ChiR in the regulation of chitinase genes in *S. marcescens*, 10 chitinase-related genes (including *chiR*) were examined by real-time PCR using the wild-type Db11 strain, the *ΔchiR* strain, and the *chiROE* strain (**Table 2**, **Table S5**, and **Table S6**).

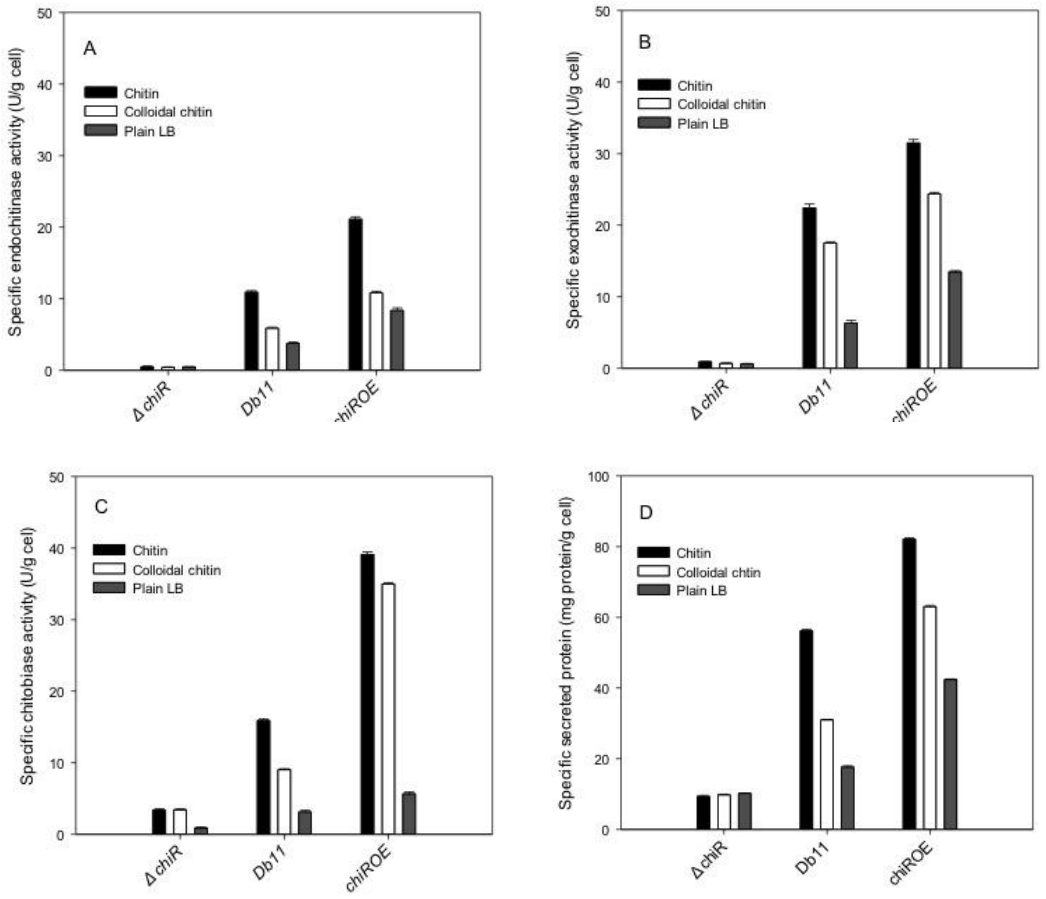


Figure 5 Specific chitinase activity and secreted protein of three *S. marcescens* strains: the *ΔchiR* strain, the wild-type Db11, and the chiROE strain. (A) specific endochitinase activity (U/g cell). (B) specific chitobiosidase activity (U/g cell). (C) Specific β-*N*-acetylglucosaminidase activity (U/g cell). (D) Specific protein secretion (mg protein/g cell) with chitin (black filled bar), colloidal chitin (unfilled bar), and plain LB (gray filled bar).

In the *chiROE* strain, *chiR* mRNA levels increased 593.59-fold as compared to the wild-type strain, as expected. Concurrently, the 9 chitinase-related genes showed significantly increased expression ranging from 2.12- to 10.93-fold increases (**Table 2**). This result showed that ChiR is an activator by upregulating all chitinase-related genes in *S. marcescens*. Five genes (*SMDB11_0468*: endochitinase, *SMDB11_2875*: exochitinase, *SMDB11_2877*: chitin binding protein, *SMDB11_1994*: chitinase, and

SMDB11_4243: exochitinase) showed greater than 5-fold increases in expression compared to the wild-type Db11. For the *ΔchiR* strain, *chiR* mRNA levels were not detected, as expected, and expression levels of the 9 chitinase-related genes were significantly decreased by 1.14- to 4.35-fold compared to the *S. marcescens* Db11 (**Table 2**). The one gene in the *chiR* operon (*cbp21*, *SMDB11_2877*) exhibited a 4.35-fold decrease in expression compared to the wild-type strain. Overall, these results demonstrated ChiR is a positive regulator of all chitinase-related genes in *S. marcescens* and were consistent with observed changes in chitinase activity conducted by assays.

Table 2 Chitinase-related gene expression in *S. marcescens* Db11, chiROE and ΔchiR by real-time PCR.

Gene no.	Locus	Gene product	Estimated binding site	Fold-change		
				Db11	*chiROE*	*ΔchiR*
1	*SMDB11_4243*	Exo-chitinase (*chiA*)	7 binding sites	1.00±0.10	10.93±0.70	0.54±0.08
2	*SMDB11_2875*	Exo-chitinase (*chiB*)	3 binding sites	1.00±0.15	5.28±0.13	0.87±0.10
3	*SMDB11_0468*	Endo-acting non-processive chitinase (*chiC*)	None	1.00±0.05	9.32±0.96	0.56±0.16
4	*SMDB11_2876*	LysR-family transcriptional regulator (*chiR*)	1 binding site	1.00±0.12	593.59±41.01	0.00±0.00
5	*SMDB11_2877*	Chitin binding protein (cbp21)	6 binding sites	1.00±0.13	9.23±0.78	0.23±0.03
6	*SMDB11_0477*	Chitobiase (*chb*)	1 binding site	1.00±0.09	3.78±0.19	0.76±0.11
7	*SMDB11_1083*	Putative chitin-binding protein	None	1.00±0.08	2.12±0.11	0.88±0.25
8	*SMDB11_1994*	Chitinase	2 binding sites	1.00±0.09	6.74±0.25	0.82±0.11
9	*SMDB11_4602*	Putative chitobiase	1 binding site	1.00±0.09	3.09±0.24	0.83±0.17
10	*SMDB11_1190*	N-beta	1 binding site	1.00±0.05	2.36±0.21	0.82±0.12

	acetylhexosaminidase (*Nag*Z)

2.3.5 Direct 2,3-butanediol production from crystal chitin

An initial baseline of 2,3-butanediol production in the *S. marcescens* Db11 was conducted for growth on glucose and *N*-acetylglucosamine. There was no difference between cell growth using glucose and *N*-acetylglucosamine (**Figure S2A**). The starting concentration of 20 g/L glucose or 20 g/L *N*-acetylglucosamine was depleted by around 30 h (**Figure S2B**). *S. marcescens* Db11 was able to produce 3.90 ± 0.03 g/L 2,3-BD from *N*-acetylglucosamine and 5.16 ± 0.05 g/L 2,3-BD from glucose (**Figure S2C**), indicating that *N*-acetyglucosamine can be used a carbon source by *S. marcescens* Db11. There was almost no difference at 2,3-BD production among the *ΔchiR* strain (3.82 ± 0.07 g/L), the *chiROE* strain (3.65 ± 0.05 g/L) and the wild-type strain (3.90 ± 0.03 g/L) when grown in either 20 g/L *N*-acetylglucosamine, implying that ChiR does not affect the 2,3-BD synthetic pathway.

To test the utility of overexpressing ChiR for a chitin-based consolidated bioprocess, untreated chitin was used as the carbon source for producing 2,3-butanediol using the wild-type Db11 strain and the *chiROE* strain. Cell growth, chitinase activity, total secreted protein and 2,3-BD production were monitored during a 3-day fermentation. Cell yields for the two strains were comparable at the end of fermentation (the *chiROE* strain cell yield of 1.03 ± 0.05 g/L and the wild-type Db11 cell yield of 1.22 ± 0.15 g/L) (**Figure S3A**). However, the *chiROE* strain showed clear increases in

endochitinase, chitobiosidase, and β-*N*-acetylglucosaminidase activity with 3-, 1.75-, and 3.26-fold increases, respectively, compared to the wild-type Db11 (**Figure S3B**). A final 2.93-fold improvement in 2,3-BD titer (1.13 ± 0.08 g/L) was achieved in the *S. marcescens chiROE* compared to 2,3-BD titer (0.4 ± 0.05 g/L) for the *S. marcescens* Db11 (**Figure 6**), indicating an improvement in utilization chitin as carbon source after overexpression *chiR*.

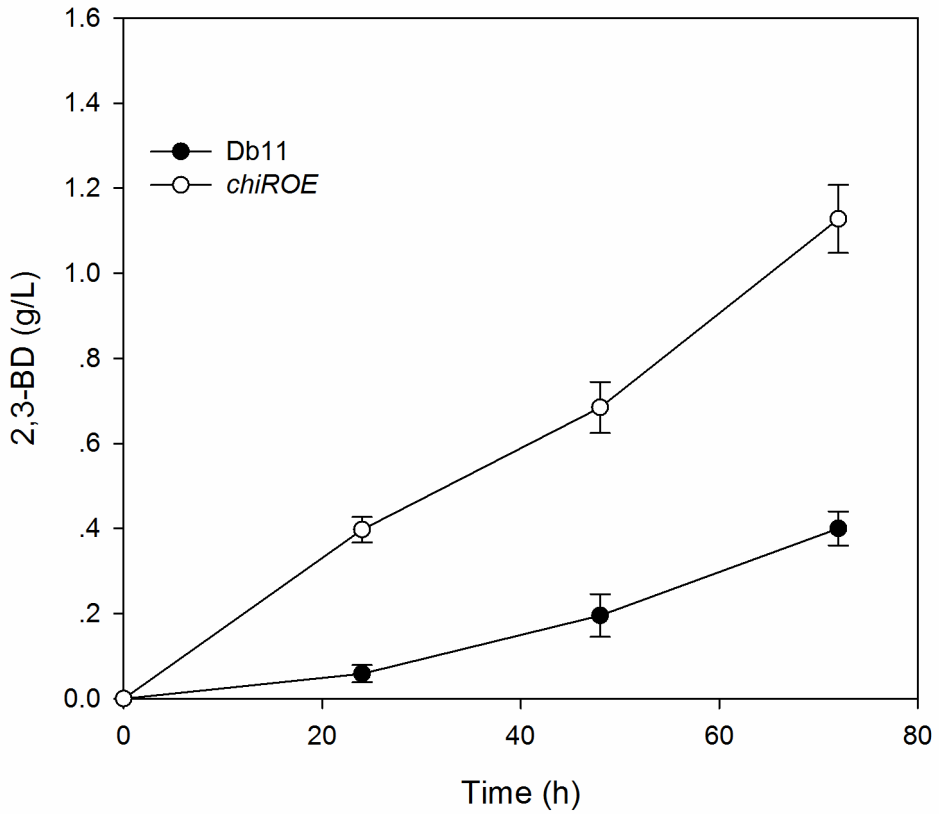


Figure 6 Production of 2,3-BD using the *S. marcescens* Db11 (filled circle) and the *S. marcescens* chiROE strain (unfilled circle) under M9 medium supplemented with 1% yeast extract and 2% chitin at 30 ℃ and 100 rpm.

2.4 Discussion

Due to the abundance of chitin in nature and its availability as an industrial waste product, use of chitinases and chitinolytic organisms have the potential to be used for

consolidated bioprocessing applications for biochemical production such as biofuels or pharmaceuticals (Aam et al. 2010; Dahiya et al. 2006; Yan and Fong 2015). *S. marcescens* has been an interesting organism for studying chitin degradation due to its native chitinolytic capabilities. Among the 10 chitinase-related genes known in the *S. marcescens* Db11, one regulatory protein, ChiR, was previously identified as a potential positive regulator of chitinase genes (Watanabe et al. 1997; (Suzuki et al. 2001). In this study, we studied ChiR function by the construction of a *chiR* overexpression (*chiROE*) strain and a *chiR* deletion (*ΔchiR*) strain. We found that ChiR significantly improved the expression of chitinolytic enzymes in *S. marcescens*. The overexpression strain, *chiROE,* had increased chitinase activity ranging from 2.14- to 6.13-fold higher than the wild-type Db11 and gene expression of 9 chitinase-related genes also increased 2.12- to 10.93-fold (compared to the wild-type Db11). Both activity and transcription data implied ChiR can be a promising target for enhancing chitin utilization. The *chiROE* strain was then grown on untreated crystal chitin and produced a maximum titer of 1.13 ± 0.08 g/L of 2,3-butanediol, a 2.93-fold improvement over the wild-type strain. To our knowledge, this is the first report of direct bioconversion of 2,3-butanediol from untreated chitin in *S. marcescens*.

Previous studies using mutagenesis and interaction assays have suggested that ChiR is a potential positive regulatory protein of chitinases in *S. marcescens* with proposing a regulatory cascade where ChiR positioned as an early phase regulator. In our study, the function of ChiR was tested directly by employing genetic tools to delete and

overexpress *chiR*. We reported 9 chitinase-related gene expression level changes in a *chiR* overexpression strain and a *chiR* deletion strain compared to the wild-type strain and corresponding chitinase activity. The *S. marcescens chiROE* strain was able to increase all 10 chitinase-related gene transcription levels. In our study, we found that 4 genes (*chiA*, *chiB*, *cbp21*, and *SMDB11_1994*) that possess more than 2 binding sites had more than 5-fold increases in transcription level (**Table 2** and **Table S3**) after *chiR* overexpression. It was also found that another 4 genes (*chb*, *SMDB11_1083*, *SMDB11_4602*, and *SMDB11_1190*) showed 2.12- to 3.78-fold increased expression and have less than 2 potential ChiR binding sites. One exception to the ChiR binding is the *chiC* gene that had 9.32-fold increased expression but no identifiable ChiR binding site. These data seem to indicate that ChiR functions in a more direct manner for regulating chitinase-related gene expression rather triggering a regulatory cascade. The real-time PCR data also identified a potential novel chitinase SMDB11_1994 that is annotated to encode a GH18 protein whose role and function is still uncharacterized. A protein blast showed this protein is 87% similarity to ChiD (PDB: 4NZC) from *Serratia proteamaculans* that is a chitinase with both hydrolytic and transglycosylation functions (Madhuprakash et al. 2013).

The overall physiological effect of manipulating ChiR was considered. The *chiROE* strain had lower biomass/cell yields when grown on glucose or *N*-acetylglucosamine compared to the wild-type Db11 and the *ΔchiR* strain. Additionally, the *ΔchiR* strain had biomass/cell yields that were increased by 1.30-fold compared to the wild-type

strain for growth on glucose or *N*-acetylglucosamine. In the *chiROE* strain, endochitinase, chitobiosidase and β-*N*-acetylglucosaminidase activity per cell were roughly 2.23-, 2.14-, and 6.31-fold higher, respectively, compared to the wild-type strain. The *chiROE* strain also had 2.41-fold higher total secreted protein per cell as compared to the wild-type strain. This increase in chitinase activity and in secreted overall protein concentration implies that an increased metabolic burden of producing chitinases production may decrease the availability of metabolic resources to produce basic biomass components (Wu et al. 2016); the *chiROE* strain likely has an increased metabolic burden (associated with chitinase production) while the *ΔchiR* strain likely has a reduced metabolic burden. These results would be consistent with the changes in biomass/cell yield for the *chiROE* and *ΔchiR* strains.

Chitin and colloidal chitin have an induction effect on chitinase production; however, their regulatory roles are still unclear. We tested *S. marcescens* Db11 transcriptomics (unpublished data) when grown at LB, LB plus colloidal chitin and LB plus chitin. First, we found that *chiR* gene mRNA level did not change significantly in these three conditions. Second, transcript levels for most chitinase-related genes showed no significant difference among the three growth conditions. One exception is that *cbp21* mRNA level increased more than 2-fold under LB with chitin/colloidal chitin compared to that grown under plain LB medium. Furthermore, in terms of overall protein secretion at the three growth conditions, there was almost no overall secreted protein changes of the *ΔchiR* strain, while an ascending trend of secreted overall

protein from plain LB, LB plus colloidal chitin, and LB plus chitin were observed in both the *chiROE* strain and the wild-type strain (**Figure 5D**). One explanation regarding the relationship between the observance of chitinase activities changes and the measurement of mRNA level changes under three conditions is that chitin or colloidal chitin may assist in inducing chitinase secretion instead of inducing chitinase overexpression.

Recently, genetic engineering of regulatory proteins for enhancing cellulase or chitinase production has been employed with various degrees of success (Aghcheh et al. 2014; Häkkinen et al. 2014; Portnoy et al. 2011; Wang et al. 2013). For instance, Deng and Fong demonstrated that CelR is a negative regulator of cellulose production in *Thermobifida fusca* and a recombinant CelR deletion strain enabled an enhancement of specific cellulase activity by 16.70-fold (Deng and Fong 2010). Multiple regulatory targets were investigated and applied to improve cellulase production in *Trichoderma reesi* (Aghcheh et al. 2014; Häkkinen et al. 2014; Portnoy et al. 2011; Wang et al. 2013). Overexpression of a single regulator gene, deletion of a single regulator gene or a combinational method can result in increasing cellulase activity range from 2- to 5-fold, shown in **Table S7**. A study that investigated tuning multiple regulator expression levels in *Penicillium oxolicum*. After overexpression of *clrB* and deletion of *bgl2* and *creA*, the cellulase activity increased from 10- to 50-fold on different substrates (Yao et al. 2015). In our study, we demonstrated that a *chiR* overexpression strain was able to increase chitinase activity 2.14- to 6.31-fold

compared to the wild-type *S. marcescens* Db11 strain as a direct result of increasing mRNA levels for 9 chitinase-related genes.

Despite high levels of interest and research effort to developing bioprocesses using renewable feedstock, very few examples exist that demonstrate direct utilization of untreated feedstock. In the present study, we report direct bioconversion of untreated crystal chitin to 2,3-butanediol with a maximum titer of 1.13 ± 0.08 g/L from 3 days batch fermentation in a ChiR overexpression strain of *S. marcescens*. Recently, several attempts have been made to utilize various natural biorenewable resources for producing 2,3-BD, shown in **Table 3**. Jiang et al. used diluted sulfide acid to hydrolyze *Jatropha* hulls, resulting in a 37.3 g/L of reducing sugar. After 60 h fermentation, 4.11 g/L 2,3-BD was produced with a yield of 0.11 g/g in a non-cellulolytic native producer, *Klebsiella oxytoca* (Jiang et al. 2012). In another study, Shin introduced a synthetic 2,3-BD pathway and overexpression of a periplasmic cellodextrinase (Ced3) in *E. coli*. Fermentation was conducted coupled with simultaneous saccharification (SSF) and a maximum titer 4.2 g/L 2,3-BD was obtained after 24 h fermentation with a yield of 0.84 g/g from cellodextrin (Shin et al. 2012). Another novel concept was demonstrated by using a cyanobacterium to produce 2,3-BD through photosynthetic pathways. In their initial study, a synthetic 2,3-BD synthetic pathway was introduced in *Synechococcus elongatus* controlled by an IPTG inducible promoter $P_L lacO_1$. As a proof of concept, 0.13 g/L 2,3-BD was produced from a recombinant *S. elongates* (Nozzi and Atsumi 2015). A follow-up

study focused on combinational optimization of sugar/light supplementation and cell density yielding a maximum titer of 3.0 g/L 2,3-BD after 10 days fermentation (McEwen et al. 2016). In our initial study, the *chiROE* strain was able to produce 1.13 ± 0.08 g/L 2,3-BD from 2% untreated crystal chitin, indicating a promising CBP scheme. This is the first study reported the feasibility using chitin as a feedstock for 2,3-BD production in *S. marcescens*.

Table 3 Microbial production of 2,3-butanediol production from biomass feedstocks.

Substrate	Microorganisms	Titer (g/L)	Yield (g/L)	Time (h)	References
Light + glucose	*Synechococcus elongatus*	3.0	n.p.	10 days	(McEwen et al. 2016)
Light	*S. elongatus*	0.12	n.p.	n.p.	(Nozzi and Atsumi 2015)
Lignocellulosic biomass hydrolysate	*Klebsiella oxytoca*	4.11	0.11	60 h	(Jiang et al. 2012)
Lignocellulosic biomass hydrolysate	*E. coli*	4.2	0.84	24 h	(Shin et al. 2012)
Crystal chitin	*S. marcescens Db11*	0.40	0.02	72 h	This study
Crystal chitin	*S.marcescens chiROE*	1.13	0.06	72 h	This study

Future studies can be made mainly based on two directions. First, this study demonstrates a proof-of-concept in producing 2,3-BD from chitin, however the 2,3-BD titer remains low, so future work can focus on improvements to increase 2,3-BD titer. For instance, increasing substrate concentration and feeding strategies can be employed to improve 2,3-BD titer. Combinational approaches of strain improvements can be made: first, adaptive evolution techniques can be applied to obtain a genomically stable mutant that is capable of fast decomposition of chitin; second, metabolic engineering strategies can be made to delete competitive metabolic

pathways (i.e. acetic acid, ethanol, lactic acid) that drain the metabolic flux away from 2,3-BD. Second, from an industrial application perspective, an efficient and economical pretreatment process from shrimp/crab shells to crystal chitin needs to be developed, since purified crystal chitin is not readily available in nature. Chitin can be extracted directly from shrimp/crab shells after removal of protein, calcium and low molecular weight compounds by treatment with sodium hydroxide and hydrochloric acid. The process kinetics of both demineralization and deproteinization have been previously considered and Percot et al. demonstrated that an optimization of the pretreatment parameters (e.g. concentration, time, temperature) can yield a low residual content of calcium (below 0.01%) and a high degree of acetylation (DA) (almost 95%) (Percot et al. 2003) that is suitable for downstream usage. In another method, Aye and Stevens investigated extraction of chitin from fresh shrimp shells through grinding and sieving (Aye and Stevens 2004). These authors demonstrated that the chitin after the dry screening method is ready to be used as a feedstock for an animal. These pretreatment methods without additional enzymes can contribute to developing a chitin biorefinery process from food wastes.

2.5 Conclusion

Microbial utilization of chitin, a potential renewable biomass feedstock, is being pursued as a means of developing novel consolidated bioprocessing for the production of chemicals. *Serratia marcescens* is a gram-negative bacterium that is known for its chitinolytic capability and as a native 2,3-butanediol producer. In *S. marcescens*, ChiR has been suggested to be a positive regulator of chitinase

production. In this study, we aim to understand the effect of ChiR in regulating nine chitinase-related genes in *S. marcescens* Db11 and demonstrate manipulation of *chiR* as a useful and efficient genetic target to enhance chitin utilization. First, a *chiR*overexpression (*chiROE*) strain and a *chiR* deletion (*ΔchiR*) strain were generated and characterized in terms of cellular growth, chitinase activity, and total secreted protein. Compared to the wild-type Db11 strain, the *S. marcescens chiROE* strain showed an increase in chitinase activity (2.14- to 6.31-fold increase). Increased transcriptional expression of chitinase-related genes was measured using real-time PCR, showing 2.12- to 10.93-fold increases. The *S. marcescens ΔchiR* strain showed decreases in chitinase activity (4.5- to 25-fold decrease), confirming ChiR's role as a positive regulator of chitinase expression. Finally, *chiR*overexpression was investigated as a means of increasing biochemical production (2,3-butanediol) from crystal chitin. The *chiROE* strain produced 1.13 ± 0.08 g/L 2,3-butanediol from 2% crystal chitin, a 2.83-fold improvement from the wild-type strain, indicating ChiR is an important and useful genetic engineering target for enhancing chitin utilization in *S. marcescens*.

Chapter 3 Design and modularized optimization of one-step production of *N*-acetylneuraminic acid from chitin in *Serratia marcescens*

Importance of this part of work:

Since I have successfully developed shuttle vectors for overexpression of genes in *S. marcescens* and demonstrated that *S. marcescens* can directly utilize chitin to produce chemicals, this part of work is to demonstrate a proof-of-concept of producing novel chemcials from chitin by introducing exogenous pathway genes in *S. marcescens*. Prior to this dissertation work, there is few research reporting about producing *N*-acetylneuraminic acid (chemical target) from chitin. Here, I demonstrate an engineered bioprocess to produce *N*-acetylneuraminic acid (Neu5Ac) directly from chitin using the chitinolytic bacterium, *Serratia marcescens* by selecting and characterizing promoters, characterization of heterologous enzyme activity, and optimization of pathway fluxes.

The conclusions and fundamental results derived from this part of work:

By generating RNA-Seq data for *S. marcescens* growth in different carbon-limited conditions (glucose, *N*-acetylglucosamine and glycerol), 12 promoters with varying strength were identified and characterized to implement for transcriptional control. Neu5Ac production was initially engineered into *S. marcescens* through heterologous expression of *N*-acetyglucosamine 2-epimerase (*slr1975*) and *N*-acetylneuraminic acid aldolase (*nanA*). Activity of both genes was characterized *in vitro* for kinetics and *in vivo* expression using promoters identified in this study. Optimization of Neu5Ac

production was accomplished by balancing pathways fluxes through promoter swapping and replacing the reversible *nanA* with the irreversible gene *neuB*. The optimized recombinant strain P_{T5}-*slr1975*-P_{rplJ}-*neuB* was able to produce 0.48 g/L Neu5Ac from 20 g/L *N*-acetylglucosamine, and 0.30 g/L Neu5Ac from 5 g/L crystal chitin. These results represent the first demonstration of direct conversion of crystal chitin to *N*-acetylneuraminic acid and illustrate the potential utility of *S. marcescens* as a chitinolytic bioprocess organism.

3.1 Introduction

Metabolic engineering approaches to chemical production from renewable biomass feedstocks (such as cellulosic biomass) have received tremendous research interest (Liao et al. 2016; Lynd et al. 2005; Olson et al. 2012). Chitin is the second most abundant carbon compound on earth that is a homopolymer of *N*-acetylglucosamine and can be degraded by chitinases as a potential renewable raw material. The special property of chitin is that it serves as a source for both carbon and nitrogen (C:N = 8:1), which is an ideal resource for production of nitrogen-rich chemical compounds (i.e. *N*-acetylneuraminic acid, C:N =11:1) (Steiger et al. 2011). Development of chitin as a potential biomass feedstock can enhance the biorefinery industry by broadening the range of practicable biomass options and lowering the environmental impact of food industry waste (Yan and Chen 2015; Yan and Fong 2015). In theory, chitin could be used in a consolidated bioprocess (CBP) (Lynd et al. 2005) in a similar fashion as cellulose, however, this has yet to be demonstrated. Two of the main areas hindering development of a chitinolytic CBP is a lack of genetic engineering tools and limited

characterization of metabolism in chitinolytic species.

Considering the various aspects of a chitin-based CBP scheme, we sought to develop the chitinolytic bacterium *Serratia marcescens*, as a potential platform organism for *N*-acetylneuraminic acid (Neu5Ac) production. *S. marcescens* has been studied and characterized for its chitinolytic capabilities (Monreal and Reese 1969; Vaaje-Kolstad et al. 2013; Watanabe et al. 1997) and can exclusively convert chitin into GlcNAc (Kim et al. 1998). The genome of *S. marcescens* Db11 has been sequenced (Iguchi et al. 2014). The main challenge is to engineer a synthetic pathway for Neu5Ac production in *S. marcescens* since the wild-type *S. marcescens* Db11 does not harbor a natural route and existing genetic engineering tools are limited. Currently only a small number of well-characterized and tractable *S. marcescens* promoters are available for *in vivo* application. A T5 promoter (Gerc et al. 2012) and a native *budAB* operon promoter (Sun et al. 2012b) were investigated either to study a multifunctional peptide synthetase (T5 promoter) or to facilitate NADH regeneration by overexpression of an oxidase (*budAB* promoter). In the previous work, we have developed an in-frame chromosomal gene deletion method for *S. marcescens* based on its native exonucleases and recombinases, and also established a stable shuttle vector thus expanding the tools available for manipulation of *S. marcescens* (Yan et al. 2017).

N-acetylneuraminic acid (Neu5Ac) is a ubiquitous sialic acids that carries out various

biological functions by acting as a receptor for microorganisms, viruses, toxins, and hormones, by masking receptors and regulation of the immune system (Chen and Varki 2010; Schauer 2000; Varki 1992). It can be used as a nutraceutical for infant brain development (Wang 2009), pharmaceutical for preventing influenza virus infections (Itzstein et al. 1993), and nanocarrier for cancer targeting and therapy (Bondioli et al. 2011).

Previous metabolic engineering studies for production of Neu5Ac have primarily used engineered strains of *E. coli* for the production of Neu5Ac from glucose by fermentation. Boddy and Lundgren assembled a six-step pathway for Neu5Ac production from fructose-6-phosphate by overexpressing genes of glucosamine synthase (*glmS*), Neu5Ac synthase (*neuB*), and UDP-GlcNAc 2-epimerase (*neuC*) from *Neisseria meningitides* (Lundgren and Boddy 2007). The *nanA* and *nanT* genes were then knocked out to abolish sialic acid catabolism. In this engineered pathway, NeuB was responsible for converting *N*-acetylmannosamine (ManNAc) and phosphoenolpyruvate (PEP) to Neu5Ac. As a result, 1.5 g/L Neu5Ac was obtained with glucose as the carbon source in a 98 h fermentation process. In another study, Kang et al. designed a Neu5Ac synthetic pathway by introducing a GlcNAc 2-epimerase (Slr1975) and glucosamine-6-phosphate acetyltransferase (GNA1) in *E. coli* DH5α (Kang et al. 2012). By deleting the feedback inhibition of glucosamine-6-phosphate synthase (Δ*nagAB*) to increase the accumulation of GlcNAc and pyruvate (Δ*ackA*Δ*poxB*Δ*ldhA*) and blocking the catabolism of Neu5Ac (Δ*nanT*),

7.85 g/L of Neu5Ac was obtained using glucose as the carbon source in a 96-h fed-batch fermentation with productivity of 0.08 g/L h. In a follow-up study, Yang et al. engineered an aptazyme-based biosensor of Neu5Ac to evolve the key enzyme Neu5Ac synthase (NeuB) (Yang et al. 2017). The modified strain was able to improve the titer by 34%. Samain et al. overexpressed NeuB and NeuC from *Campylobacter jejuni* in a *neuA*, *nanA*, *nanK* and *nanT* deletion strain of *E. coli* producing 39 g/L of Neu5Ac with glycerol as the carbon source (Samain 2008).

Here we investigate developing a chitin-based CBP scheme to utilize chitin-based feedstock to produce *N*-acetylneuraminic acid through a single-unit operation (chitinase production and secretion, hydrolysis of chitin, and fermentation of available sugars to produce Neu5Ac). The main advantage of this CBP process is to reduce the substrate cost and simplify the process while still producing a value-added product. A rough analysis of the economic difference between a glucose-based process and the CBP is shown as follows. Assuming 1 g chitin can produce 1 g *N*-acetylglucosamine (a highest yield reported 0.76 GlcNAc g/g chitin (Husson et al. 2017)), the carbon cost to produce 1 kg Neu5Ac from dried shrimp shells ($100-120/ton (Yan and Chen 2015)) is around $0.071 based on the theoretical yield of 1.40 g Neu5Ac/g GlcNAc; while for the glucose process, the carbon cost to produce 1 kg Neu5Ac from the glucose ($55.5/kg, Sigma) is about $32.27 based on the theoretical yield of 1.72 g Neu5Ac/g glucose.

In this study, we applied the recently developed systems for overexpression of genes in *S. marcescens* to generate designed strains that directly converts crystal chitin to Neu5Ac. To identify reliable promoters, we started with characterizing *S. marcescens* gene expression levels (RNA-Seq) of the wild-type strain grown on various carbon sources (glucose, *N*-acetylglucosamine, and glycerol) and tested the activity of 12 potential constitutive promoters by measuring the mRNA (real-time PCR) and protein expression levels (fluorescence intensity) of a *gfp* reporter gene. Subsequently, a two-step synthetic pathway (*N*-acetylglucosamine 2-epimerase and *N*-acetylneuraminic aldolase) for Neu5Ac production was introduced in *S. marcescens* and the functionality of the pathway enzymes was validated by real-time PCR and enzymatic activity assays. Improvement of Neu5Ac titer (0.48 g/L Neu5Ac from 20 g/L *N*-acetylglucosamine) was achieved by tuning mRNA level (promoter swapping) and by utilizing an irreversible enzyme to replace a reversible enzyme (switch *N*-acetylneuraminic aldolase to *N*-acetylneuraminic acid synthase). Thus, *S. marcescens* has been engineered to directly convert untreated chitin to Neu5Ac.

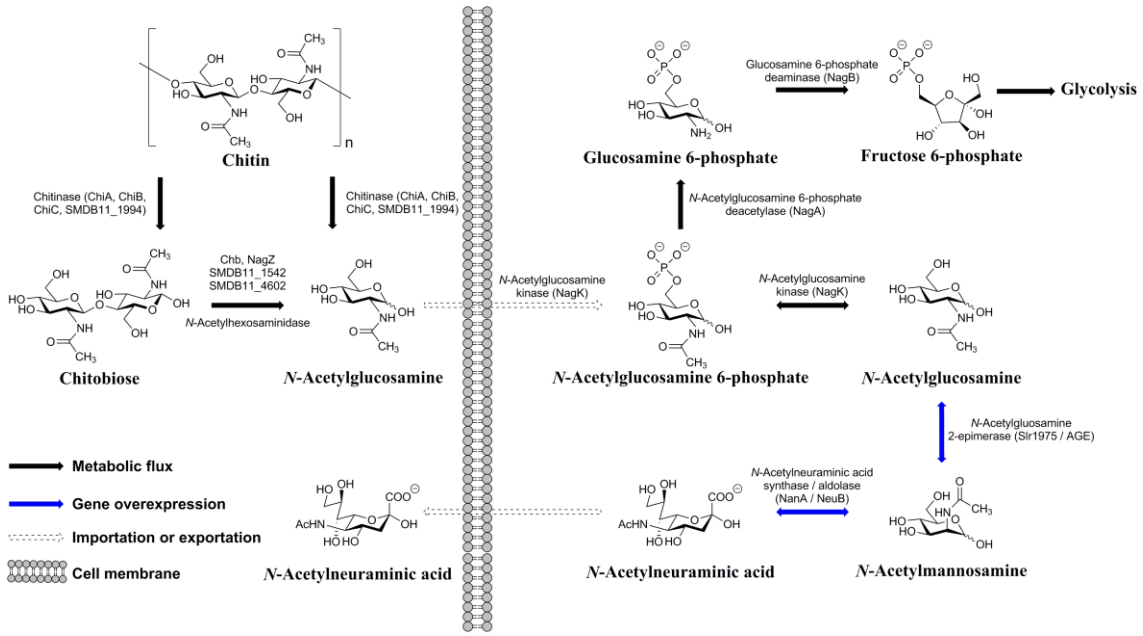


Figure 7 A scheme of *N*-acetylneuraminic acid from chitin by *Serratia marcescens*.

3.2 Materials and methods

3.2.1 Bacterial strains and plasmids

S. marcescens Db11 is the wild-type strain and was purchased from Caenorjabditis Genetics Center (Twin City, MN https://cbs.umn.edu/cgc) (Flyg et al. 1980). *E. coli* NEB10β (New England Biolabs, Ipswich, MA) was used as a host for pUC19-backbone plasmid construction and propagation. *E. coli* TOP10 (Thermo Fisher Scientific, Waltham, MA) was used as a host for pEXP-backbone plasmids plasmid construction and propagation. Strain and plasmids used in this study are listed in **Table 4**.

Table 4 Strains and plasmids used in this study.

Strain name	Genotype	Resources
S. marcescens Db11	Wild-type	Caenorjabditis Genetics Center
S. marcescens pQY04	Db11 strain harbors pUC19 inserted by P_{T5}-*GFP*, amp^R	This study
S. marcescens pQY05	Db11 strain harbors pUC19 inserted by P_{ampR}-*GFP*, amp^R	This study

S. marcescens pQY06	Db11 strain harbors pUC19 inserted by P_{budAB}-*GFP*, amp^R	This study
S. marcescens pQY19	Db11 strain harbors pUC19 inserted by P_{ompA}-*GFP*, amp^R	This study
S. marcescens pQY20	Db11 strain harbors pUC19 inserted by P_{rplJ}-*GFP*, amp^R	This study
S. marcescens pQY21	Db11 strain harbors pUC19 inserted by P_{rpsL}-*GFP*, amp^R	This study
S. marcescens pQY24	Db11 strain harbors pUC19 inserted by P_{gpmA}-*GFP*, amp^R	This study
S. marcescens pQY25	Db11 strain harbors pUC19 inserted by P_{metF}-*GFP*, amp^R	This study
S. marcescens pQY27	Db11 strain harbors pUC19 inserted by P_{phlA}-*GFP*, amp^R	This study
S. marcescens pQY23	Db11 strain harbors pUC19 inserted by P_{cysP}-*GFP*, amp^R	This study
S. marcescens pQY22	Db11 strain harbors pUC19 inserted by P_{rpsG}-*GFP*, amp^R	This study
S. marcescens pQY26	Db11 strain harbors pUC19 inserted by P_{SMDB11_4509}-*GFP*, amp^R	This study
S. marcescens pQY28	Db11 strain harbors pUC19 inserted by P_{cysP}-*slr1975*-P_{cysP}-*nanA*, amp^R	This study
S. marcescens pQY29	Db11 strain harbors pUC19 inserted by P_{rplJ}-*slr1975*-P_{rplJ}-*nanA*, amp^R	This study
S. marcescens pQY30	Db11 strain harbors pUC19 inserted by P_{T5}-*slr1975*-P_{T5}-*nanA*, amp^R	This study
S. marcescens pQY36	Db11 strain harbors pUC19 inserted by P_{T5}-*slr1975*-P_{rplJ}-*nanA*, amp^R	This study
S. marcescens pQY37	Db11 strain harbors pUC19 inserted by P_{rplJ}-*slr1975*-P_{T5}-*nanA*, amp^R	This study
S. marcescens pQY43	Db11 strain harbors pUC19 inserted by P_{T5}-*AGE*-P_{rplJ}-*nanA*, amp^R	This study
S. marcescens pQY42	Db11 strain harbors pUC19 inserted by P_{T5}-*slr1975*-P_{rplJ}-*neuB*, amp^R	This study
S. marcescens pQY46	Db11 strain harbors pUC19 inserted by P_{T5}-*AGE*-P_{rplJ}-*neuB*, amp^R	This study
E. coli NEB10β	Δ*(ara-leu) relA1 endA1* (Str^R)	NEB
E. coli pQY16	*E. coli* NEB10β harbors pUC19 inserted by P_{T5}-*GFP*, amp^R	This study
E. coli pQY17	*E. coli* NEB10β harbors pUC19 inserted by P_{ampR}-*GFP*, amp^R	This study
E. coli pQY18	*E. coli* NEB10β harbors pUC19 inserted by P_{budAB}-*GFP*, amp^R	This study
E. coli pQY19	*E. coli* NEB10β harbors pUC19 inserted by P_{ompA}-*GFP*, amp^R	This study
E. coli pQY20	*E. coli* NEB10β harbors pUC19 inserted by P_{rplJ}-*GFP*, amp^R	This study
E. coli pQY21	*E. coli* NEB10β harbors pUC19 inserted by P_{rpsL}-*GFP*, amp^R	This study
E. coli pQY24	*E. coli* NEB10β harbors pUC19 inserted by P_{gpmA}-*GFP*, amp^R	This study
E. coli pQY25	*E. coli* NEB10β harbors pUC19 inserted by P_{metF}-*GFP*, amp^R	This study
E. coli pQY27	*E. coli* NEB10β harbors pUC19 inserted by P_{phlA}-*GFP*, amp^R	This study
E. coli pQY23	*E. coli* NEB10β harbors pUC19 inserted by P_{cysP}-*GFP*, amp^R	This study
E. coli pQY22	*E. coli* NEB10β harbors pUC19 inserted by P_{rpsG}-*GFP*, amp^R	This study
E. coli pQY26	*E. coli* NEB10β harbors pUC19 inserted by P_{SMDB11_4509}-*GFP*, amp^R	This study
E. coli pQY28	*E. coli* NEB10β harbors pUC19 inserted by P_{cysP}-*slr1975*-P_{cysP}-*nanA*, amp^R	This study
E. coli pQY29	*E. coli* NEB10β harbors pUC19 inserted by P_{rplJ}-*slr1975*-P_{rplJ}-*nanA*, amp^R	This study
E. coli pQY30	*E. coli* NEB10β harbors pUC19 inserted by P_{T5}-*slr1975*-P_{T5}-*nanA*, amp^R	This study
E. coli pQY36	*E. coli* NEB10β harbors pUC19 inserted by P_{T5}-*slr1975*-P_{rplJ}-*nanA*, amp^R	This study
E. coli pQY37	*E. coli* NEB10β harbors pUC19 inserted by P_{rplJ}-*slr1975*-P_{T5}-*nanA*, amp^R	This study
E. coli pQY43	*E. coli* NEB10β harbors pUC19 inserted by P_{T5}-*AGE*-P_{rplJ}-*nanA*, amp^R	This study
E. coli pQY42	*E. coli* NEB10β harbors pUC19 inserted by P_{T5}-*slr1975*-P_{rplJ}-*neuB*, amp^R	This study
E. coli pQY46	*E. coli* NEB10β harbors pUC19 inserted by P_{T5}-*AGE*-P_{rplJ}-*neuB*, amp^R	This study
E. coli TOP10	*rec*A1 *ara*D139 Δ(*ara leu*) (*StrR*) *end*A1	Invitrogen
E. coli pQY32	*E. coli* TOP10 harbors pEXP-5NT P_{T7}-*slr1975*, amp^R	This study
E. coli pQY33	*E. coli* TOP10 harbors pEXP-5NT P_{T7}-*nanA*, amp^R	This study
E. coli pQY35	*E. coli* TOP10 harbors pEXP-5NT P_{T7}-*neuB*, amp^R	This study
E. coli pQY40	*E. coli* TOP10 harbors pEXP-5NT P_{T7}-*AGE*, amp^R	This study

Plasmid	Feature	Resources
pUC19	Used for gene expression in *S. marcescens*, ampR	NEB
pEXP 5NT/TOPO®	Used for cell-free protein expression, ampR	Invitrogen

All plasmids were constructed by DNA assembly techniques (Gibson et al. 2009). The vector backbones were PCR-amplified using Q5® Hot Start DNA polymerase (New England Biolabs, Ipswich, MA). The inserts (promoter sequences and coding sequences) were chemically synthesized from Integrated Device Technology (San Jose, CA) and the sequences and primers can be found in **Table S8** and **Table S9**, respectively. The PCR products were purified by a PCR purification kit (Zymo Research, Irvine, CA). The vector backbone were mixed with Gibson Assembly® Master Mix (New England Biolabs, Ipswich, MA) and incubated at 50 ℃ for 1 h. Subsequently, the assembly mixture was transformed to *E. coli* NEB10β or TOP10 strain. The presence of correctly cloned insertion was validated by colony PCR and DNA sequencing (Eurofins Genomics, Louisville, KY).

3.2.2 Chemicals and reagents

All chemicals unless otherwise specified were acquired from Sigma-Aldrich (St. Louis, MO) or Thermo Scientific (Waltham, MA).

3.2.3 Media and cultivation

All the *E. coli* strains were grown in LB or SOC media containing appropriate antibiotics at 37 ℃ on a rotary shaker (250 rpm). Antibiotics were used at the following concentrations: ampicillin, 200 μg/mL; streptomycin, 40 μg/mL.

All *S. marcescens* strains were grown in a M9 medium containing 0.1% yeast extract

and various carbon sources at 30 ℃ on a rotary shaker (200 rpm). Carbon sources were used at the following concentration: *N*-acetylglucosamine, 20 g/L; glucose, 20 g/L; glycerol, 20 g/L; chitin, 20 g/L; colloidal chitin, 20 g/L. Antibiotics were used at the following concentrations: ampicillin, 800 μg/mL. The 20 mL pre-cultures of *S. marcescens* strains were grown through overnight in a 125 mL Erlenmeyer flask at 30 ℃ on a rotary shaker (200 rpm). For fermentative production of Neu5Ac, the *S. marcescens* preculture was inoculated at 5% in 2 L Erlenmeyer flask and cells were grown at 30 ℃ on a rotary shaker (200 rpm).

Stock cultures of *E. coli* and *S. marcescens* were maintained at -80 ℃ in 26% (v/v) glycerol.

3.2.4 mRNA sequencing

The total RNA of *S. marcescens* Db11 was isolated using mid-log phase cell cultures by QIAGEN RNA protect reagent (Venlo, Netherlands) in combination with the QIAGEN RNeasy Mini kit (Venlo, Netherlands). The mRNA was isolated, enriched, and reverse transcribed into cDNA. The resulting cDNA was sequenced by a pair-ended reads using Illumina Hiseq 2500 (San Diego, CA). An average insert size of 500 bp was created and draft mRNA-Seq data were generated. CLC Genomics Workbench version 10 (QIAGEN, Venlo, Netherlands) and FASTQC were applied to trim reads for quality sequence data. The *S. marcescens* Db11 genome (GeneBank: HG326223) was used as a reference. Each mRNA sample was tested with two runs. The gene expression level was calculated by normalizing each gene's reads per

kilobase per million mapped sequence reads (RPKM) to overall gene numbers, namely:

$$\text{Gene expression level} = \frac{\text{RPKM value}}{4831}$$

3.2.5 Measurement of cell density

The measurement of *S. marcescens* strains culture density was generally quantified at OD600 using a Biomate3 UV/VIS spectrophotometer (Thermo Fisher Scientific, Waltham, MA). The growth of *S. marcescens* strains on crystal chitin and colloidal chitin was measured by testing cytoplasmic protein content (Yan et al. 2017). The correlation factor of the overall protein content to the dry cell weight (DCW) was DCW (g/L) = $1.20 \pm 0.03 \times$ Protein (g/L).

3.2.6 Quantification of gene transcription using real-time PCR

The expression levels of *sfgfp*, *slr1975*, and *nanA* in the *S. marcescens* recombinant strains and the wild-type Db11 strain were measured by real-time PCR. A housekeeping gene (*luxS*, *SMDB11_0167*) was used as a reference gene. Primers used for real-time PCR can be found in **Table S10**. The relative mRNA level of the targeted gene was quantified by normalizing to the reference gene.

3.2.7 Measurement of GFP fluorescence intensity

The GFP fluorescence intensity of the *S. marcescens* recombinant strain cultures were measured by flow cytometry. The protocol is according to our previous study with minor modifications (Yan and Fong 2017b). Briefly, the cell cultures were harvested at exponential phase (OD600 ≈ 0.4 at 2.5 h) and stored on ice for at least 1 h to cease cell growth. Cell pellets were then re-suspended with PBS buffer to an optical density

(OD600 0.1 ~ 0.2) in a 1.5 mL centrifuge tube. The flow cytometry analysis was performed using BD Accuri^{TM} C6 flow cytometer (San Jose, CA). The fluorescence intensity of each *S. marcescens* recombinant strains was normalized to the fluorescence intensity of the wild-type Db11.

3.2.8 Quantification of Neu5Ac, GlcNAc and by-products

The concentration of GlcNAc, Neu5Ac, and all by-products were analyzed by an Ultimate 3000 HPLC (Dionex, Sunnyvale, CA), equipped with a Bio-Rad Aminex HPX-87H column (Hercules, CA), a UV detector (199 nm) (Dionex, Sunnyvale, CA), and a refractive index detector (Shodex, Japan). The condition was run with 5 mM H_2SO_4 at 0.6 mL/min and 55 ℃.

3.2.9 Measurement of *in vitro* Neu5Ac production

The *in vitro* Neu5Ac production by the *S. marcescens* recombinant strains was measured by feeding the cell lysates with GlcNAc. In brief, the *S. marcescens* cell cultures were harvested by centrifuging at 14,000×*g* and 4 ℃ for 5 min. The cell pellets were washed twice by PBS buffer and lysed by 1×Bugbuster (EMD Millipore, Billerica, MA) for 30 min. The cell lysate was centrifuged at 14,000×*g* and 4 ℃ for 5 min and the supernatant was ready for *in vitro* Neu5Ac production assay. For strains overexpressing NeuB, the assay was performed in a total volume of 200 μL including 12.5 mM GlcNAc, 10 mM PEP, 12.5 mM $MnCl_2$, 100 mM Bicine buffer (pH 8.0), 50 μL cell lysate and 2.5 mM ATP. For strains overexpressing NanA, the assay was performed in a total volume of 200 μL containing 12.5 mM GlcNAc, 10 mM pyruvate, 12.5 mM $MgCl_2$, 100 mM Tris-HCl buffer (pH 7.5), 50 μL cell lysate and 2.5 mM

ATP. All the reactions were incubated at 37 ℃ and samples were taken every 2 - 4 h.

3.2.10 Cell-free overexpression and purification of Slr1975, AGE, NanA, and NeuB

The GlcNAc 2-epimerases (Slr1975 and AGE) and Neu5Ac aldolase/synthase (NanA and NeuB) were produced using a Cell-Free Protein Expression kit (Thermo Fisher Scientific, Waltham, MA) and purified in a His-Spin Protein Miniprep^{TM} kit (Zymo Research, Irvine, CA). The enzymes were purified in homogeneity and analyzed by protein electrophoresis (Bio-Rad, Hercules, CA). Protein concentration was determined by Bradford assay (Bio-Rad, Hercules, CA)

3.2.11 *N*-acetylglucosamine 2-epimerase enzyme assay

The GlcNAc 2-epimerase enzyme assay was carried out by measuring ManNAc production. The reaction mixture was carried out in a 200 μL volume reaction containing 12.5 mM GlcNAc, 12.5 mM $MnCl_2$, Tris-HCl buffer (pH 7.5), 50 μL enzyme and 2.5 mM ATP. The reaction was incubated at 37 ℃ for 2 h. Each reaction was run at least three times. One unit of the enzyme activity was defined as mM ManNAc produced per hour.

3.2.12 *N*-acetylneuraminic acid aldolase/synthase enzyme assay

The forward/reverse reaction of NanA or NeuB was carried out by measuring Neu5Ac/ManNAc production. The enzyme assay reactions were carried out in a 200 μL volume reaction containing appropriate substrates, 12.5 mM $MnCl_2$, 100 mM Bicine buffer (pH 8.0) and 50 μL enzyme. For the forward reaction of NanA/NeuB enzyme assay, the substrates are 12.5 mM ManNAc and 10 mM pyruvate/10 mM PEP,

respectively; for the reverse reaction of NanA/NeuB enzyme assay, the substrate is 12.5 mM Neu5Ac. Each reaction was run at least three times. One unit of the enzyme activity was defined as mM Neu5Ac/ManNAc produced per hour.

3.3 Results

To develop and optimize an engineered strain of *S. marcescens* for production of Neu5Ac from chitin, we conducted RNA-Seq experiments to identify potential native constitutive promoters, characterized transcriptional strength of identified promoters, measured activity of heterologous proteins for Neu5Ac production, implemented and balanced heterologous expression in *S. marcescens*, and swapped heterologous proteins to favor Neu5Ac production.

3.3.1 Characterization of *S. marcescens* constitutive promoter

In order to identify functional native *S. marcescens* promoters, we targeted native *S. marcescens* promoters based upon a mRNA sequencing (RNA-Seq) of the wild-type *S. marcescens* Db11 strain using different growth conditions (glucose, *N*-acetylglucosamine, glycerol). The total number of reads for each condition were: 5,986,581 (glucose), 7,820,963 (*N*-acetylglucosamine) and 19,025,155 (glycerol). The average coverage per gene (4831 total genes) was about 1239, 1619, and 3938 for glucose, *N*-acetylglucosamine and glycerol, respectively, indicating a high depth of coverage. Since our goal is to drive the exogenous Neu5Ac synthetic pathway gene expression, we specifically targeted promoter sequences with various strengths (high, medium, and low) based on the range of RPKM values. Based upon mRNA levels in all three conditions (looking for constitutive promoters), we selected two genes with

high expression (*ompA* and *rplJ*; RPKM > 1.0), 4 genes with medium expression (*rpsG*, *rpsL*, *gpmA*, and *metF*; 0.1 < RPKM < 1.0), and 4 genes with low expression (*cysP*, *phlA*, *SMDB11_4509*, and *budAB*; RPKM < 0.1), shown in **Figure 8A**. In addition to the 10 potential promoters, two promoters were used as a high strength positive control (P_{T5}) (Gerc et al. 2012) and a medium strength positive control (P_{ampR}) (Yan et al. 2017) that have been used in previous studies.

The functionality of the 12 promoters was examined by measuring both transcriptional and translational levels of GFP gene (**Figure 8B** and **Figure S4**). The T5 promoter was designated as high strength promoter, whose RFU (2328.32 ± 76.8) and GFP mRNA value (515876 ± 23800) were ranked highest. There were 5 promoters (P_{rplJ}, P_{rpsL}, P_{gpmA}, P_{metF}, and P_{ampR}), designated as medium strength promoters, whose RFU ranged from 83 to 402 and GFP mRNA level ranged from 341 to 43237. There were 3 promoters (P_{rpsG}, P_{phlA} and P_{cysP}), designated as low strength promoters, with RFU ranged from 2.65 to 4.13 and GFP mRNA level ranged from 8.81 to 21.51. Overall, we obtained a library of constitutive promoters with various promoter strengths (1 high-strength promoter, 5 medium-strength promoters, and 3 low-strength promoters).

3.3.2 Initial construction of Neu5Ac synthetic pathway

he wild-type *S. marcescens* Db11 strain does not natively produce Neu5Ac (confirmed by HPLC). Since *S. marcescens* Db11 can natively produce *N*-acetylglucosamine, only two exogenous pathways are required to produce Neu5Ac

(Figure 1). Several candidate enzymes have been reported to potentially achieve the desired two biochemical reactions: 1) AGE from *Anabaena* sp. CH1 (Lee et al. 2007) or Slr1975 from *Synechocystis* sp. PCC6803 (Tabata et al. 2002) were two well-characterized *N*-acetylglucosamine 2-epimerases with high reported activity to convert *N*-acetylglucosamine (GlcNAc) to *N*-acetylmannosamine (ManNAc) and 2) Neu5Ac aldolase NanA from *E. coli* MG1655 (Mahmoudian et al. 1997) or a Neu5Ac synthase NeuB from *Campylobacter jejuni* NCTC11168 (Sundaram et al. 2004) for conversion of ManNAc to Neu5Ac. Initially the following three recombinant strains were constructed and tested: P_{T5}-*slr1975*-P_{T5}-*nanA*, P_{rplJ}-*slr1975*-P_{rplJ}-*nanA*, and P_{cysP}-*slr1975*-P_{cysP}-*nanA* based on promoter strengths (high-high, medium-medium, low-low), shown in **Figure 9A**.

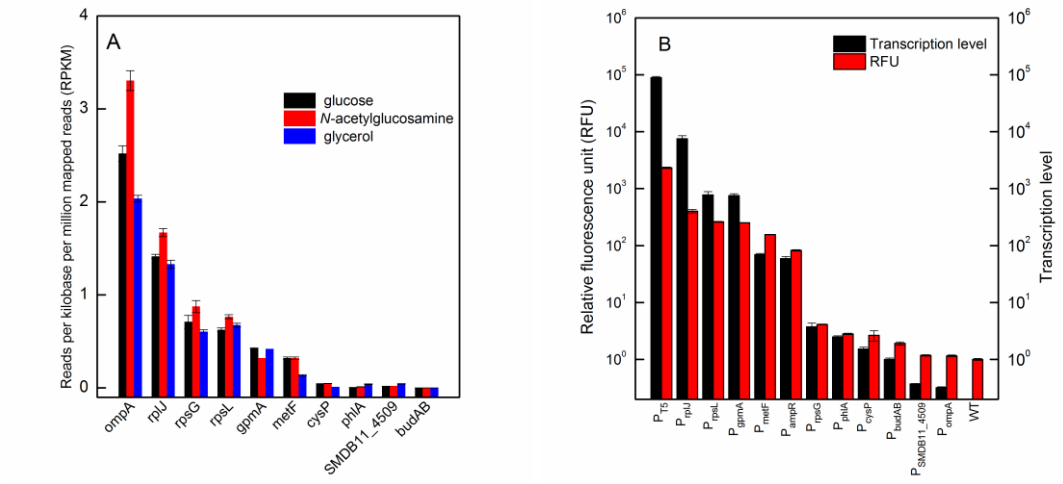


Figure 8 Characterization of *S. marcescens* constitutive promoters. (A) RPKM values of 10 native genes of *S. marcescens* Db11; (B) GFP expression level of *S. marcescens* constitutive promoters by flow cytometry and quantification of GFP mRNA level of *S. marcescens* recombinant strains by real-time PCR.

Real-time PCR was employed to measure the mRNA level of *slr1975* and *nanA* in the three recombinant strains, shown in **Figure 9B**. The wild-type Db11 strain was used

as a negative control where *slr1975* and *nanA* mRNA transcripts were not detected. Compared to the mRNA level of *slr1975* and *nanA* of the P_{cysP}-*slr1975*-P_{cysP}-*nanA* strain, the P_{T5}-*slr1975*-P_{T5}-*nanA* strain showed 29.45- and 28.57-fold increases, respectively, and the P_{rplJ}-*slr1975*-P_{rplJ}-*nanA* strain showed 14.79- and 12.29-fold increases, respectively.

To test if Slr1975 and NanA were functional and to evaluate their activity, a cell-free Neu5Ac production assay was conducted by incubating 2.5 g/L GlcNAc with the cell lysates of the recombinant strains. After 2 h incubation, the P_{T5}-*slr1975*-P_{T5}-*nanA* strain exhibited 154 ± 13.2 mg/L Neu5Ac, the P_{rplJ}-*slr1975*-P_{rplJ}-*nanA* strain exhibited 78.9 ± 8.76 mg/L Neu5Ac, and the P_{cysP}-*slr1975*-P_{cysP}-*nanA* strain showed 34.1 ± 2.45 mg/L Neu5Ac (**Figure 9C**). The *in vitro* Neu5Ac production titer results were consistent with the measured mRNA levels where an increase transcription level by using stronger promoters led to an increase the Neu5Ac production rate.

While using the strong promoter (P_{T5}) to drive expression of both genes led to the highest Neu5Ac production *in vitro*, implementation *in vivo* showed that this design had a strong negative effect on cell growth where the P_{T5}-*slr1975*-P_{T5}-*nanA* strain had a growth rate of μ_{max} 0.19 ± 0.03 /h. This is significantly slower when compared to the P_{rplJ}-*slr1975*-P_{rplJ}-*nanA* strain (μ_{max} 0.53 ± 0.01 /h), the P_{cysP}-*slr1975*-P_{cysP}-*nanA* strain (μ_{max} 0.86 ± 0.02 /h), and the wild-type strain (μ_{max} 0.88 ± 0.01 /h), shown in **Figure 9D**. One possible explanation is that the high overexpression of the two exogenous

genes in the *S. marcescens* P_{T5}-*slr1975*-P_{T5}-*nanA* strain adds significant metabolic burden to the cells, hindering basal metabolic needs that are required to support growth.

3.3.3 Neu5Ac titer improvement by tuning promoter strength

The combination of cellular growth and product yield affect the overall cellular productivity, so an iteration on the strain design was conducted to attempt to balance cell growth and Neu5Ac titer by tuning promoter strengths. The P_{T5}-*slr1975*-P_{rplJ}-*nanA* strain and P_{rplJ}-*slr1975*-P_{T5}-*nanA* strain were created by adapting a high (P_{T5}) and medium strength (P_{rplJ}) promoter combination while retaining the same coding sequences (*slr1975* and *nanA* genes), shown in **Figure 9A**. Compared to the P_{cysP}-*slr1975*-P_{cysP}-*nanA* strain, the mRNA levels of *slr1975* and *nanA* were 25.76- and 8.64-fold, and 10.06- and 27.31-fold for the P_{T5}-*slr1975*-P_{rplJ}-*nanA* strain and P_{rplJ}-*slr1975*-P_{T5}-*nanA* strain, respectively, shown in **Figure 9B**. The *in vitro* Neu5Ac production result showed that the P_{T5}-*slr1975*-P_{rplJ}-*nanA* strain increased 1.7-fold Neu5Ac titer compared to the P_{T5}-*slr1975*-P_{T5}-*nanA* strain and had the highest *in vitro* titer of any of the strains tested to this point; however, the P_{rplJ}-*slr1975*-P_{T5}-*nanA* strain decreased Neu5Ac titer compared to the P_{T5}-*slr1975*-P_{T5}-*nanA* strain (**Figure 9C**). The growth rate of the P_{T5}-*slr1975*-P_{rplJ}-*nanA* strain and P_{rplJ}-*slr1975*-P_{T5}-*nanA* strain were 0.49 ± 0.01 /h and 0.50 ± 0.02 /h, which was significantly improved from the P_{T5}-*slr1975*-P_{T5}-*nanA* strain (**Figure 9D**). The P_{T5}-*slr1975*-P_{rplJ}-*nanA* strain showed the highest *in vitro* titer and reasonable growth rates and thus looked to be the most promising strain at this

point.

While conducting the *in vitro* Neu5Ac production assays, several observations were made related to Neu5Ac production. First, while Neu5Ac was produced within 2 h, the concentration decreased between 2 h and 4 h and reached a plateau after 8 h. Also, detectable amounts of GlcNAc and pyruvate were left after 8 h. These observations were consistent during testing of all the recombinant strains. One possible reason could be an exhaustion of ATP that hinders conversion of GlcNAc to ManNAc. To test this possibility, ATP was fed at 2 h after ATP might have been exhausted. This did lead to an increase in the final Neu5Ac concentration (605.26 ± 38.67 mg/L), shown in **Figure 10B**. Since ATP can be regenerated *in vivo*, we believed that the designs would have sufficient ATP available once implemented in *S. marcescens* to maintain Neu5Ac production. The observation of leftover GlcNAc and pyruvate implied that the *in vitro* system was hitting a chemical equilibrium between forward and reverse reactions since both pathways (*slr1975* and *nanA* genes) were reversible.

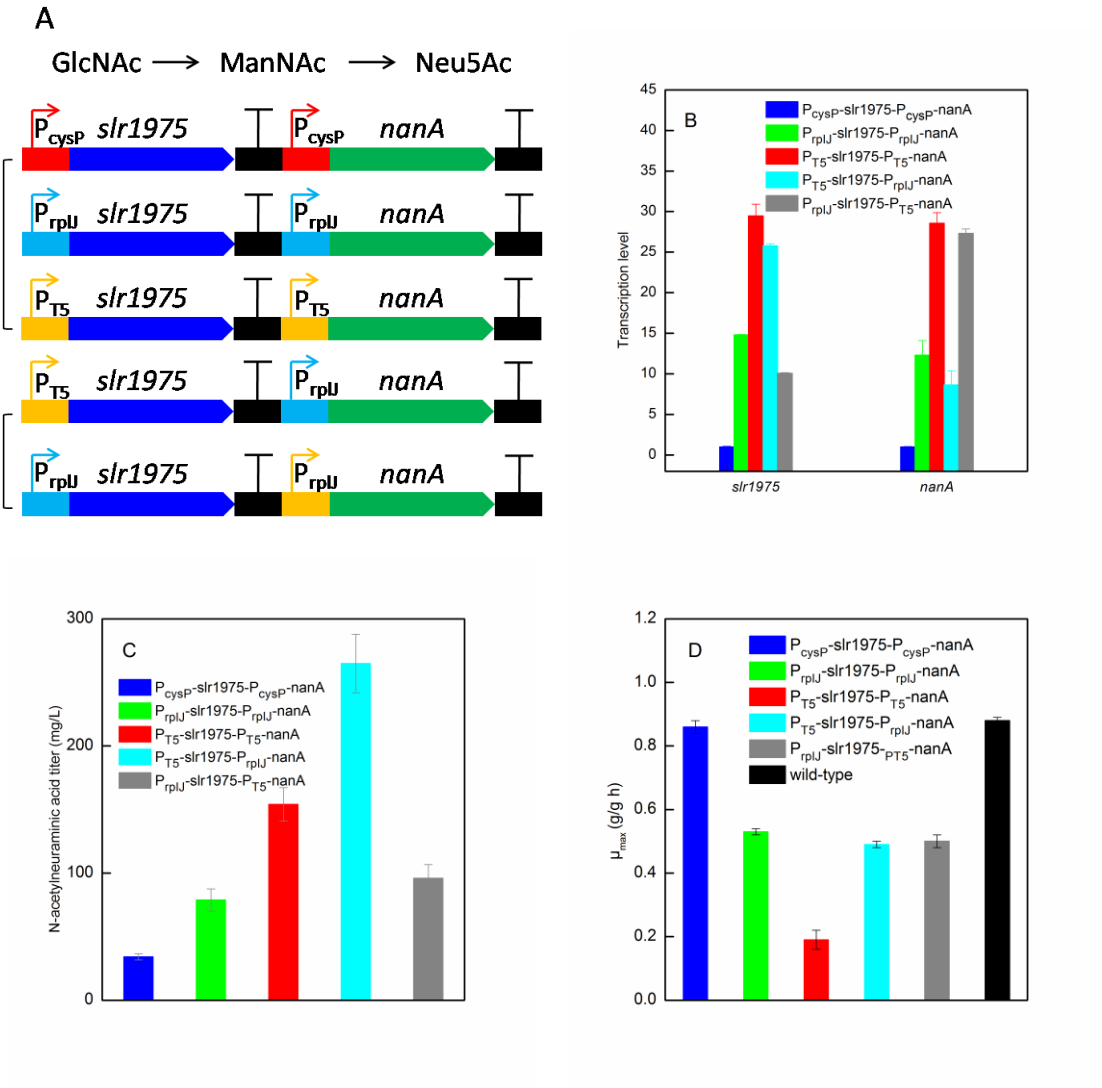


Figure 9 The Neu5Ac production by tuning promoter strength of the pathway genes. (A) The design strategy of tuning promoter strength to improve Neu5Ac production; (B) Quantifications of the mRNA levels of *slr1975* and *nanA* of the *S. marcescens* recombinant strains by real-time PCR; (C) *In vitro* production of Neu5Ac by using the *S. marcescens* recombinant strain cell lysates after 2 h incubation with 2.5 g/L GlcNAc; (D) Cell growth rates of the *S. marcescens* recombinant strains at M9 medium with 20 g/L GlcNAc and 0.1% yeast extract at 30 ℃ and 200 rpm.

3.3.4 Neu5Ac improvement by swapping irreversible enzyme

To further investigate the possible effects of the reversibility of reactions on productivity and details of enzyme activity, two different GlcNAc 2-epimerases (AGE from *Anabaena* sp. CH1 and Slr1975 from *Synechocystis* sp. PCC6803), a Neu5Ac aldolase (NanA from *E. coli* MG1655), and a Neu5Ac synthase (NeuB from

Campylobacter jejuni NCTC11168) were synthesized, produced and purified using *E. coli*.

For the last step in production of Neu5Ac, ManNAc is the precursor compound so *in vitro* assays were run to test activity and reversibility by incubating with either ManNAc or with Neu5Ac as the starting substrate. Assays with NanA produced Neu5Ac from ManNAc with activity of 31.44 ± 0.48 U/mg protein and produced ManNAc from Neu5Ac with activity of 60.08 ± 0.76 U/mg protein, indicating that NanA functions in a reversible manner, shown in **Figure 10B** and **Figure S5**. Assays with NeuB produced Neu5Ac from ManNAc with activity of 95.12 ± 1.32 U/mg protein but no production of ManNAc from Neu5Ac was found, indicating that NeuB converts ManNAc to Neu5Ac in an irreversible manner. For the forward reaction (ManNAc to Neu5Ac), the enzyme activity of NeuB is around 3.0-fold higher than NanA. For the enzyme candidates tested for conversion of GlcNAc to ManNAc, no significant difference of activity or reversibility were found between AGE and Slr1975 (data not shown).

Based upon this information, four additional recombinant strains (P_{T5}-*slr1975*-P_{rplJ}-*nanA*, P_{T5}-*slr1975*-P_{rplJ}-*neuB*, P_{T5}-*AGE*-P_{rplJ}-*nanA*, and P_{T5}-*AGE*-P_{rplJ}-*neuB*) were designed and built to test the different enzymes in combination, shown in **Figure 10A**. After 60 h *in vitro* reaction with supplemented sufficient ATP, Neu5Ac production reached a plateau. A maximum titer of 1.5 ± 0.09 g/L Neu5Ac was obtained from the P_{T5}-*slr1975*-P_{rplJ}-*neuB* construct, that represents

approximately a 2.5-fold increase above the previous best results of the P_{T5}-*slr1975*-P_{rplJ}-*nanA* construct, shown in **Figure 10C**. This result indicated that Neu5Ac production could be significantly improved after swapping the reversible enzyme NanA to the irreversible enzyme NeuB. After swapping the first step enzyme from Slr1975 to AGE, there was no significant Neu5Ac production difference, which is in consistent with having no appreciable difference in enzymatic activity results.

3.3.5 *In vivo* Neu5Ac production

All of the *S. marcescens* recombinant strains and the wild-type strain Db11 were initially grown on 20 g/L GlcNAc to test for Neu5Ac production to test strain functionality without the complicating factor of chitinase activity. After 48 h of growth, GlcNAc was almost depleted for the wild-type strain Db11 and all recombinant strains except for the P_{T5}-*slr1975*-P_{T5}-*nanA* strain (**See Figure S6A**). The cell growth and metabolic end-products of all *S. marcescens* strains were monitored (**See Figure S6B and Table S11**). Neither intracellular nor extracellular Neu5Ac was detected during cell culture of the wild-type Db11 strain as expected for the negative control. A maximum titer of 0.48 ± 0.06 g/L Neu5Ac was produced by the P_{T5}-*slr1975*-P_{rplJ}-*neuB* strain (**Figure 11**). Compared to the Neu5Ac titer of the initial P_{T5}-*slr1975*-P_{T5}-*nanA* strain, the Neu5Ac production was significantly improved by about 11.7-fold after balancing the pathways using promoter swapping and switching the final reaction step to use a faster, irreversible enzyme (*neuB*).

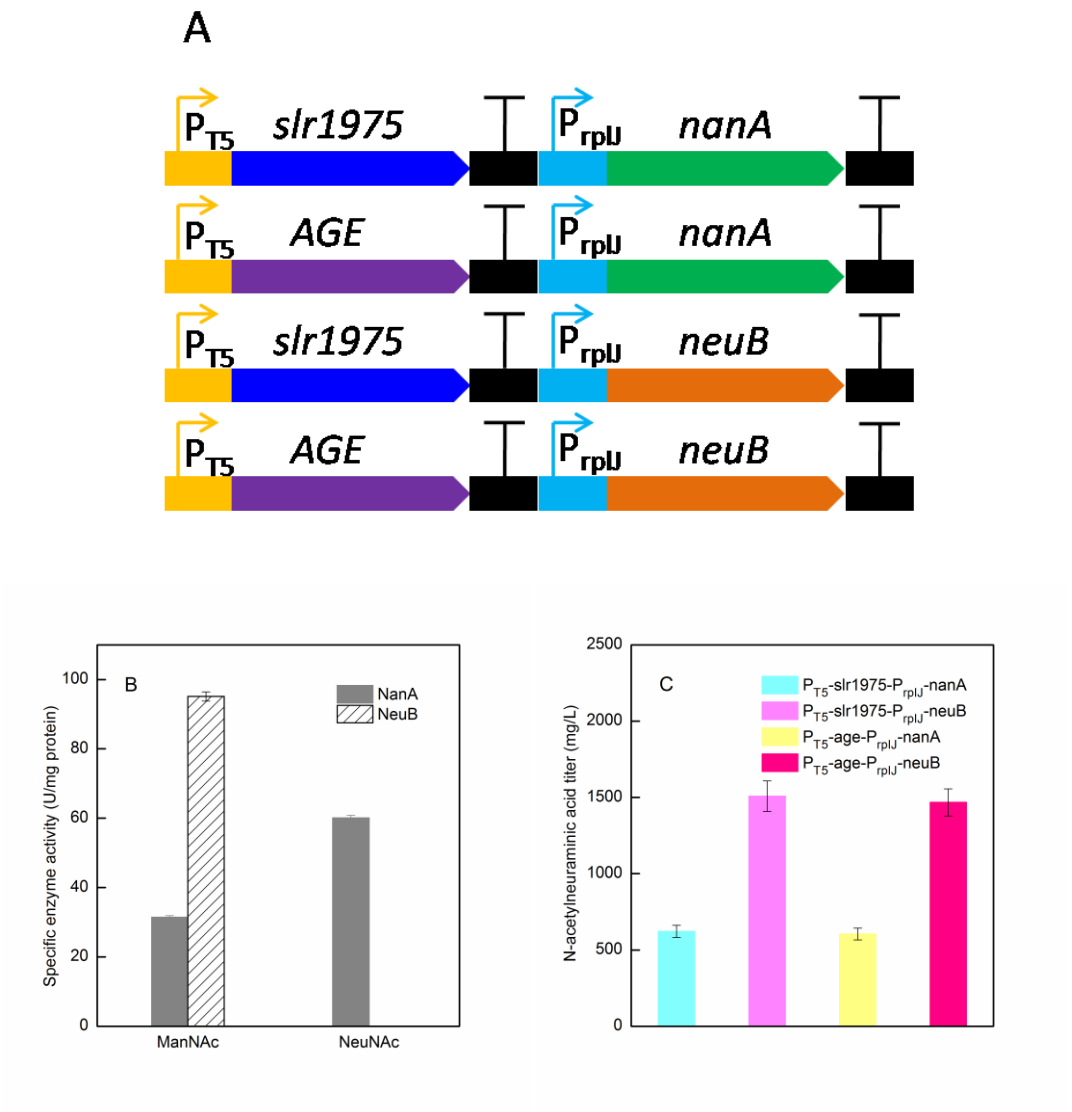


Figure 10 The Neu5Ac production by utilizing different enzyme candidates. (A) The design strategy of different activity enzymes to improve Neu5Ac production; (B) Specific enzyme activities of NanA and NeuB by feeding either ManNAc or Neu5Ac as a substrate; (C) *In vitro* production of Neu5Ac by using the *S. marcescens* recombinant strains cell lysates from 2.5 g/L GlcNAc after 48 h incubation with supplement of 10 mM ATP at 2 h.

3.3.6 Neu5Ac production directly from chitin

The final component was to test for production of Neu5Ac for growth directly on chitin. The *S. marcescens* P_{T5}-*slr1975*-P_{rplJ}-*neuB* strain and the wild-type strain were grown on 5 g/L, 10 g/L and 20 g/L crystal chitin or 5 g/L, 10 g/L and 20 g/L colloidal chitin. After 1 week culture, the *S. marcescens* P_{T5}-*slr1975*-P_{rplJ}-*neuB* strain produced

0.32 ±0.02 g/L and 0.30 ±0.05 g/L from 5 g/L crystal chitin and 5 g/L colloidal chitin, respectively, shown in **Table 5** and **Figure 12**. Overall, our results showed that the engineered *S. marcescens* strain was capable of producing Neu5Ac directly from chitin and the maximum titer is on the same order of magnitude as reported from glucose as a carbon source in *E. coli*.

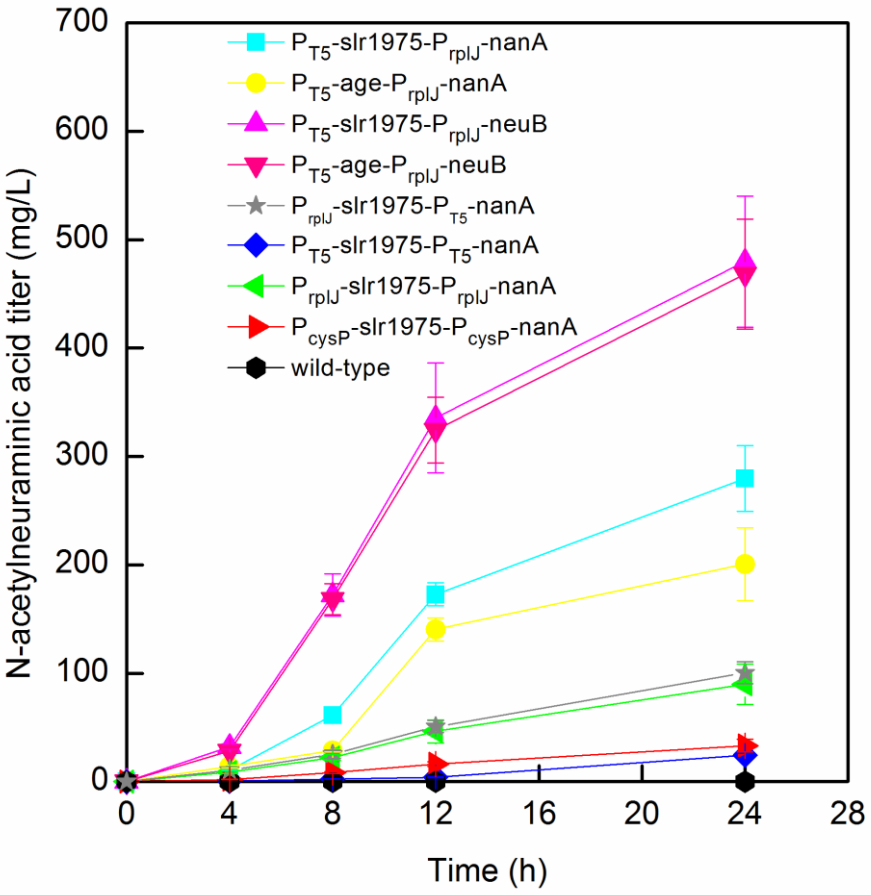


Figure 11 *In vivo* Neu5Ac production from the *S. marcescens* recombinant strains at M9 medium with 20 g/L GlcNAc and 0.1% yeast extract at 30 ℃ and 200 rpm.

Table 5 Neu5Ac production from crystal chitin and colloidal chitin using *S. marcescens* P$_{T5}$-*slr1975*-P$_{rplJ}$-*neuB* and the wild-type Db11 strain after 1-week

fermentation in a rotator flask.

Strain	Substrate	Neu5Ac (g/L)	Biomass (g/L)	2,3-BD (g/L)	Acetic acid (g/L)	Lactic acid (g/L)
S. marcescens P_{T5}-*slr1975*-P_{rplJ}-*neuB*	crystal chitin (20 g/L)	0.37 ±0.02	0.83 ±0.05	0.20 ±0.09	0.13 ±0.01	0.08 ±0.00
	colloidal chitin (20 g/L)	0.33 ±0.05	0.87 ±0.04	0.18 ±0.10	0.11 ±0.01	0.06 ±0.00
S. marcescens Db11	crystal chitin (20 g/L)	-	1.08 ±0.01	0.40 ±0.04	0.26 ±0.03	0.15 ±0.01
	colloidal chtin (20 g/L)	-	0.99 ±0.03	0.38 ±0.02	0.23 ±0.02	0.10 ±0.01
S. marcescens P_{T5}-*slr1975*-P_{rplJ}-*neuB*	crystal chitin (10 g/L)	0.30 ±0.06	0.80 ±0.04	0.19 ±0.02	0.12 ±0.00	0.06 ±0.00
	colloidal chitin (10 g/L)	0.32 ±0.05	0.82 ±0.03	0.16 ±0.02	0.10 ±0.00	0.06 ±0.00
S. marcescens P_{T5}-*slr1975*-P_{rplJ}-*neuB*	crystal chitin (5 g/L)	0.32 ±0.05	0.81 ±0.03	0.20 ±0.02	0.11 ±0.00	0.07 ±0.00
	colloidal chitin (5 g/L)	0.30 ±0.03	0.82 ±0.05	0.18 ±0.02	0.12 ±0.00	0.06 ±0.00

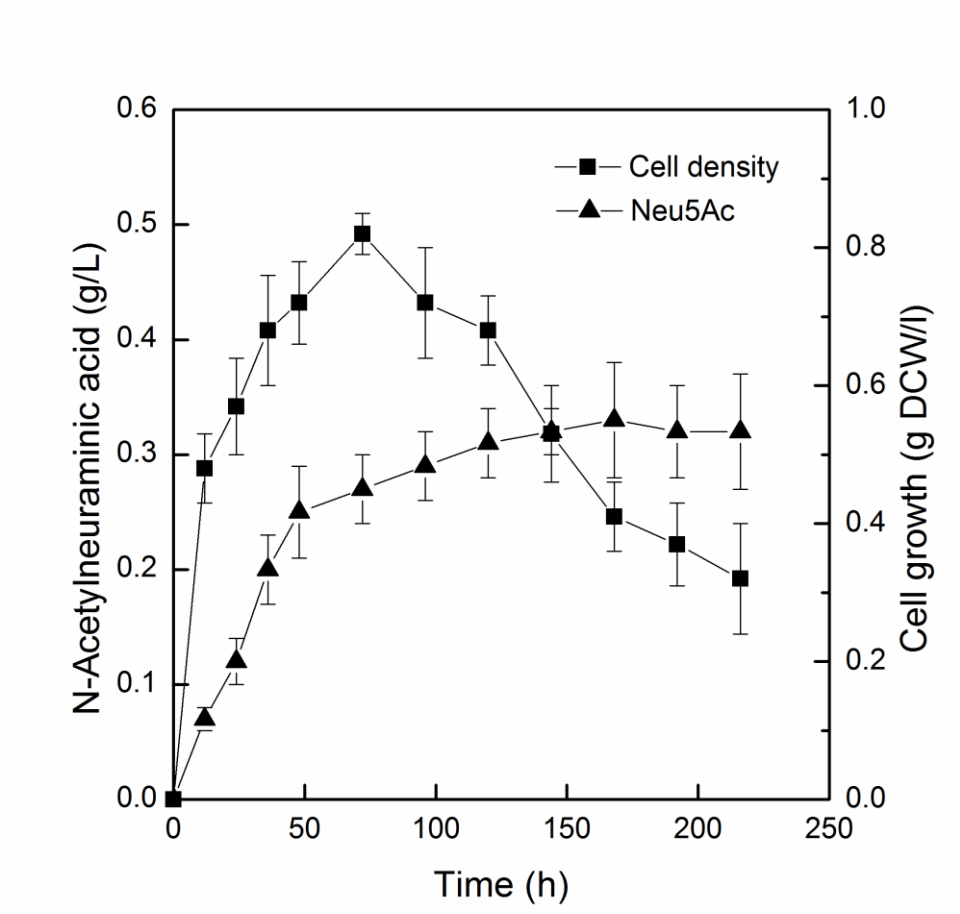


Figure 12 *N*-Acetylneuraminic acid production and cell growth under 5 g/L crystal chitin M9 medium.

3.4 Discussion

When considering the application of bioprocessing for industrial-scale production of chemicals, a diversity of issues are involved. A large and growing body of work utilizes the main workhorse microorganisms (*Escherichia coli*, *Saccharomyces cerevisiae*, *Bacillus subtilis*) for many applications, but it is becoming necessary to develop additional platform organisms that can utilize specialized cellular activities or functions that are not found in the established workhorse organisms. In this study, we have focused on the chitinolytic organism, *Serratia marcescens*, to demonstrate its potential utility as a platform strain for chitin-based CBP applications. Overall, we

have engineered *S. marcescens* to directly convert crystal chitin to *N*-acetylneuraminic acid. Native constitutive promoters for *S. marcescens* were identified and tested (1 high-strength promoter, 5 medium-strength promoters, and 3 low-strength promoters) that were used for this study and can help future engineering of *S. marcescens*. Here, the promoters were applied for Neu5Ac production in *S. marcescens* by overexpressing two exogenous genes. Tuning the mRNA levels of the pathway genes resulted in an 11.56-fold Neu5Ac titer increase that balanced pathways flux and growth considerations. Furthermore, pulling the reaction equilibrium toward Neu5Ac by use of a fast, irreversible enzyme further increased Neu5Ac titer 1.72-fold. Finally, the optimized strain (P_{T5}-*slr1975*-P_{rplJ}-*neuB*) was able to produce 0.48 ± 0.06 g/L Neu5Ac from 20 g/L GlcNAc and 0.32 ± 0.05 g/L Neu5Ac was produced from 5 g/L crystal chitin.

Heterologous expression of genes and tuning gene expression levels in *S. marcescens* is limited due to the lack of a well-characterized promoter library and additional genetic tools. Use of RNA-Seq data from various growth conditions (substrates, temperature, culture time) can provide useful information to identify native constitutive promoters and it has been applied to balancing pathways to improve production titers and to activating cryptic pathways to produce novel chemicals (Li et al. 2015; Luo et al. 2015; Sun et al. 2012a). It has been found that the wild-type *S. marcescens* has mechanisms for catabolite repression by various carbon sources (Monreal and Reese 1969). To account for catabolite repression, we generated

RNA-Seq data for growth on glucose, *N*-acetylglucosamine, and glycerol. Promoters associated with 10 genes were targeted based upon the RNA-Seq data due to the similarity on the expression level at the three substrates. To build a promoter library, two non-native promoters (P_{T5} and P_{ampR}) were also considered and overall, 12 promoters were characterized (transcriptional and translational levels) resulting in a verified promoter collection containing 1 high-strength promoter, 5 medium-strength promoters and 3 low-strength promoters. Previous studies have used P_{budAB} as the primary native promoter in *S. marcescens*. It should be noted that the P_{rplJ} promoter had a 210.7-fold increase in transcription over the P_{budAB} promoter and is thus the highest reported native *S. marcescens* promoter to date.

Promoter engineering as a strategy to improve chemical production by relieving metabolic burden (Wu et al. 2016), minimizing unnecessary mRNA (Blazeck et al. 2012; Jones et al. 2015), and pulling more metabolic flux toward target chemicals (Ma et al. 2011; Nowroozi et al. 2014; Pitera et al. 2007; Xu et al. 2011), has been implemented in many applications (Alper et al. 2005). In this study, we found that balancing the promoter strengths of the pathway genes can improve Neu5Ac production (11-fold) for two reasons: 1) It can increase cell growth by minimizing unnecessary mRNA production and reducing metabolic burden (cell growth rate increased 2.5-fold) and 2) It can help balance pathway fluxes to pathways in series when there an inbalance in enzyme activity. In *S. marcescens*, after GlcNAc is transported into the cell, there is an interconversion reaction between GlcNAc-6P and

GlcNAc. The GlcNAc is either used in glycolysis as glucosamine 6-phosphate (GlcN-6P) for cell growth or flows to the Neu5Ac synthetic pathway by converting GlcNAc to ManNAc. Future work to enhance flux to the first-step reaction can reinforce the flux toward the intermediate ManNAc formation and favor accumulation of Neu5Ac.

Targeting specific enzyme activity to improve microbial biochemical production by switching a more efficient enzyme at a key-step reaction has been demonstrated in other studies (Bormann et al. 2014; Shen et al. 2011; Yan and Fong 2017a). For the case of Neu5Ac synthesis, the bioprocess involving the reversible enzyme, Neu5Ac aldolase (NanA), usually requires feeding of high concentration of GlcNAc (~ 200 g/L) with a large amount of residual GlcNAc (Kang et al. 2012; Lin et al. 2013; Tao et al. 2011) to skew the chemical equilibrium. The last step of Neu5Ac formation was also reported as a bottleneck step of the bioconversion process due to the lower activity of the aldolase (NanA 0.57 ± 0.09 U/mg protein) than the activity of epimerase (AGE 37.1 ± 3.6 U/mg protein) (Lin et al. 2013). Yang et al. investigated the Neu5Ac production by engineered *E. coli* from the glucose pathways and found that Neu5Ac titer can be improved by 1.33-fold after swapping NeuB to NanA (Yang et al. 2017). In this study, we found that NeuB is more efficient and superior than NanA at both reaction mechanism (irreversible versus reversible) and activity (95.12 ± 1.32 U/mg protein versus 31.44 ± 0.48 U/mg protein) for Neu5Ac synthesis. The P_{T5}-*slr1975*-P_{rplJ}-*neuB* strain produced 1.58-fold more Neu5Ac *in vivo* compared to

the P_{T5}-*slr1975*-P_{rplJ}-*nanA* strain. Future experiments on protein engineering of the NeuB could be conducted to investigate NeuB's effect on improving Neu5Ac production on a molecular basis.

The engineered *S. marcescens* P_{T5}-*slr1975*-P_{rplJ}-*neuB* from this work can produce Neu5Ac from GlcNAc, colloidal chitin, and crystal chitin. The Neu5Ac maximum titers achieved by our designed strain was 0.48 g/L from GlcNAc. 0.33 g/L from colloidal chitin, and 0.37 g/L from crystal chitin. The Neu5Ac titer achieved by our designed strain is on the same order of magnitude as the highest reported from *E. coli* using fermentation approaches (1.62 g/L from 30 g/L glucose) (Kang et al. 2012). The Neu5Ac titer obtained by our designed strain (0.37 g/L from crystal chitin) is the highest reported so far (Steiger et al. 2011).

3.4 Conclusion

In this study, we demonstrated the direct conversion of untreated chitin to *N*-acetylneuraminic acid (Neu5Ac) using an engineered chitinolytic bacterium, *Serratia marcescens*. To add to the genetic tools and facilitate engineering of *S. marcescens*, we used RNASeq data to identify and characterize a set of native constitutive promoters. Production of Neu5Ac was evaluated by functional transcription and protein activities after heterologous expression of *N*-acetylglucosamine 2-epimerase (Slr1975) and Neu5Ac aldolase (NanA). Subsequently, Neu5Ac production titers were improved 11.56- fold by promoter engineering to minimize unnecessary mRNA production and to pull carbon flux

toward ManNAc accumulation. Further optimization was achieved (1.56-fold increase) by adapting an irreversible Neu5Ac synthase (NeuB) to push the reaction equilibrium toward Neu5Ac. The optimized recombinant strain P_{T5}-*slr1975*-P_{rplJ}-*neuB* was able to produce 0.48 g/L Neu5Ac from 20 g/L *N*-acetylglucosamine and 0.32 g/L Neu5Ac from 5 g/L crystal chitin. This is the first study to report *de novo* production of Neu5Ac from chitin in *S. marcescens*.

Chapter 4 Adaptive evolution *S. marcescens* to increase chitin utilization efficiency and genome resequencing platform reveals insightful mechanism

Importance of this part of work:

I have demonstrated production of 2,3-butanediol and *N*-acetylneuraminic acid from chitin using *S. marcescens*. One of the challenges when considering *S. marcescens* as a workhorse for consolidated bioprocessing is that degradation of recalcitrant chitin is a long process. Laboratory adaptive evolution has been proved to be a powerful tool for improving organisms' substrate utilization. The fact that *S. marcescens* can grow on chitin makes it an ideal candidate to conduct adaptive evolution experiments. Since chitinolytic organism are not fully characterized and identified, to the best of my knowledge, this is the first of study to conduct adaptive evolution research using chitin as a substrate.

The conclusions and fundamental results derived from this part of work:

I conducted serial passage of *Serratia marcescens* in colloidal chitin M9 medium for four months, propagating 288 generation, yielding an evolved strain *S. marcescens* EPS. Physiological changes including increased chitinase activities, decreased secreted proteins, increased cell yield on chitin, and metabolic end-products changes from the wild-type strain to the EPS strain were observed. Cellular morphological characterization showed the EPS strain exhibits a large cell size (length and height). The EPS strain was capable of producing 1.41 g/L 2,3-butanediol from 2% crystal chitin, a 3.7-fold increase compared to that of the wild-type strain. In addition, the molar ratio of 2,3-butanediol to byproducts is improved from the wild-type strain

(2.5:2.5:1) to the EPS strain (4.8:2.7:1). Key mutations occurred in two gene products (SMDB11_2700 and SMDB11_1516) that encode a LysR-family regulator and a fimbrial adhensin may be accounting for the increased chitinase activities and overproduction of 2,3-butanediol.

4.1 Introduction

Chitin is the second most common organic compound on earth and can be potentially used as an inexpensive and renewable raw material for many applications (Yan and Chen 2015; Yan and Fong 2015). One of potentially interesting schemes to utilize chitin for biochemical production is using chitinolytic organisms to perform one-step conversion, known as consolidated bioprocessing (CBP) (Yan and Fong 2017a). To more fully utilize chitin, efforts are being made to identify, characterize, and utilize chitinolytic organisms. Collectively, chitinolytic microorganisms are a small fraction of overall known species with different species found in many different branches of life such as *Serratia marcescens* (Vaaje-Kolstad et al. 2013), *Bacillus circulans* (Jee et al. 2002), and *Acinetibacter parvus* (Kim et al. 2017). While generally poorly characterized, chitinolytic microbes possess great physiological and biochemical diversity making them interesting organisms to study from both a basic and applied perspective. *Serratia marcescens* is a particularly interesting chitinolytic bacterium that has been well-studied and more efficient at producing chitinases than most other chitinolytic bacteria.

One of the main challenges to utilizing chitin is that chitin degradation/hydrolysis is a

slow and difficult process. The roles, activity, and regulatory mechanisms controlling expression of different chitinases are all being studied with a goal of being able to understand and increase the rate of chitin hydrolysis (Vaaje-Kolstad et al. 2013; Yan et al. 2017). Another potential approach to increase the rate of chitin hydrolysis is to conduct adaptive evolution experiments with a chitinolytic organism. By serial passage of a microorganism growing in one constant substrate medium, a mutated daughter cell whose growth outpaces other cells can eventually take over the whole population by its own daughter cells. Laboratory adaptive evolution has been applied to many cellulolytic microorganisms as a means for inducing systemic changes to improve cell growth on cellulosic materials (Deng and Fong 2011a; Lin et al. 2016; Patyshakuliyeva et al. 2016). In cases where selection is based upon growth rate that is related to increased consumption of the limiting substrate, thus, laboratory evolution may be a means for increasing the rate of chitin hydrolysis in a chitinolytic organism.

In this study, our goal is to improve *S. marcescens* chitin utilization efficiency by propagating *S. marcescens* Db11 in a minimum M9 medium containing colloidal chitin as a sole carbon source. After 4-month laboratory adaptive evolution experiments, an end-point evolved strain was isolated with clear cellular physiological changes including increase of chitinase activities and cell growth, decrease of overall secreted proteins, and changes of secreted end-point metabolites and cellular morphology. Whole genome sequencing of the evolved strains was performed to seek

single nucleotide variations (SNVs) and identified key mutations in a regulator gene upstream region and a fimbrial adhensin gene. In addition, the evolved strain was capable of improving 2,3-butanediol production by 3.7-fold from 2% crystal chitin compared to that of the wild-type strain.

4.2 Materials and Methods

4.2.1 Adaptive evolution of *S. marcescens*

Starting with the wild-type strain of *S. marcescens* Db11, whose genome has recently been sequenced (Iguchi et al. 2014), laboratory evolution experiments were conducted for growth in M9 medium with 2% colloidal chitin as a sole carbon source for four months. The *S. marcescens* strains were serial passage into the fresh medium while cells remained in exponential growth. The phenotypes (cell growth, chitinase activity and overall secreted proteins) were measured every five days and the amount of dilution at each passage was adjusted daily to account for changes in growth rate. After around four months of cell propagation, the cell growth and chitinase activity were stable and the laboratory evolution experiment was stopped. The end-point strains were propagated for nearly 287 generations and the middle-point strains were propagated for nearly 143 generations. For subsequent testing, isogenic colonies were selected by diluting *S. marcescens* on agar plates containing M9 medium with 2% colloidal chitin and picking single colonies for culturing.

4.2.2 Growing *S. marcescens* with glucose, *N*-acetylglucosamine and colloidal chitin

S. marcescens strains preculture were grown in LB medium at 30℃ and 220 rpm for

overnight. The preculture was harvested and washed three times using purified water. Then, 2.5% (v/v) inoculation was added in a 50 mL Erlenmeyer flask containing 35 mL M9 medium with 10 g/L glucose, 10 g/L *N*-acetylglucosamine, or 20 g/L colloidal chitin. The *S. marcescens* strains were incubated at 30 ℃ and 100 rpm for 48 h or 7 days, respectively. The cell growth were monitored either by testing OD600 every 4 hours when grown in glucose or *N*-acetylglucosamine M9 medium or by measuring overall cell lysates protein concentration every 24 hours when grown in colloidal chitin M9 medium.

4.2.3 Cell density measurement

The optical density (OD) was measured as absorbance at 600 nm every 4 h for *S. marcescens* growing on glucose and *N*-acetylglucosamine M9 medium. Due to the physiology of colloidal chitin (milk white turbid liquid), the cytoplasmic protein content of *S. marcescens* strains when grown in colloidal chitin medium was tested according to the protocol described in our previous publication (Yan et al. 2017). In brief, 1 mL cell culture was centrifuged at 10,000 ×*g* for 5 min. The cell pellets were re-suspended in fresh media and repeated centrifugation at 10,000 ×*g* for 5 min. The samples were resuspended in 100 μL 1X Bugbuster (EMD Millipore, Billerica, MA) and lysed by incubating at room temperature for 30 min. Cell wall debris were removed by centrifuging at 10,000 ×*g* for 5 min and the supernatant was used to measure cytoplasmic protein concentration by the Bradford protein assay (Bio-Rad, Hercules, CA). The dry cell weight (DCW) is proportionally related to the overall protein content according to a correlation factor equation: DCW (g/L) = 1.20 ± 0.03 ×

Protein (g/L)).

4.2.4 Secreted end-products measurement

1 mL cell culture was harvested and centrifuged at 10,000×*g* for 5 min. The supernatant was filtered using a 0.22 μm filter and subsequently, the samples were ready for HPLC measurements. Metabolic end-products were quantified using an Ultimate3000 HPLC system (Dionex, Sunnyvale, CA), equipped with a Bio-Rad HPX-87H ion exclusion column (Hercules, CA) and a refractive index detector (Shodex, Japen). The condition was run with 5 mM H_2SO_4 at 0.6 mL/min and 55 ℃.

4.2.5 Chitinase activity assay

Three substrates were used to measure the chitinase activities according to different enzyme acting mechanisms. For instance, 0.2 mg/mL 4-Nitrophenyl β-D-*N*,*N*′,*N*″-triacetylchitotriose was used to detect endochitinase activity; 0.5 mg/mL 4-Nitrophenyl *N*,*N*′-diacetyl-β-D-chitobioside was used to determined chitobiosidase activity; 1 mg/mL 4-Nitrophenyl *N*-acetyl-β-D-glucosaminide was used to test β-*N*-acetyl-glucosaminidase activity. Briefly, 1 mL cell culture was centrifuged at 10,000×*g* for 5 min. Each reaction contained 10 μL supernatant samples and 90 μL substrate dissolved in assay buffer (Sigma-Aldrich, St. Louis, MO). The reaction was incubated in a 96-well plate at 37 ℃ for 30 min and was terminated by adding 200 μL 0.4 M sodium carbonate solutions. The absorbance at 405 nm was measured immediately using a microplate reader (VERSAmax, San Jose, CA). Chitinase activity was defined as one unit of activity will release 1.0 μmole of 4-Nitrophenol from appropriate per minute at pH 4.8 at 37 ℃.

4.2.6 Cell morphological characterization

All *S. marcescens* strains cell morphology was conducted and characterized using Atom Force Microscope (AFM). In brief, *S. marcescens* strains were streaked from freezer stocks on agar plates containing M9 medium with 2% colloidal chitin and were incubated at 30 ℃ for 1 week. After 1 week culture, *S. marcescens* colonies grew saturated on the agar plates and formed clear zone around them. A single isogenic colony was picked up and diluted in a 1.5 mL centrifuge tube containing 1 mL DI H_2O. One drop of the cell resuspension solution was placed on the surface of a sample plate. When the samples were dried at room temperature, they were readily visualized and characterized by an AFM. Detailed procedures of AFM characterization were according to a previous publication by a research group at VCU (Wang et al. 2015).

4.2.7 Whole genome resequencing

Genomic DNA was isolated using the DNeasy Blood & Tissue kit (Qiagen, USA) according to the manufacturer's instructions. All the genomic sequencing libraries were prepared according to the manufacturer's instructions. We sequenced the isolates from *S. marcescens* MPS and EPS using sequencer and then these two isolates were also sequenced by Illumina Hiseq3000. In all cases we generated pair-end reads.

4.2.8 Single nucleotide variations (SNVs) detection from sequence data

Sequencing reads obtained using the Roche 454 FLX platform were analyzed by manufacturer-supplied software and processed by a Technology Center for Genomics and Bioinformatics at University of California at Los Angelas. Illumina reads were mapped directly to the *S. marcescens* reference sequence using a somatic mutation

calling program. During the alignment of the Illumina data, sequence fragments were called to a unique genome position and only fragments with a perfect sequence match or a single base mismatch were used. In general, there was an average coverage of 50 for each gene. The reference genome sequence of *S. marcecens* Db11 is HG326223.

4.2.9 2,3-butanediol fermentation

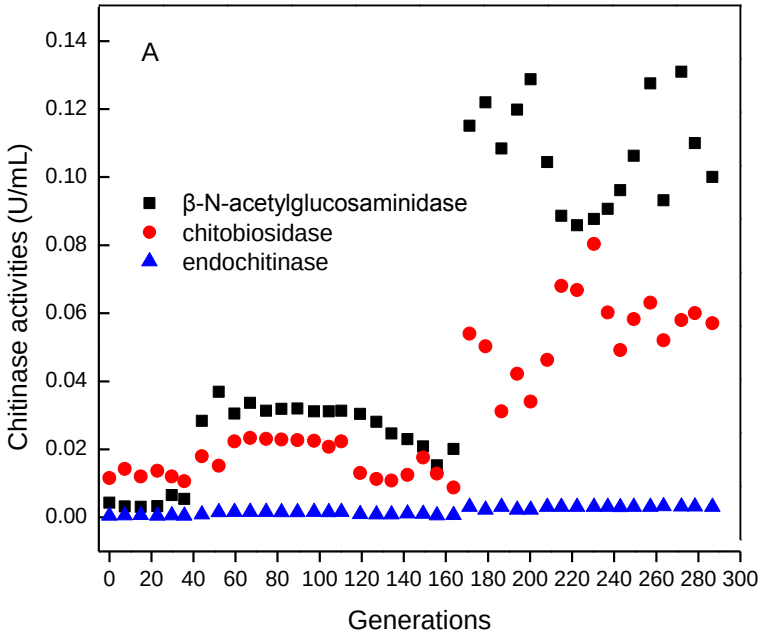
S. marcescens pre-culture was performed in LB medium at 30℃ and 220 rpm for overnight. 10% (v/v) seed culture was inoculated for 2,3-butanediol fermentation. *S. marcescens* 2,3-butanediol fermentations were conducted in 250-mL Erlenmeyer flask containing 150 mL M9 minimum medium with 1% yeast extract, 5 g/L N-acetylglucosamine, and 20 g/L crystal chitin at initial pH 7.5, 30℃ and 100 rpm.

4.3 Results and Discussion

4.3.1 Adaptive evolution of *S. marcescems*

After around four-month serial passage of cells at exponential phage into fresh M9 colloidal chitin medium, the end-point strain, designated as *S. marcescens* EPS, was propagated for ~287 generations while the middle-point strain, designated as *S. marcescens* MPS was propagated for ~143 generations. By monitoring the overall population phenotypes, we observed the changes of cell yields, chitinase activities and overall secreted proteins tend to be steady (**Figure 13A, 13B and 13C**). Compared to the parental strain, after 287 generations propagation, the end-point populations showed a trend of increased chitinase activities (β-*N*-acetylglucosaminidase, from 4.259×10^{-3} to 0.10 U/mL; endochitinase, from 4.77×10^{-4} to 0.0031 U/mL; chitobiosidase from 0.0116 to 0.057 U/mL); a trend of decreased overall secreted

protein (from 71.36 mg/L to 51.63 mg/L); and a trend of increased overall cell yield (from 6.06 g/L to 7.26 g/L periplasmic protein). Both increases of chitinase activities and cell yields indicate a potential improvement of S. marcescens chitin utilization capability.



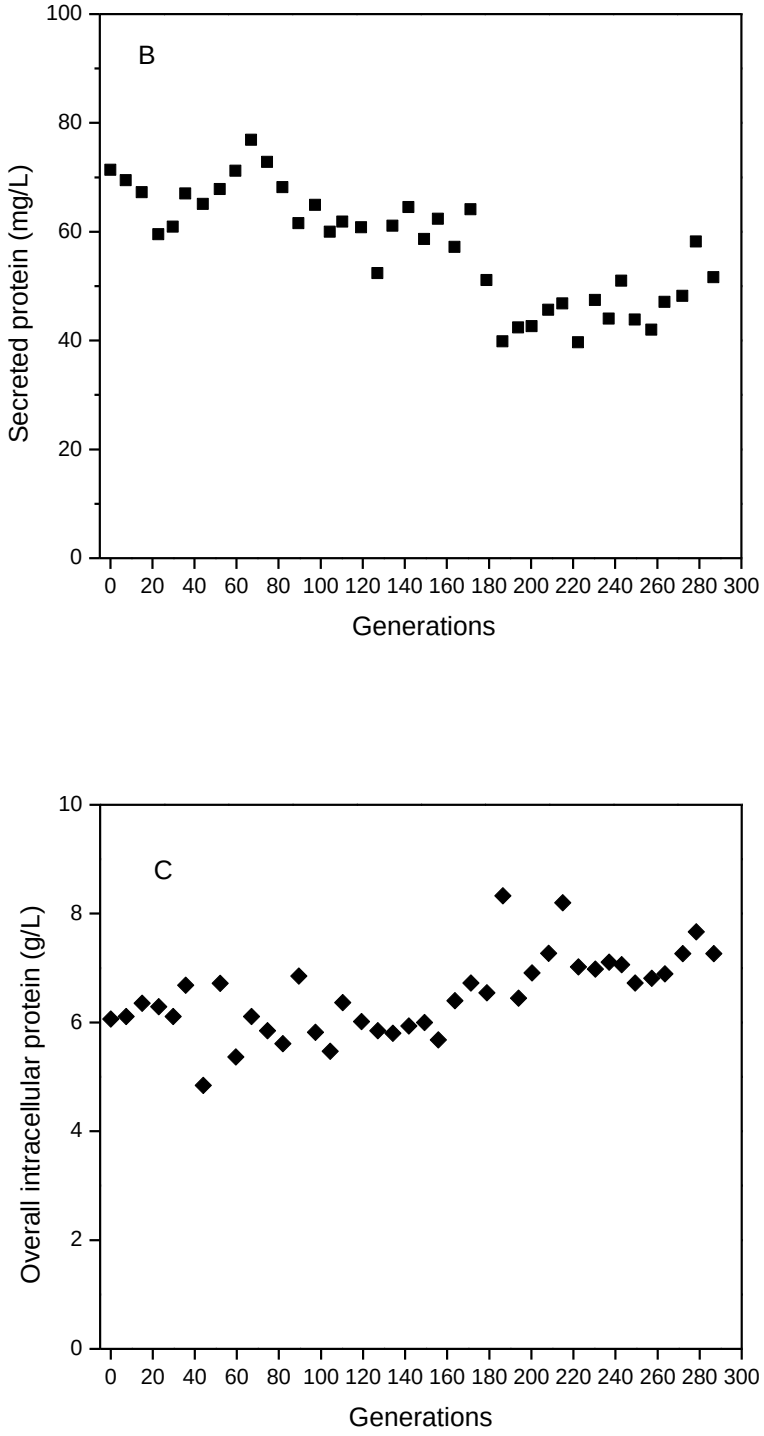
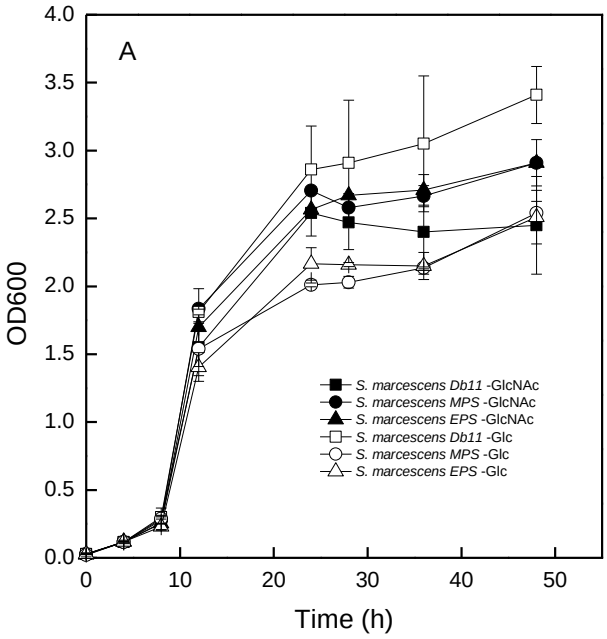


Figure 13 Physilogical changes after serial passage of *S. marcescens* in 2% colloidal chitin M9 medium. (A) Chitinase activities (U/mL): β-*N*-acetylglucosaminidase (black square), chitobiosidae (red circle), endochitinase (blue triangle); (B) Overall secreted proteins (mg/L); (C) Overall intracellular protein concentration.

4.3.2 Growth changes

Next, we isolated the evolved middle-point strain (MPS) and end-point strain (EPS) by diluting on colloidal chitin agar plates. Ten colonies for 143 and 287 generation populations were picked and tested for chitinase activities, cell growth and secreted overall protein (data not shown). The chitinase activities were consistent to the mixed population trend. Thus, we isolated two isogenic strains: a *S. marcescens* MPS and a *S. marcescens* EPS. The two strains in combination with the wild-type Db11 strain were used to conduct physiological characterization.

Compared to cell growth of wild-type Db11 strain grown in glucose medium, the *S. marcescens* MPS and *S. marcescens* EPS showed slightly decreased cell yields (**Fig. 14A**). Compared to cell growth of wild-type strain grown in *N*-acetylglucosamine medium, the *S. marcescens* MPS and the *S. marcescens* EPS growth did not change apparently. Compared to the cell yield (2.52 g/L) of wild-type strain grown in colloidal chitin medium, the *S. marcescens* MPS and *S. marcescens* EPS exhibited an increased overall cell yield (3.69 g/L and 3.89 g/L, respectively) (**Fig. 14B**). The increase of overall cell yields from wild-type strain to *S. marcescens* MPS and EPS are consistent to the overall population trend.



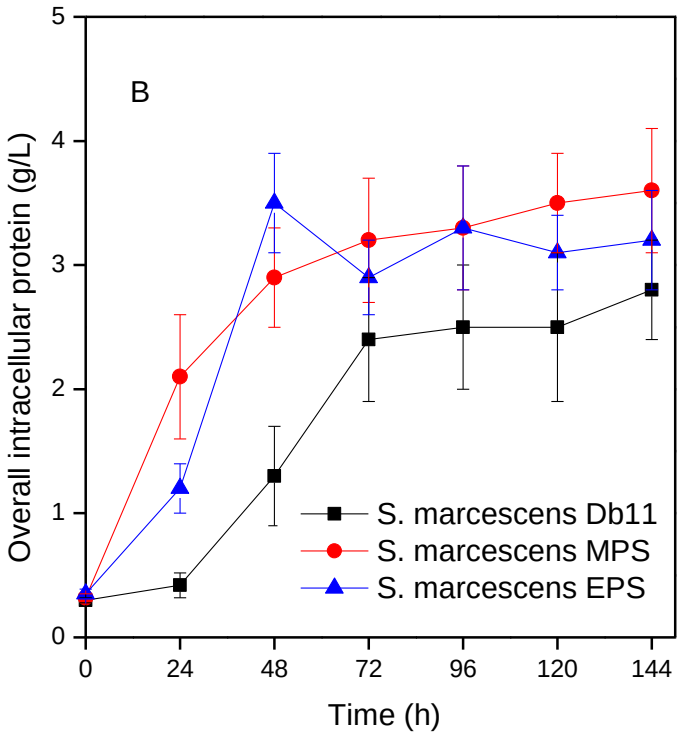


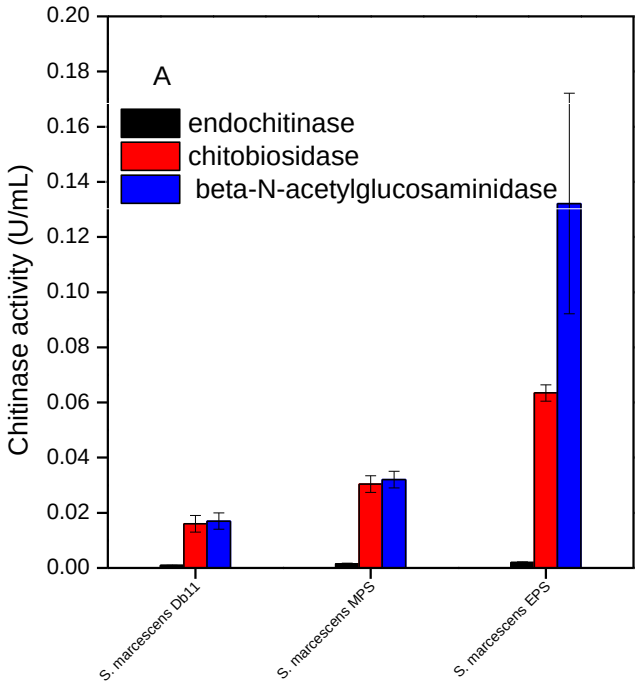
Figure 14 Time course of cell growth in M9 medium of different carbon souces at 220 rm and 30℃. (A) 10 g/L glucose and 10 g/L *N*-acetylglucosamine, (B) 2% colloidal chitin.

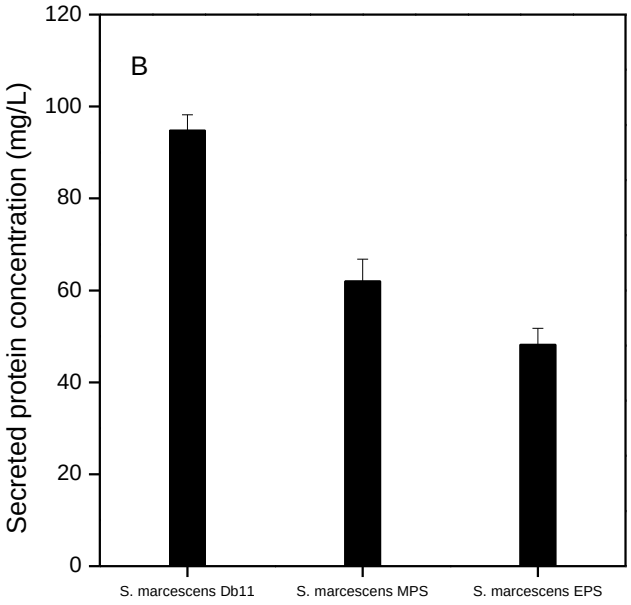
4.3.3 Chitinase activity and metabolic end-products changes

Compare to the chitinase activities of the wild-type strain, the *S. marcescens* EPS and the *S. marcescens* MPS chitinase activities were significantly increased compared to the wild-type strain (**Fig. 15A**). For instance, the endochtinase activity of *the S. marcescens* EPS (0.00197 U/mL) and the *S. marcescens* MPS (0.001518 U/mL) increased 2- and 1.5-fold, respectively. The chitobiosidase activity of the *S. marcescens* EPS (0.0634 U/mL) and the *S. marcescens* MPS (0.0304 U/mL) increased 3.9- and 1.9-fold, respectively. The β-*N*-acetylglucosaminidase activity of the *S. marcescens* EPS (0.132 U/mL) and the *S. marcescens* MPS (0.032 U/mL) increased 7.73- and 1.88-fold, respectively. Compared to the overall secreted proteins (94.81 mg/L) of the wild-type strain, the *S. marcescens* EPS (48.18 mg/L) and the *S. marcescens* MPS (62.0 mg/L) secreted overall protein was decreased 1.96- and 1.53-folds, respectively (**Fig. 15B**).

In addition, the main secreted metabolic end-products of the *S. marcescens* strains were measured when grown in M9 *N*-acetylglucosamine medium. The *S. marcescens* EPS completely consumed the initial 10 g/L *N*-acetylglucosamine after 24 h; the *S. marcescens* MPS consumed 8.51 g/L *N*-acetylglucosamine after 24 h; and the wild-type strain consumed 6.24 g/L *N*-acetylglucosamine after 24 h. Therefore, the *S. marcescens* EPS exhibited more actively at uptaking *N*-acetylglucosamine than the *S. marcescens* MPS and Db11. Compared to the secreted metabolic end-products of the

wild-type strain, *S. marcescens* EPS produced 1.93-fold increased 2,3-butanediol, 1.84-fold increased lactic acid, and 1.46-fold increased acetic acid (**Fig. 15C**). The overall increased metabolic end-products from the wild-type strain to the MPS and EPS strain is the consequence of improvement of *N*-acetylglucosamine uptake rates while no apparent differences of cell biomass formation among the three strains were observed. In addition, the molar end-production ratio of *S. marcescens* EPS (2,3-butanediol:acetic acid:lactic acid =5:3.3:1) is slightly higher than that of the wild-type strain (2,3-butanediol:acetic acid: lactic acid =4.5:3.8:1). The increased molar ratio of 2,3-butanediol:acetic acid:lactic acid from the EPS strain to the wild-type strain can be due to balance intracellular acidic conditions caused by production of acetic acid and lactic acid since *S. marcescens* prefers to produce 2,3-butanediol under an acidic condition (Ji et al. 2011).





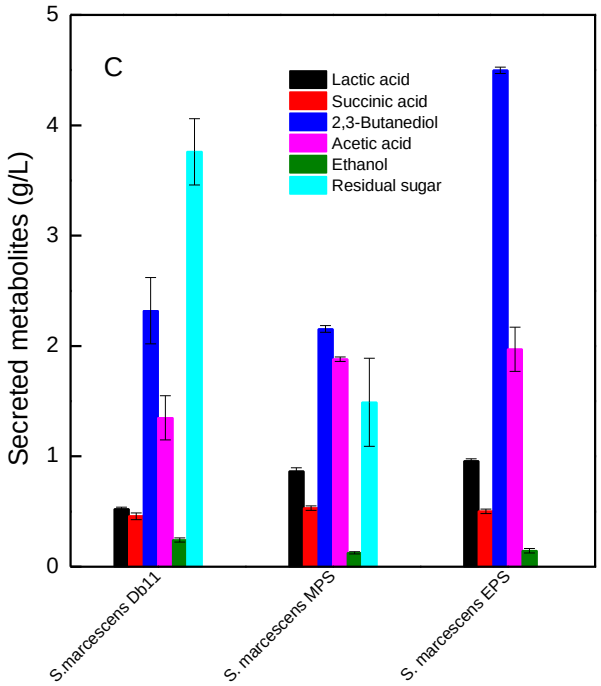


Figure 15 Characterization of phenotypical changes of the evolved strains and parental strain. (A) Chitinase activity, (B) overall secreted protein, (C) end-point metabolites.

4.3.4 Cell morphological changes

All *S. marcescens* strains were imaged by AFM and the detailed procedures can be found in the Materials and methods. Under a scope of 50 μm image, compared to *S. marcescens* Db11 cellular length and height, there is a generally increase trend of cellular length and height for *S. marcescens* MPS and EPS strains, shown in **Fig. 16A**, **16C**, and **16E**. Under a scope of 3.0 μm image where single cells were targeted and centered, the trend of increase at cellular length and height become clear ranging from *S. marcescens* Db11 to *S. marcescen* MPS and *S. marcescen* EPS (shown in **Fig. 16B**, **16D**, and **16E**). On average, *S. marcescens* EPS strains have a cellular length of 2.19 ± 0.19 μm and height of 331 ± 30 nm; *S. marcescens* MPS strains have a cellular length of 1.33 ± 0.08 μm and height of 197 ± 20 nm; *S. marcescens* Db11 strains have a cellular length of 1.31 ± 0.21 μm and height of 206 ± 34 nm. Overall, the end-point evolved strains exhibited larger cellular sizes (length and height) than the wild-type strains. There are two possible reasons that are responsible for the increased cellular sizes after evolutionary experiments: 1) it might be due to an increased secretion of more surrounded polysaccharides since *Serratia* sp. are known to form biofilms during cell growth; 2) it might be due to the increase of metabolic capabilities where facilitate nutrients and metabolites mass transformation by increasing outer or inner membrane areas. Since the evolution experiment was conducted by serial passage of *S. marcescens* under chitin as the sole carbon conditions, it may not be surprising that cells evolved to maximize the cellular membrane to uptake and secrete nutrients and metabolites to survive under the harsh conditions.

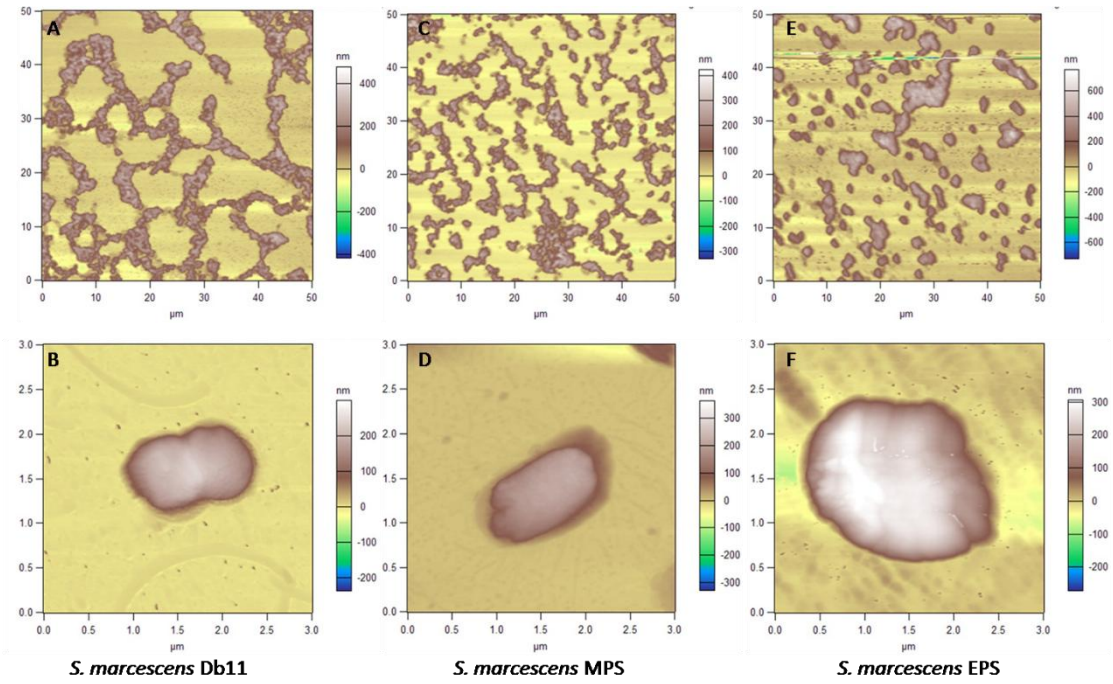
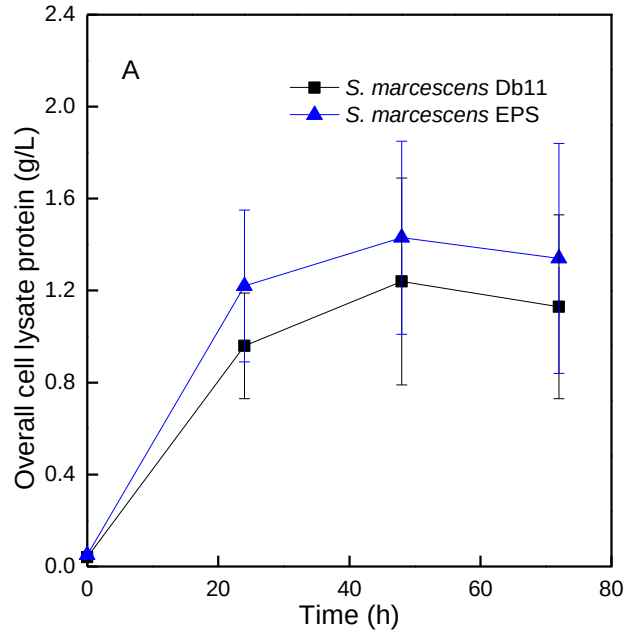


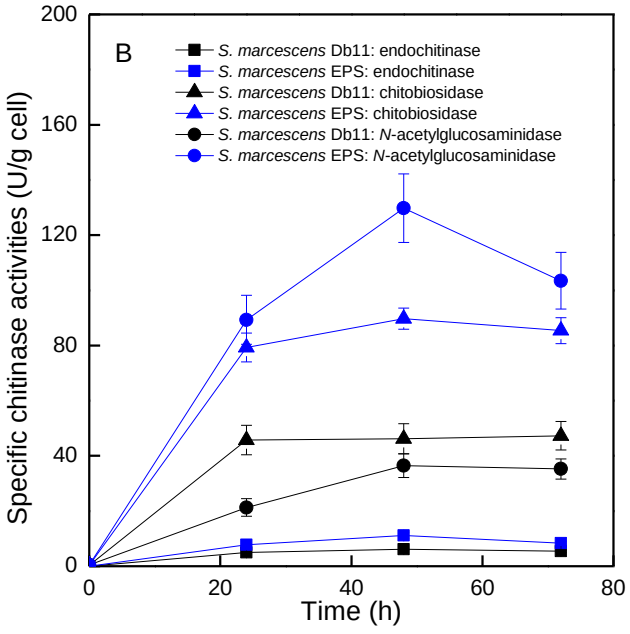
Figure 16 Cellular morphological characterization of *S. marcescens* strains by AFM. (A, C, and E) Image of *S. marcescens* Db11, MPS and EPS under scope of 50 μm; (B, D, and F) single cell image of *S. marcescens* Db11, MPS, and EPS under scope of 3 μm.

4.3.5 2,3-butanediol production from chitin

In order to evaluate the *S. marcescens* EPS 2,3-butanediol production capability from crystal chitin, the *S. marcescens* EPS and Db11 were cultured in 2% crystal chitin medium for 1-week. Compared to *S. marcescens* Db11 fermentation, the *S. marcescens* EPS shows higher cell yields (1.34 g/L vs 1.13 g/L) (**Figure 17A**), specific endochitinase activity (8.31 U/g cell vs 5.42 U/g), specific chitobiosidase activity (85.4 U/g cell vs 47.3 U/g cell), and specific β-*N*-acetylglucosaminidase (103.48 U/g cell vs 35.23 U/g) (**Figure 17B**). The growth and chitinase activity phenotypic differences are consistent to the above-mentioned physiological characterization results. After 72 h incubation, the *S. marcescens* EPS was able to produce 1.41 g/L 2,3-butanediol from 20 g/L crystal chitin, 3.7-fold improvement

compared to that of *S. marcescens* Db11 (0.38 g/L), shown in **Figure 17C**.





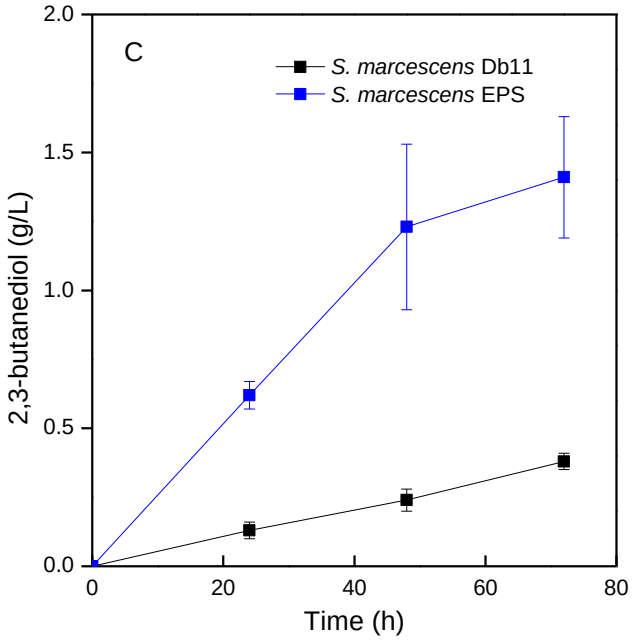


Figure 17 Cell growth (A), chitinase activities (B), production of 2,3-BD (C) using the *S. marcescens* Db11 (black color) and the *S. marcescens* EPS strain (blue color) under M9 medium supplemented with 1% yeast extract and 2% chitin at 30 ℃ and 100 rpm.

The *S. marcescens* EPS showed generally better growth on chitin than the wild-type strain and is able to produce 1.41 g/L 2,3-butanediol from 20 g/L crystal chitin due to more efficiently utilizing chitin as a substrate. Compared to a previous report of 2,3-butanediol titer (1.13 g/L) from a *S. marcescens chiR* overexpression strain (chiROE) (Yan et al. 2017), *S. marcescens* EPS is able to produce 2,3-butanediol titer 25% higher. In addition, the molar ration of 2,3-butanediol:acetic acid:lactic acid of *S. marcescens* EPS (4.8:2.7:1) is higher than the *S. marcescens* chiROE strain (3:2.8:1) because propagating cells under chitin medium is advantageous to provide selected pressure for cells to induce better phenotypes, namely growth and metabolic changes.

4.3.6 Whole genome sequencing of the evolved strains

Whole genome sequencing of the *S. marcescens* MPS and EPS genomes were conducted and single nucleotide variations of the evolved strains compared to the wild-type strain were listed in **Table 6**. Three SNVs were identified in *S. marcescens* MPS genome and eight SNVs were identified in *S. marcescens* EPS genome compared to the reference genome. The three SNVs of *S. marcescens* MPS are transport proteins. The SNVs of *S. marcescens* EPS contain transport proteins, global regulatory proteins, and pathway proteins. Two key mutations of the *S. marcescens* EPS were identified to be responsible for the physiological changes: 1) one mutation occurred at upstream of *SMDB11_2700* gene, encoding a LysR-family regulatory protein; 2) one mutation occurred in the coding sequence of SMDB11_1055, encoding a fimbrial adhensin.

SMDB11_1055 locates downstream of a chaperone-usher fimbriae operon (SMDB11_1052, encoding a putative fimbrial protein; SMDB11_1053, encoding a putative fimbrial chaperone; SMDB11_1054, putative fimbrial usher protein; and SMDB11_1055, a putative fimbrial adhensin). Through running protein sequence analysis against other organisms, SMDB11_1055 sequence is highly conserved among *Serratia* sp. and under the family of *Enterobacteriaceae*. Although the functionality of SMDB11_1055 has yet been characterized, it is generally believed that a chaperone-usher fibriae forms an outer membrane pore and secretes adhensin protein (i.e. SMDB11_1055) as a tip for reaching out and targeting substrates. The mutation of the putative fimbrial adhensin (SMDB11_1055) may affect *S. marcescens*

swarming motility phenotype and potentially increase the possibility of accessing to the substrate. In other studies, *S. marcescens* fimbrial proteins have been demonstrated essential for swarming motility toward its substrates. For instance, overexpression flagellum dependent genes (i.e. FlhD, encoding the flagellar expression master operon, and FliA, encoding an alternative sigma factor during flagellar assembly) can facilitate greatly *S. marcescens* swarming and swimming to kill against some fungal hyphae (Hover et al. 2016).

Table 6 Single nucleotide variations (SNVs) of the evolved strains compared to the wild-type strain after whole genome sequencing.

Position	Mutation	Gene	Annotations
S. marcescens MPS			
1121161	A→C	*SMDB11_1055*	Putative fimbrial adhensin
3667780	A→T	*SMDB11_3454*	unknown
1579431	C→T	*SMDB11_1516*	*rnfC*, electron transport complex protein
S. marcescens EPS			
1121161	A→C	*SMDB11_1055*	Putative fimbrial adhensin
1579431	C→T	*SMDB11_1516*	*rnfC*, electron transport complex protein
1579454	C→G	*SMDB11_1516*	*rnfC*, electron transport complex protein
2713402	A→G	*SMDB11_2558*	*fruA*, fused fructose-specific PTS enzymes: IIB component/IIC component
2861957	C→A	*SMDB11_2700*	Upstream of *SMDB11_2700*
2861957	C→A	*SMDB11_2700*	Upstream of *SMDB11_2702*
3667780	A→T	*SMDB11_3454*	Unknown
4366216	A→G	*SMDB11_4092*	*glnA*, glutamine synthetase

SMDB11_2700 encodes a LysR-family transcription regulator. Protein sequence similarity analysis of SMDB11_2700 showed that it is 67% indentify compared to a *Yersinia pseudotuberculosis* LysR-type transcription regulator RovM (Quade et al. 2011), , which controls gene transcription by attaching and changing DNA structures.

The LysR family regulatory protein functions in variety of categories such as a local regulator (i.e. chitinase secretion) and as a global regulator (i.e. quorum sensing) (Maddocks and Oyston 2008). In terms of the phenotypic changes and genome sequencing results, it is likely that SMDB11_2700 serves as a global regulatory protein and the mutation occurred in the upstream region of the global regulatory protein (SMDB11_2700) may affect a cascade effect of regulation in *S. marcescens*. In addition, *Serratia* sp. quorum sensing systems have been reported directly/indirectly controlling various phenotypes including sliding/swarming motility, production of chitinase, and butanediol fermentation (Van Houdt et al. 2007).

The mutation of SMDB11_2700 may also be responsible for 2,3-butanediol overproduction. In general, the overproduction of 2,3-butanediol can be a result of neutralize the acidic environment caused by the acidic end-products (e.g. acetic acid, lactic acid and succinic acid). In both *Serratia marcescens* MG1 and *Serratia plymuthica* RVH1, inactivation of the acylated homoserine lactones (AHL) synthase encoding gene leads to a reduced production of 2,3-butanediol and to a continued production of acidic end-products at the end of the exponential and throughout the stationary growth phase, which in turn leads to early growth arrest in the presence of fermentable sugars (Johansen et al. 1975; Magee and Kosarie 1987). Thus, SMDB11_2700 may play a key role in quorum sensing systems in *Serratia marcescens* Db11 that the mutation at the promoter region of SMDB11_2700 may account for increased chitinase activities and overproduction of 2,3-butanediol.

4.4 Conclusion

The newly evolved strain *S. marcescens* EPS exhibits increased chitinase activities, decreased secreted proteins, increased cell yield on chitin, and metabolic end-products changes compared to the wild-type strain. The EPS strain was capable of producing 1.41 g/L 2,3-butanediol from 2% crystal chitin, a 3.7-fold improvement compared to that of the wild-type strain. In addition, the molar ratio of 2,3-butanediol to byproducts is improved from the wild-type strain (2.5:2.5:1) to the EPS strain (4.8:2.7:1). Key mutations occurred in two gene products (SMDB11_2700 and SMDB11_1516) that encode a LysR-family regulator and a fimbrial adhensin may be accounting for the increased chitinase activities and overproduction of 2,3-butanediol.

Chapter 5 Metabolite profile of a chitinolytic bacterium *Serratia marcescens* using a transcriptomics-based genome-scale metabolic model

Importance of this part of work:

Since metabolic network of *S. marcescens* is complex and undertermined, genome-based constraint metabolic model is a powerful tool to understand metabolic profile of *S. marcescens* in a systematic manner. After construction of a computational model of *S. marcescens*, it can be readily deployed for predicting genetic targets to improve chemical production. Prior to my dissertation, no genome-scale metabolic model of *S. marcescens* is developed and intracellular metabolic profile remains uncharacterized.

The conclusions and fundamental results derived from this part of work:

We constructed a constraint-based genome-scale metabolic model (*i*SR929) including 929 genes, 1185 reactions and 1164 metabolites based on genomic annotation of *S. marcescens* Db11. The model performed reasonably comparing simulated cell growth and metabolites secretion with experimental acquisition, and 138 essential genes were predicted. The model *i*SR929 was curated by integrating mRNA-seq data of *S. marcescens* growth at three carbon sources (glucose, *N*-acetylglucosamine, and glycerol). Major metabolic differences of utilizing the three carbon sources rely on the utilization of citric acid cycle. Under growth of *N*-acetylglucosamine, *S. marcescens* exhibits high activity of pentose phosphate pathway and nucleotide synthesis but low activity of citric acid cycle. Three target chemicals (2,3-butanediol, *N*-acetylneuraminic acid, and *n*-butanol) can be implemented for metabolic

engineering applications by *S. marcescens* and their pathway gene activities were given.

5.1 Introduction

With recent advances in genomics, omics, and bioprospecting, the breadth of novel and interesting biochemistry continues to explore. In terms of the large amount of data and resources available, one current challenge is to conduct elaborate analysis to translate raw data into knowledge that provides functional insight. Genome-scale metabolic modeling (GSMM) facilitates metabolic analysis and dovetails well with genomic and biochemical information. While GSMM have advantages of being easily accessible and providing gene-protein-reaction level specificity, there are still a number of limitations (Yan and Fong 2017a). For instance, because metabolic networks are underdetermined, there exist alternative flux states with different pathway usage that produce indistinguishable cellular phenotypes. This is an underlying problem with GSMM that impacts multiple portfolios of these models including the initial reconstruction (association of specific reactions with annotated genes) and running simulation predictions (presence of alternate optimal solutions). In this study, we consider the metabolically under-characterized chitinolytic bacterium, *Serratia marcescens*, and utilize mRNA-seq data to run computational simulation to gain a better understanding of its metabolic network.

S. marcescens is unique among enteric bacteria in many aspects. It secretes extracellular DNase, gelatinase, lipase, several proteases, a red pigment (prodigiosin),

chitinases and a chitin binding protein. It is believed to be one of the most efficient chitin-degrading bacteria in the environment (Monreal and Reese 1969; Vaaje-Kolstad et al. 2013). *S. marcescens* Db11 contains ten chitinase-related proteins (Yan et al. 2017). Numerous studies recently have investigated on various facets of the *S. marcescens* chitinolytic mechanisms (Hamilton et al. 2014; Kim et al. 1998; Mekasha et al. 2017; Vaikuntapu et al. 2016). Due to the high efficiency of processing chitin materials, several individual *S. marcescens* chitinase genes have been cloned into those model bacterial species (e.g. *Escherichia coli*) (Brurberg et al. 1995; Suzuki et al. 2002; Tuveng et al. 2017). The cloned enzymes were isolated in good concentrations but failed to show similar level of chitinolytic activity as is found in *S. marcescens*. This may be due to the complexity of chitin degradation systems that are not defined by a few genes but is an interwined network of various enzymes (Hamilton et al. 2014; Suzuki et al. 2016; Suzuki et al. 2001; Yan et al. 2017). Hence, it is suggested that the best way to fully utilize the chitinolytic capabilities of *S. marcescens* may be developing *S. marcescens* rather than moving its chitinases into other systems by heterologous expression.

The ability to produce chemicals of industrial importance using inexpensive chitinolytic biomass has been a recent focus (Dahiya et al. 2006; Yan and Chen 2015; Yan and Fong 2015). Microbial conversion of chitin waste into value-added chemicals, known as consolidated bioprocessing (CBP), can help solve potentially social problems and make economic benifits. For *S. marcescens* Db11, the sequencing of its

genome in 2014 sets up a milestone towards understanding this industrially applicable microbe (Iguchi et al. 2014). It also showed success towards utilization of crystal chitin material (Yan et al. 2017). The initial developments in the characterization of *S. marcescens* are listed in **Table 7**. This is a promising development toward making use of the chitinolytic capabilities of this microbe to reduce the complex multi-step bioprocess to CBP.

Table 7 Significant milestone for *S. marcescens* research and characterization.

Year	Development	Group	Citation
1980	Genus Constructed	FlyG and Xanthopoulos	(Flyg et al. 1980)
1991-1999	Physical Characterization	Brurberg et al.; Suzuki et al.	(Brurberg et al. 1995; Watanabe et al. 1997)
1998	Shrimp/crab chitin degradation study to produce *N*-acetylglucosamine	Kim et al.	(Kim et al. 1998)
2014	Genome sequenced	Hayashi et al.	(Iguchi et al. 2014)
2017	Development of genetic modification method	Yan and Fong	(Yan et al. 2017)

With the availability of genomic sequences, it has become possible to use genome annotation and biochemical information to reconstruct cellular metabolic networks. Whole genome sequence and annotation of *S. marcescens* Db11 (Iguchi et al. 2014) was used to build GSMMs. The reaction database used for drafting the model includes but is not limited to sources such as KEGG (Kanehisa and Goto 2000), BiGG (Schellenberger et al. 2010), rBioNet (Thorleifsson and Thiele 2011), UniProt (Consortium 2008) and MBRole (Chagoyen and Pazos 2011).

The process of compiling biochemical reactions leads to the formulation of a draft model of *S. marcescens*, however, at this stage the represented metabolic network inevitably is incomplete and has numerous metabolic gaps. Thus, any initial model that is generated undergoes a gap-filling step. For our models, gap-filling was initially performed using a computational gap-filling algorithm (Brooks et al. 2012; Gowen and Fong 2010; Roberts et al. 2010; Vanee et al. 2014; Vanee et al. 2010). After gap-filling, simulations can be run by specifying input constraints such as substrate uptake rate, metabolites production rate, or oxygen uptake rate.

Once the model provides a framework to understand the cellular process, it can be used to find target for focused metabolic engineering to yield products of biotechnological value (Roberts et al. 2010). The most widely-used algorithms for design and simulation of genome-scale constraint-based metabolic models such as OptKnock (Burgard et al. 2003), OptForce (Ranganathan et al. 2010), EMiLio (Yang et al. 2011) are based on flux distribution and flow through chemically balanced reactions (FBA). FBA uses linear programming to optimize the objective function as follows:

$$\text{Maximize: } Z$$

$$\text{Subject to: } Sv = 0,$$

$$a_i \leq v_i \leq b_i \text{ for all reactions } i,$$

where, Z is the flux through objective function (biomass production and production

optimization), S: stoichiometry of the reactions represented as matrix, v is reaction flux vector, a_i and b_i are the constraints placed on the flux v_i of the reaction i (Palsson 2015).

Even after compiling biochemical information and gap-filling a model, there often are discrepancies between the computational model results and *in vivo* states due to difficulty in identifying using only computational approaches. A second step of model curation can be done by integrating high-throughput omics data with the framework of a computational model to realize "content in context" (Palsson 2015). The high-throughput omics data can be obtained by multi-scale high-throughput sequencing techniques such as transcriptomics, proteomics and metabolomics. This step reconciles computational predictions with experimental data and thereby helps enhance the characterization of the metabolic activity.

A variety of applications using incorporating omics data with the framework of GSMM have been reported for many prokaryotic microbes to more closely match cellular processes (Gowen and Fong 2010; Vanee et al. 2017; Vanee et al. 2014). Once the model closely resembles a biological system, it can be optimized for defined objective function. This objective function may range from production of biomass to production of a chemical target. Following the *in silico* optimization of yields the computational design may eventually be replicated for applications in industry, therapeutics or health-related predictions.

In this study, we aim at understanding metabolite profile of *S. marcescens* Db11 and its chitinolytic regulation systems. Therefore, a genome-scale metabolic model, *i*SR929, consisting of 1185 reactions, 1164 metabolites and 929 genes was constructed. The constructed model was analyzed for 1) accuracy compared to experimental results; 2) prediction of essential genes; 3) metabolic differences under different carbon growth. We also proposed three potential chemical compounds for metabolic engineering implementation.

5.2 Materials and Methods

5.2.1 Microbial growth

The *Serratia marcescens* Db11 was purchased from the Caenorhabditis Genetics Center (Twin City, USA http://www.cbs.umn.edu/CGC) (Flyg et al. 1980). The *S. marcescens* strain was grown in a M9 medium or LB medium containing various carbon sources. Carbon sources in the M9 medium were used at the following concentration: glucose, 5 g/L; *N*-acetylglucosamine, 5 g/L; glycerol 5%. *S. marcescens* pre-culture was performed in LB medium at 30℃ and 250 rpm for overnight. Then, 2.5% seed culture was inoculated at 50 mL M9 minimum medium with a 250 mL Erlenmeyer flask at 30℃, initial pH 7.5, and 220 rpm. Stock cultures of *S. marcescens* were maintained at -80℃ in a 26% (v/v) glycerol.

5.2.2 Construction of the model *i*SR929

The core of the stoichiometric metabolic model is a list of metabolic reactions occurring in *S. marcescens*, compiled based on evidence from genome annotations

and experimental observations. An initial list of biochemical reactions was assembled based on predicted enzymatic functions in the genomic annotations available from IMG, UniProt, and KEGG (Consortium 2008; Kanehisa and Goto 2000; Markowitz et al. 2008). Specifically, Enzyme Commission (EC) numbers of annotated *S. marcescens* genes were used to select reactions from a set of database reactions.

5.2.3 Linear programming for flux balance analysis

In-house python scripts were used to run the FBA simulations using the linear programming algorithm as shown in the introduction.

5.2.4 Objective function of FBA: biomass equation

The draft model aims at the growth optimization and the best estimates used in metabolic modeling scenario are biomass equation, which was designed by slightly manipulating the biomass equations from closely related species and available published information about *S. marcescens* growth conditions (Yan et al. 2017). The box below shows the biomass equation used for the simulation of all the version of the model.

5.2.5 Biomass Equation:

0.05 5mthf + 5.0E-5 accoa + 0.488 ala_L + 0.0010 amp + 0.281 arg_L + 0.229 asn_L + 0.229 asp_L + 45.7318 atp + 1.29E-4 clpn_SM + 6.0E-6 coa + 0.126 ctp + 0.087 cys_L + 0.0247 datp + 0.0254 dctp + 0.0254 dgtp + 0.0247 dttp + 1.0E-5 fad + 0.25 gln_L + 0.25 glu_L + 0.582 gly + 0.154 glycogen + 0.203 gtp + 45.5608 h2o + 0.09 his_L + 0.276 ile_L + 0.428 leu_L + 0.0084 lps_SM + 0.326 lys_L + 0.146 met_L + 0.00215 nad + 5.0E-5 nadh + 1.3E-4 nadp + 4.0E-4 nadph + 0.001935

pe_SM + 0.0276 peptido_SM + 4.64E-4 pg_SM + 0.176 phe_L + 0.21 pro_L + 5.2E-5 ps_SM + 0.035 ptrc + 0.205 ser_L + 0.0070 spmd + 3.0E-6 succoa + 0.241 thr_L + 0.054 trp_L + 0.131 tyr_L + 0.0030 udpg + 0.136 utp + 0.402 val_L --> 45.5608 adp + 45.56035 h + 45.5628 pi + 0.7302 ppi

5.2.6 Gap analysis

The draft model consists of the list of reactions however there are patches in the network that obstruct continuous flow of the flux through the pathway. These links are filled in by using the reaction databank and suggesting the list of reactions required to complete the network. FBA-GAP is used to suggest the connection nodes/reactions that are missing. The use and application of this framework have been described in past by Roberts et al., Vanee et al., and Gowen and Fong (Gowen and Fong 2010; Roberts et al. 2010; Vanee et al. 2014). FBA-GAP takes a draft model and biomass reaction and uses distances in the reaction network and mathematical optimization to produce a list of metabolites that are necessary for biomass production but cannot be produced or consumed by the cell. Reactions producing and consuming this list of metabolites are obtained from a reference database. These potentially gap-filling reactions were manually checked for relevant evidence such an associated protein/enzyme characterized or gene annotations. On detecting the specific evidence, these reaction additions to the model were accepted. The process is repeated until a positive biomass flux value is obtained.

5.2.7 Essentiality

Constraint-based models can be analyzed by FBA to make comprehensive *in silico*

gene essentiality predictions (Oh et al. 2007). For each gene in the model, that gene is assumed to be deleted or nonfunctional. Then, using the GPR relationships, the effect of the *in silico* gene knockout on reaction activities is assessed. If the gene is crucial to the activity of a reaction in the model, then that reaction is constrained to have zero flux to simulate the effect of the gene deletion. Finally, the model with new constraint is analyzed by FBA, maximizing the growth objective for batch growth on glucose. If the maximum flux on the biomass reaction is zero, then the deleted gene is predicted to be essential. If the maximum flux on the biomass reaction is greater than zero, the deleted gene is predicted to be nonessential. All constraints are reset to their default values, and the process is repeated for the next gene in the model.

5.2.8 mRNA sequencing

The total RNA of *S. marcescens* Db11 was isolated using mid-log phase cell cultures by QIAGEN (Venlo, Netherlands) RNA protect reagent in combination with the QIAGEN (Venlo, Netherlands) RNeasy Mini kit. The mRNA was isolated, enriched, and reverse transcribed into cDNA. The resulting cDNA was sequenced by a pair-ended reads using Illumina Hiseq 2500 (San Diego, CA). An average insert size of 500 bp was created and draft mRNA-Seq data were generated. CLC Genomics Workbench version 10 (QIAGEN, Venlo, Netherlands) and FASTQC were applied to trim reads for quality sequence data. The *S. marcescens* Db11 genome (GeneBank: HG326223) was used as a reference. Each mRNA sample was tested with two runs. The gene expression level was calculated by normalizing each gene's reads per kilobase per million mapped sequence reads (RPKM) to overall gene numbers.

5.2.9 Measurement of cell density

The measurement of *S. marcescens* strains culture density was generally quantified at OD600 using a Biomate3 UV/VIS spectrophotometer (Thermo Fisher Scientific, Waltham, MA). The growth of *S. marcescens* strains on crystal chitin and colloidal chitin was measured by testing cytoplasmic protein content (Yan et al. 2017). The concentration factor of the overall protein content to the dry cell weight (DCW) was DCW (g/L) = $1.20 \pm 0.03 \times$ protein (g/L).

5.2.10 Quantification of secreted metabolites

The concentration of *N*-acetylglucosamine, glucose, glycerol and all by-products (acetic acid, succinic acid, 2,3-butanediol, and ethanol) were analyzed by an Ultimate 3000 HPLC (Dionex, Sunnyvale, CA), equipped with a Bio-Rad Aminex HPX-87H column (Hercules, CA), a UV detector (199 nm) (Dionex, Sunnyvale, CA), and a refractive index detector (Shodex, Japan). The condition was run with 5 mM H2SO4 at 0.6 mL/min and 55 ℃.

5.2.11 Integration of gene expression data with genome-scale reconstruction

The integration of mRNA-seq data with the *i*SR929 model was analyzed using the method described by (Shlomi et al. 2008), in which the Boolean mapping from genes to reactions in *i*SR929 was used to translate gene expression data into predicted reaction activity states. Based on the relative gene expression levels calculated from RNAseq data, gene expression states were determined according to the following rules:

$$\text{gene state} = \begin{cases} -1 & g = 0 \\ 0 & 0 < g \leq \gamma \\ 1 & \gamma < g \end{cases}$$

Where g is the normalized expression level as determined by RNAseq (e.g. RPKM) and γ is the threshold above which a gene is called 'on'. The resulting gene states were then mapped using the gene-protein-reaction (GPR) relationships in *i*SR929 to generate lists of reactions predicted to be high or low, as determined by a flux with an absolute value higher than a given reaction threshold epsilon (ε). MILP was then used as in to find a network flux state that maximizes agreement with the expression based reaction activities, where a reaction was confirmed to be on when its flux was greater than or equal to a given reaction threshold ε (Shlomi et al. 2008). This method is distinct from traditional FBA because it is able to predict flux state for the network independent of a biomass objective function, which may be valuable in cases where cells may not have been evolved towards a maximum growth rate and, therefore, are not operating optimally.

5.2.12 Determination of intracellular citric acid and isocitric acid concentration

Intracellular levels of citric acid and isocitric acid were evaluated using a citrate assay kit (Sigma-Aldrich, St. Louis, MO) and an isocitric acid kit (BioVision, Milpitas, CA), respectively. In brief, *S. marcescens* cells growth were ceased at 2.5 h (OD600 about 0.4) and were collected by centrifuging at 4 ℃, 15,000 rpm, and 10 min. Citric acid and isocitric acid assays were conducted according to the manual's instruction. 10 μL supernatant samples were added in a total volume of 100μL reaction. Citric acid and

isocitric acid concentrations were determined at 570 nm and 450 nm by a VERSAmax microplate reader (Molecular Devices, Sunnyvale, CA), respectively.

5.3 Results and discussion

5.3.1 Construction of *S. marcescens* GSMM

Based upon genomic and available physiological evidence, a genome-scale constraint-based metabolic model for *S. marcescens*, hereafter denoted *i*SR929, was constructed. *i*SR929 contains 1185 reactions representing the function of 929 genes and 1164 metabolites (see **Table 8**). Among the 1100 gene associated reactions of the total 1185 reactions, there are 795 reactions that are associated with only one reaction and there are 305 reactions associated with more than one enzyme, meaning either isozymes or enzyme complexes. Of the 1164 metabolites, 1099 metabolites are intracellular, 43 metabolites are extracellular, and 22 metabolites are boundary. Furthermore, we added chitobiose degradation reactions to the model based upon the annotation and experimental evidence (EC 3.2.1.52, SMDB11_0477, SMDB11_1190 SMDB11_1542 and SMDB11_4602).

Table 8 Overview of genome-scale constraint-based model of *S. marcescens* iSR929.

	Serratia marcescens
Genome size	5.11 Mb
ORFs	4832
Included genes	929
Reactions associated with only 1 gene	795
Reactions associated with more than 1 gene	305
Reactions with gene associated	1185
Intracellular metabolites	1099
Extracellular metabolites	43
Boundary metabolites	22

The breakdown of *i*SR929 by functional categories is represented in **Figure 18**. The amino acid metabolism forms the largest of these, with 253 reactions. Experimentally, *S. marcescens* appears to have the capacity to synthesize all 20 amino acids (Yan et al. 2017), and *i*SR929 reflects this. Other large groups of reactions include subsystems related to carbohydrate metabolism (glycolysis, pentose phosphate pathway, pyruvate metabolism, and citric acid cycle), cofactor and vitamin metabolism (nicotinate and nicotinamide metabolism, folate biosynthesis, and prophyrin and chlorophyll metabolism) and nucleotide metabolism (including reactions related to purine and pyrimidine synthesis). The model also includes reactions relating to the synthesis of lipids, including fatty acid, glyceropospholipid, and glycerolipid metabolism.

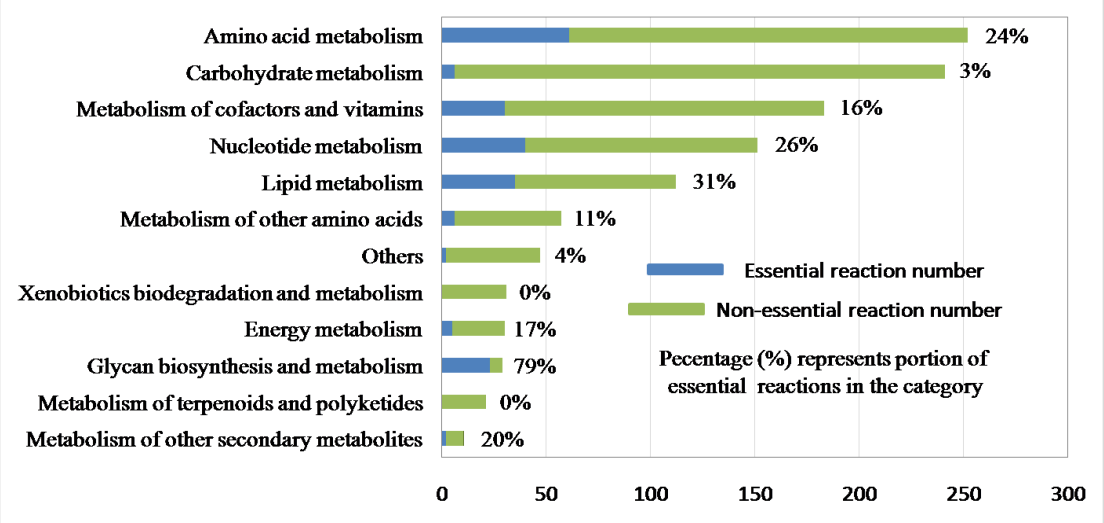


Figure 16 Reactions by functional category with number of reactions in model iSR929.

To test the predictions of the *i*SR929 model, we simulated growth of *S. marcescens* by applying FBA, assuming minimal media conditions with one of three possible carbon sources (glucose, *N*-acetylglucosamine, or glycerol). We then compared model

predictions to experimentally observed growth rates and fermentation product secretion profiles of *S. marcescens* grown in batch culture. In addition to the carbon source, the *in silico* minimal medium used for simulations contained water (H_2O), ammonia (NH4), sulfate (SO4), phosphate (pi), calcium (Ca2), ferrous iron (Fe3), hydrogen sulfide (H2S), potassium (K), magnesium (Mg2), pantothenate (pnto-r), and nicotinate D-ribonucleotide (nmn). In each of the three simulation conditions (glucose, *N*-acetylglucosamine, and glycerol), we applied progressively more experimentally determined constraints associated with by-product secretion rates to determine how closely the computational results could match the experimental results given the possibility of alternate optimal solutions. **Figure 19** shows the simulation results for each growth condition when exchange rates for the carbon sources, acetic acid, and succinic acid were constrained to match experimental observations.

5.3.2 Computational prediction of gene knockout targets

We conducted comprehensive *in silico* single gene deletions with *i*SR929, using glucose as a carbon source and the other minimal media components (as described above), and constraining glucose uptake to its experimentally observed value for batch growth (10.53 mmol/gDW/h). Gene essentiality results are shown in **Figure 18**. In the case of growth on glucose, we found that 138 (14.9%) of *S. marcescens* genes included in *i*SR929 were predicted to be essential. We also examined which subsystems of *i*SR929 contained the highest percentage of essential reactions (‘vulnerable subsystems’). Among the most vulnerable subsystems are the amino acid metabolism, nucleotide metabolism, lipid metabolism, and metabolism of cofactors

and vitamins.

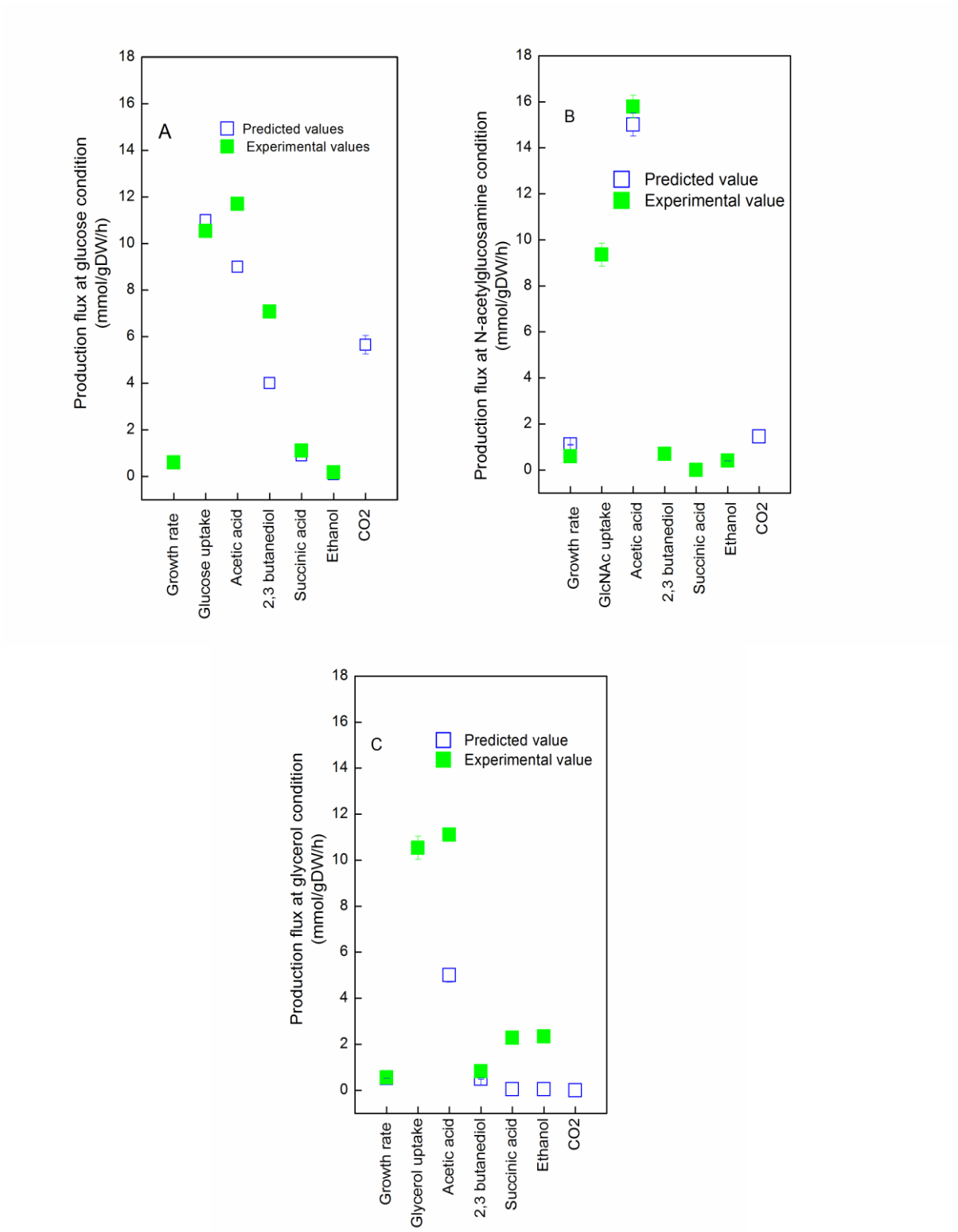


Figure 17 Comparison of model predictions to experimental values. *S. marcescens* *i*SR929 was used to simulate growth in multiple conditions. Actual and predicted flux rates are shown and predicted fermentation product production rates are shown as ranges as determined by flux variability analysis. For each simulation, the boundary fluxes for growth rate, GlcNAc uptake rate, growth rate and glycerol uptake rate were constrained to match the measured fluxes during (A) glucose, (B) N-acetylglucosamine, (C) glycerol conditions, respectively.

5.3.3 mRNA sequencing of *S. marcescens* under growth of different carbon sources

mRNA samples of *S. marcescens* Db11 were prepared from cell growth on six conditions: M9 medium with glucose, *N*-acetylglucosamine, or glycerol, plain LB medium, LB medium with crystal chitin or colloidal chitin. The reason for choosing these conditions are: 1) there is a general catabolite repression when *S. marcecens* growing in glucose or *N*-acetyglucosamine; 2) the three carbon sources can help identify *S. marcescens* active metabolic fluxes and metabolic differences. Overall, the average coverage per gene (4831 genes) of each sample's mRNA-seq implies a high-depth-coverage (generally 250).

5.3.4 Integration of mRNA-seq data into the GSMM

The integration of mRNA-seq data of *S. marcescens* at M9 medium with glucose, *N*-acetylglucosamine, and glycerol was deployed in the *i*SR929 based on the mixed integer linear programming algorithm approach (Shlomi et al. 2008). In order to setup a cutoff value of gamma, we initially categorized gene expression level into 58 cutoffs based on an exponential of 1.2, and a distribution of gene numbers against each gene expression level range was then plotted, a representative figure of M9 glucose condition was shown in **Figure S7**. The lower bound is 3.00, which precludes around 10% of the total number of genomic genes across all the conditions, and the upper bound is 850.56, which circled around 5% of the total number of genomic genes across all the conditions, shown in **Table S12**.

The model of *S. marcescens* was then run a simulation by integrating the transcriptomics data using a linear programming under the three conditions separately, where a reaction flux is 1000 if the corresponding gene's RPKM is over 850.56 and a reaction flux is 0 if the gene's RPKM is less than 3.00. A summary of the active metabolic reactions under three carbon sources were categorized in different metabolic pathway modules (**Table 9**). When growing at glucose condition, there are 150 active reactions. Among these 150 reactions, the majority is in carbohydrate metabolism (42), amino acid metabolism (38), lipid metabolism (12), cofactor metabolism (12), and nucleic acid metabolism (6). When growing at *N*-acetylglucosamine condition, there are 132 active reactions. Among these 132 active reactions, the majority is in carbohydrate metabolism (39), amino acid metabolism (28), lipid metabolism (6), cofactor metabolism (9), and nucleic acid metabolism (10). When growing at glycerol condition, there are 146 active reactions. Among these 146 active reactions, the majority is in carbohydrate metabolism (44), amino acid metabolism (41), lipid metabolism (6), cofactor metabolism (8), and nucleic acid metabolism (7).

Table 9 Numbers of active metabolic reactions in *i*SR929 by running simulation based on the transcriptomic data.

Carbohydrate metabolism	glucose	*N*-acetylglucosamine	glycerol
Butanoate metabolism	4	4	2
Citrate cycle (TCA cycle)	13	0	10
Glycolysis / Gluconeogenesis	7	10	9
Glyoxylate and dicarboxylate metabolism	3	2	1
Pentose and glucuronate interconversions	0	2	2
Pentose phosphate pathway	6	14	11

Propanoate metabolism	1	0	1
Pyruvate metabolism	6	5	6
Amino sugar and nucleotide sugar metabolism	2	2	2
Amino acid metabolism			
Alanine, aspartate and glutamate metabolism	4	3	5
Arginine and proline metabolism	9	10	23
beta-Alanine metabolism	0	0	2
Glutathione metabolism	2	2	0
Taurine and hypotaurine metabolism	2	2	2
Valine, leucine and isoleucine biosynthesis	5	2	2
Glycine, serine and threonine metabolism	16	9	7
Nucleic acid metabolism			
Purine metabolism	6	1	6
Pyrimidine metabolism	0	9	1
Lipid metabolism			
Fatty acid degradation	2	0	0
Glycerophospholipid metabolism	8	4	4
alpha-Linolenic acid metabolism	2	2	2
Energy metabolism			
Oxidative phosphorylation	3	3	4
Nitrogen metabolism	0	0	1
Metabolism of cofactors and vitamins			
Pantothenate and CoA biosynthesis	3	1	0
Riboflavin metabolism	2	2	2
One carbon pool by folate	6	6	6
Biosynthesis of other secondary metabolites			
Monobactam biosynthesis	2	2	2
Streptomycin biosynthesis	2	2	2
Xenobiotics metabolism			
Benzoate degradation	1	0	0
Nitrotoluene degradation	2	2	2
Transportation	31	31	29
Total	150	132	146

The main active pathways for the metabolism of glucose as a sole carbon source were shown in **Figure 20A**. The glucose trespasses the cell membrane and flows through glycolysis and pentose phosphate pathway to yield pyruvate. Then, the main flux goes to citrate cycle, butanonate metabolism (yielding 2,3-butanediol), amino acid synthesis, lipoylprotein synthesis, and pyruvate metabolism (yielding ethanol, formate, lactate, and acetate).

The main active pathway for the metabolism of *N*-acetylglucosamine as a sole carbon source was shown in **Figure 20B**. For the metabolism of *N*-acetylglucosamine, the carbon source was hypothesized first to be phosphorylated by a membrane type II phosphotransferase (SMDB11_0473). The *SMDB11_0473* gene expression level under *N*-acetylglucosamine condition (RPKM 967.20) was 12.09- and 5.13-fold than those under glucose (RPKM 80.23) and glycerol (RPKM 188.36) conditions, respectively. The carbon flows through glycolysis metabolism from fructose-6P to regenerate glucose-6P. The glucose-6P flows to the pentose phosphate pathway to yield pyruvate. Then, the majority carbon was distributed into purine and pyrimidine metabolism, butanonate metabolism (secreted 2,3-butanediol as an end-product), pyruvate metabolism (yielding ethanol, formate, lactate, and acetate as end-products), amino acid synthesis, and lipoylprotein synthesis.

The main active pathways for the metabolism of glycerol as a sole carbon source were shown in **Figure 20C**. Glycerol metabolism is thought to convert to pyruvate by three steps: a membrane facilitator protein (SMDB11_4021), an ATP-dependent glycerol kinase (SMDB11_4022), and a glycerol-3-phosphate dehydrogenase (SMDB11_3886) (Murarka et al. 2008). The SMDB11_4021 expression level of glycerol condition (RPKM 1772.81) is 89.5- and 70.8-fold higher than those of glucose and *N*-acetylglucosamine condition. The *SMDB11_4022* expression level of glycerol condition (RPKM 4535.25) is 76.6- and 63.6-fold higher than those of glucose

(RPKM 59.20) and *N*-acetylglucosamine (RPKM 71.28) condition. The gene expression level of *SMDB11_3886* of glycerol condition (RPKM 1683.89) is 59.9- and 54.7-fold higher than those of glucose (RPKM 28.10) and *N*-acetylglucosamine (RPKM 30.78) condition. The main active flux goes to glycolysis, pentose phosphate pathway, amino acid synthesis, and pyruvate metabolism.

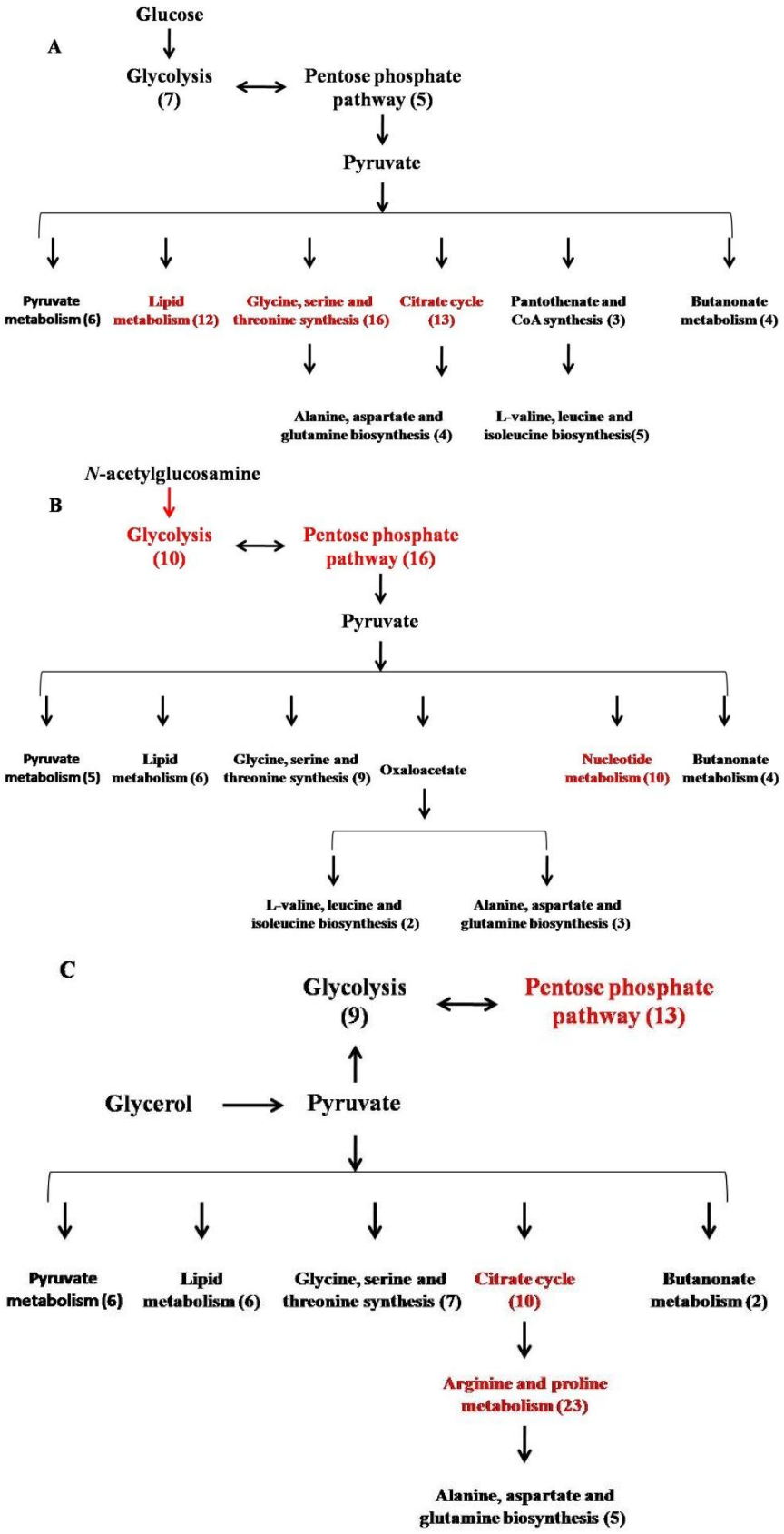


Figure 18 Summary of active pathways after running Ruppin algorithm using model *i*SR929 at each growth conditions. (A) M9 glucose; (B) M9 *N*-acetylglucosamine; (C) M9 glycerol. The most active metabolism pathways were marked in red.

5.3.5 *S. marcescens* exhibits lower activity in TCA cycle under *N*-acetylglucosamine as a sole carbon source

According to the model prediction, the main metabolic differences among the three carbon sources imply the utilization of TCA cycles. When glucose is the sole carbon source, the citrate cycle is fully active; under glycerol condition, the citrate cycle is partially active (R00361_smac, R01082_smac, R02164_smac, R00405_smac, R02570_smac, and R00621_smac); while under *N*-acetylglucosamine condition, the TCA cycle is not active. Intracellular citric acid and isocitric acid level can be indicators of activities of the TCA cycles and have been examined before. The fact that intracellular citric acid and isocitric acid concentration under growth of *N*-acetylglucosamine were lower than those under growth of glucose and glycerol supports the model prediction results (**Figure 21A** and **Figure 21B**). When *N*-acetylglucosamine is the sole carbon source, *S. marcescens* tends to utilize intermediates (e.g. oxaloacetate, fumarate, and 2-oxoglutarate) as precursors for amino acids synthesis. At the meantime, the majority carbon source flows for synthesis of nucleotides (e.g. purine and pyrimidine) from the ribose-5P generated from the pentose phosphate pathways since *N*-acetylglucosamine is an ideal carbon source for amino-sugar and nucleotides synthesis (e.g. C:N ratio 8:1).

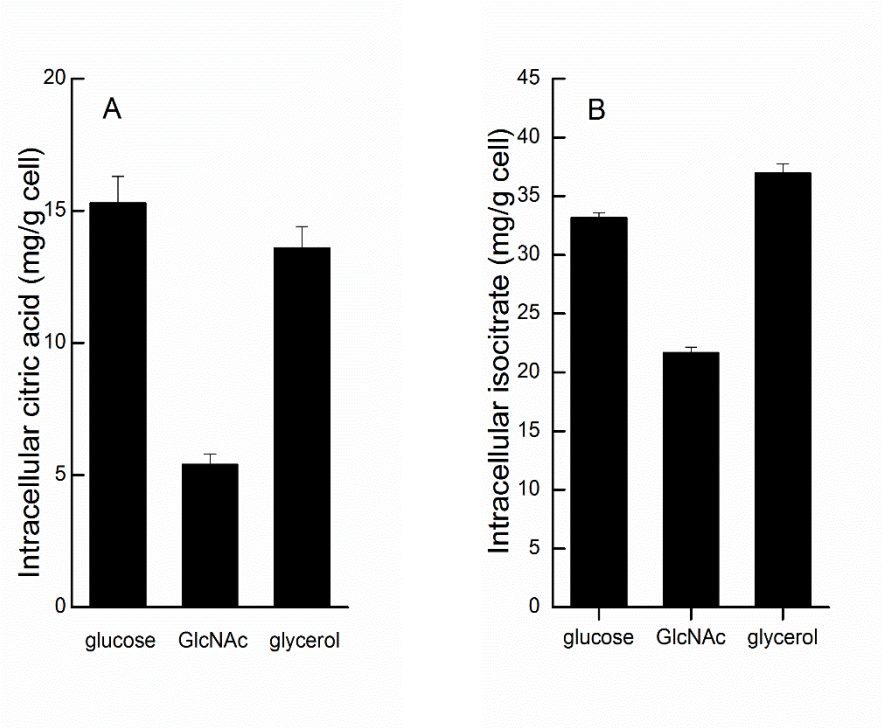


Figure 19 Intracellular citric acid (A) and isocitric acid (B) concentrations of *S. marcescens* Db11 grown at glucose, *N*-acetylglucsamine, and glycerol.

5.3.6 Insights of metabolic engineering application of *S. marcescens* from the model *i*SR929

Recently, developing the chitinolytic bacterium, *S. marcescens* as a workhorse to directly utilize chitin as carbon source to produce value-added chemicals is an interesting topic. The process includes directly degrading of chitin yielding fermentable sugar (*N*-acetylglucosamine), subsequently conversion of *N*-acetylglucosamine producing target chemicals. Based on the aforementioned experimental observation and computational simulation results, we target three chemical compounds (2,3-butanediol, *n*-butanol, and *N*-acetylneuraminic acid) and illustrate possible routes for deploying the pathway implementation.

S. marcescens Db11 is a native 2,3-butanediol producer. Direct production of

2,3-butanediol from crystal chitin by *S. marcescens* Db11 have been experimentally validated and the model *i*SR929 reflected it across all conditions (Yan et al. 2017). The proposed pathways of 2,3-butanediol and the corresponding annotated genes were shown in **Figure 22A**. Unlike other reported *S. marcescens* 2,3-butanediol producers (Bai et al. 2015; Zhang et al. 2016), *S. marcescens* Db11 harbors only two 2,3-butanediol dehydrogenases (meso and (2S,3S)) and relatively small amount of (2S,3S)-2,3-butanediol dehydrogenase was expressed compared to the meso-2,3-butanediol dehydrogenase, indicating meso-2,3-butanediol is the major product.

The second target of interest is a biofuel chemical, *n*-butanol. The *i*SR929 incorporates most of the reactions present in the butanoate metabolism. Besides the 2,3-butanediol synthetic pathway, two active reactions (R01171 and R01975) were found under glucose conditions and one active reactions (R01171) was observed under glycerol conditions. One consideration of the computational simulation is that butanoate is a product of a secondary pathway and *S. marcescens* may not overproduce this pathway associated genes. Thus, it is not necessarily surprising that growth simulations with no genetic designs incorporated may not show flux through secondary pathways. The proposed *n*-butanol synthetic pathway and their gene expression levels in *S. marcescens* Db11 were listed in **Figure 22B**. The annotated pathway toward *n*-butanol is partially completed but missing the last step enzyme (EC 1.2.1.10 alcohol dehydrogenase). A bifunctional alcohol dehydrogenase (AdhE2) from

Clostridium acetobutylicum is reported with highest activity of alcohol production (Deng and Fong 2011b) and can be an appropriate enzyme candidate for heterologous expression in *S. marcescens*. Furthermore, from the mRNA level of the *n*-butanol pathway genes, it is also unclear that the pathway is readily active when overexpression of *adhE2* gene in *S. marcescens*. For instance, the first step enzyme (SMDB11_4664 acetoacetate-CoA transferase) mRNA level is relative low and up-regulation of this gene mRNA level may also be needed.

The third chemical target of interest is *N*-acetylneuraminic acid. The *S. marcescens* Db11 doesn't harbor an endogenous pathway for *N*-acetylneuraminic acid synthesis. Compared to the glucose-based *N*-acetylneuraminic acid synthesis, *N*-acetylglucsaomine/chitin as a carbon sources are advantageous: 1) it requires only two steps for *N*-acetylneuraminic acid synthesis from *N*-acetylglucosamine compared to glucose; 2) no additional nitrogen source is needed (Steiger et al. 2011; Yan and Fong 2015). The growth of *S. marcescens* under *N*-acetylglucosamine condition showed that the first-step enzyme (SMDB11_0473) of *N*-acetylglucosamine uptake is highly active (RPKM 967.20). The SMDB11_0473 (*N*-acetylglucosamine transferase) is a reversible *N*-acetylglucosamine transferase that can generate intracellular *N*-acetylglucosamine (Kang et al. 2012), shown in **Figure 22C**. The *N*-acetylglucosamine can be further converted to *N*-acetylneuraminic acid by additional two steps (*N*-acetylmannosamine 2-epimerase and *N*-acetylneuraminic acid synthase) (Kang et al. 2012). Further overproduction of *N*-acetylneuraminic acid may

require carefully balance the heterogeous pathways since heterologous expression may generate metabolic burden.

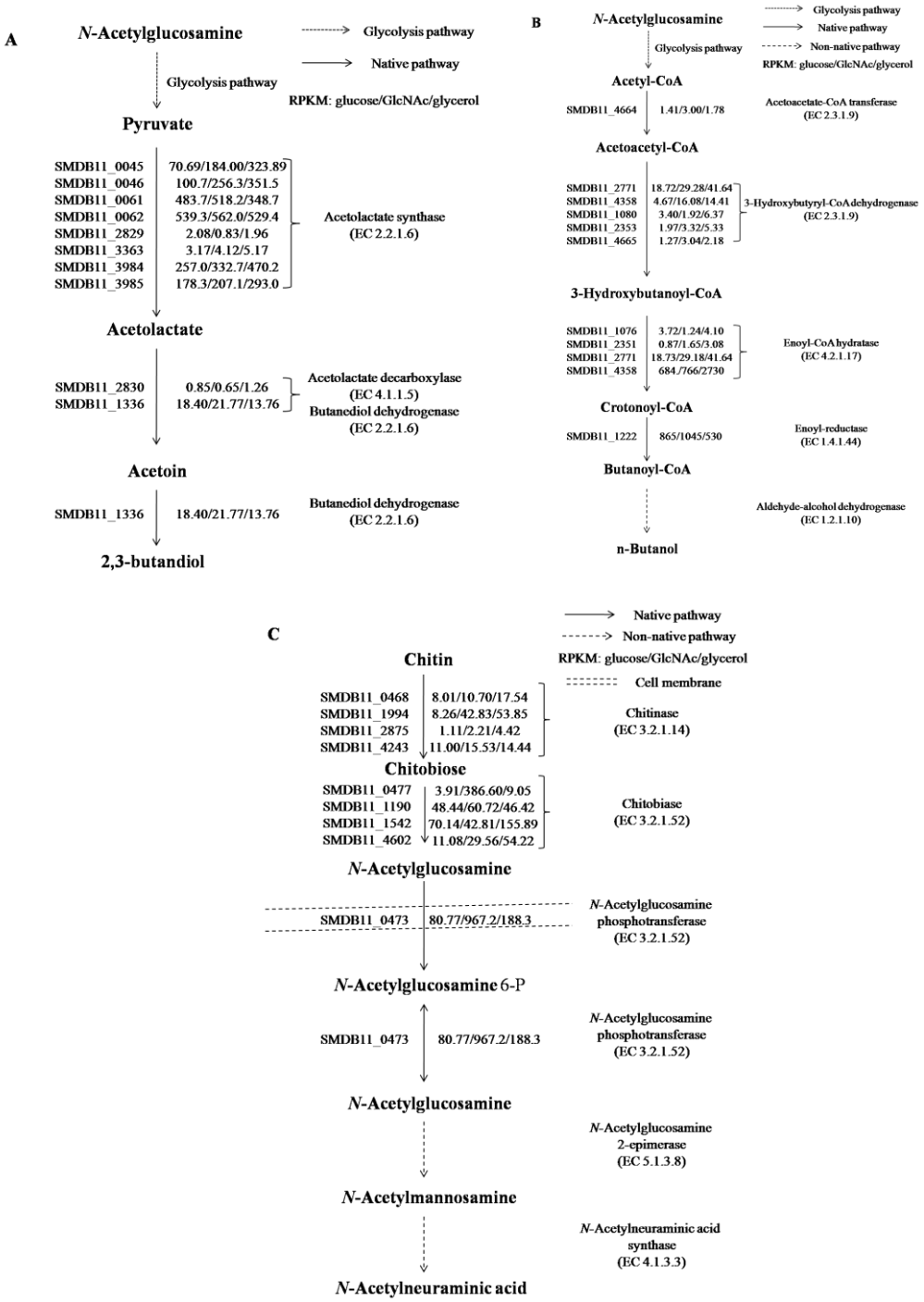


Figure 20 Proposed pathways for potential chemicals of interest production as metabolic engineering targets by *S. marcescens* and their corresponding gene expression values obtained from mRNA-Seq data

5.4 Conclusion

In this study, we constructed a *S. marcescens* genome-scale metabolic model *i*SR929. The model contains 929 genes, 1185 reactions and 1164 metabolites. The iSR929 model is able to provide reasonable predictions that match with experimental acquisition. Integrating transcriptomics data of different carbon sources (e.g. glucose, *N*-acetylglucosamine and glycerol) to the model reveals major metabolic difference is the utilization of TCA cycles. Three biochemicals (2,3-butanediol, *n*-butanol, and *N*-acetylneuraminic acid) were proposed as promising metabolic engineering implementation targets and pathway activities were given. Overall, our results show metabolite profile of *S. marcescens* and the model iSR929 can be applied to identifying genetic targets for chemical overproduction.

Chapter 6 Conclusion

6.1 Summary

Serratia marcescens is a chitinolytic bacterium. Its most important characteristics were able to produce kinds of enzymes especially chitinases. The chitinases of *S. marcescens* were identified to have broad manners of acting on crystalline chitin. The highly efficient chitinases can be used in various industries especially the chemical production industry, because chitinases can degrade chitin to sugars which are used by other microorganism to produce chemicals. Although there are some other bacteria which are able to produce chitinases, *S. marcescens* is much easier to be cultured and cheaper in the industry. Thus, the study of *S. marcescens* is of great importance. My dissertation work aims at characterizing a *S. marcescens* strain and developing it as a platform strain for chemical production from chitin.

Although *S. marcescens* chitinases have been characterized, the *S. marcescens* chitinolytic regulatory mechanisms remain unclear. ChiR is the only regulator protein that is believed to be essential for chitinase production, however, prior to this dissertation, its functionality of regulating chitinases is still not known. By generating a ChiR deletion strain and a ChiR overexpression strain allows me to characterize *S. marcescens* physiological changes of cell growth, chitinase production and activities. My results showed that ChiR is a positive regulator in controlling all chitinase transcription in *S. marcescens*. The *chiROE* strain produced 1.13 ± 0.08 g/L 2,3-butanediol from 2% crystal chitin, a 2.83-fold improvement from the wild-type

strain, indicating ChiR is an important and useful genetic engineering target for enhancing chitin utilization in *S. marcescens*.

Since I have demonstrated *S. marcescens* can grow on chitin and produce 2,3-butanediol from crystal chitin, I further exploit possibilities of producing novel chemicals (e.g. *N*-acetylneuraminic acid) using *S. marcescens*. Chitin may be an ideal substrate for producing *N*-acetylneuraminic acid since no additional nitrogen source is needed. Prior to this study, production of *N*-acetylneuraminic acid in *S. marcescens* has not been reported yet. In order to introduce an exogenous pathway for production of *N*-acetylneuraminic acid from *N*-acetylglucosamine, we systematically characterized *S. marcescens* promoters and heterologously overexpression an *N*-acetylglucosamine 2-epimerase and *N*-acetylneuraminic acid synthase. After transcriptional and translational balancing the pathways, the optimized recombinant strain P_{T5}-*slr1975*-P_{rplJ}-*neuB* was able to produce 0.48 g/L Neu5Ac from 20 g/L *N*-acetylglucosamine, and 0.30 g/L Neu5Ac from 5 g/L crystal chitin. These results represent the first demonstration of direct conversion of untreated chitin biomass to *N*-acetylneuraminic acid and illustrate the potential utility of *S. marcescens* as a chitinolytic bioprocess organism.

The capability of *S. marcescens* growing on chitin allows me to further improve *S. marcescens* growth phenotype on chitin as a substrate since the first step of chitin degradation is usually the rate-limiting step. Therefore, I conducted serial passage of

Serratia marcescens grown in colloidal chitin medium for 4 months, propagating 288 generation. The newly evolved strain *S. marcescens* EPS exhibits increased chitinase activities, decreased secreted proteins, increased cell yield on chitin, and metabolic end-products changes compared to the wild-type strain. The EPS strain was capable of increasing 2,3-butanediol production from 2% crystal chitin, a 3.7-fold improvement compared to that of the wild-type strain. In addition, the molar ratio of 2,3-butanediol to byproducts is improved from the wild-type strain (2.5:2.5:1) to the EPS strain (4.8:2.7:1). Key mutations occurred in two gene products (SMDB11_2700 and SMDB11_1516) that encode a LysR-family regulator and a fimbrial adhensin may be accounting for the increased chitinase activities and overproduction of 2,3-butanediol. To the best of my knowledge, this is the part of dissertation work is the frst adaptive evolution study using chitin as a substrate.

In previous parts of my dissertation, I have mainly focused on experimental implementation of *S. marcescens*. Computational tool is a powerful tool for mimicking the complexity of metabolic networks since the intraceullar fluxes are always undertermined. Before my dissertation study, there is no GSMM construction of *S. marcescens* and its metabolic profile remains uncharacterized. In this part of dissertation, a constraint-based genome-scale metabolic model (*i*SR929) including 929 genes, 1185 reactions and 1164 metabolites based on genomic annotation of *S. marcescens* Db11 was constructed. The model performed reasonably comparing simulated cell growth and metabolites secretion with experimental acquisition, and

138 essential genes were predicted. I have learned from the model prediction guided by mRNA-seq data that major metabolic usage differences are the utilization of citric acid cycle when GlcNAc is a substrate compared to glycerol or glucose as a substrate. On the contrary, *S. marcescens* exhibits high activity of pentose phosphate pathway and nucleotide synthesis but low activity of citric acid cycle when GlcNAc is the substrate. In addition, I have demonstrated that three target chemicals (2,3-butanediol, *N*-acetylneuraminic acid, and *n*-butanol) can be implemented for metabolic engineering applications by *S. marcescens* by giving their pathway gene activities.

6.2 Future work

Although a lot of work has been done in *S. marcescens* which has been described in this dissertation, there are still some work in *S. marcescens* that can be done in the future. First of all, the transcriptomics and metabolimics of the evolved *S. marcescens* EPS and MPS strains are needed to be tested and with these data, the systems level mechanisms could be discovered and so that the better way to increase chitinase production and 2,3-butanediol production could be discovered. Second, transforming the optimized constructs P_{T5}-*slr1975*-P_{rplJ}-*neuB* into the evolved strain may be helpful to increase Neu5Ac titer from chitin. Third, the constructed genome-scale metabolic model may be used to predict some genetic targets to simulate Neu5Ac production by implementing Neu5Ac production as an objective function. At the meantime, experimental implementation can be employed and the design-build-test cycle can be conducted to improve Neu5Ac titer.

6.3 Commercial feasibility

During my research, it was found that the chitinase production in *S. marcescens* is very powerful and the chitinase activity keeps highly active. Because *S. marcescens* is a Gram-negative bacterium and can use crystal chitin to grow, the cultivation of *S. marcescens* on the industrial level is relatively lower than those model organisms (e.g. *Escherichia coli* and *Sacchamyces cerevisiae*). Due to the special properties tested in the dissertation, *S. marcescens* could be used in the industry to produce chitinases in large scale. In order to produce a large-scale chitinases, we probably need to obtain some data of the cell physology and chitinase acitivity from bioreactors (i.e. 5 liter), then transform the information and fermentation parameters into even large scale (i.e. 100 liter). To the fact that S. marcescen is a native 2,3-butanediol producer, it may be particular of interest for the 2,3-butanediol chemical industry. Besides the experimental demonstrations of chemicals (e.g. Neu5Ac and 2,3-BD), *n*-butanol can be another target for experimental implementation. My work provided deep insight, valuable knowledge and useful tools that should be able to help a potential chitin degradation, chitinase production, or even chitin-based CBP industry.

Chapter 7 List of references

Aam BB, Heggset EB, Norberg AL, Sorlie M, Varum KM, Eijsink VG (2010) Production of chitooligosaccharides and their potential applications in medicine. Mar Drugs 8(5):1482-517 doi:10.3390/md8051482

Adrangi S, Faramarzi M (2013) From bacteria to human: a journal into the world of chitinases. Biotechnol Adv 31:1786-1795

Aghcheh RK, Nemeth Z, Atanasova L, Fekete E, Paholcsek M, Sandor E, Aquino B, Druzhinina IS, Karaffa L, Kubicek CP (2014) The VELVET a orthologue VEL1 of *Trichoderma reesei* regulates fungal development and is essential for cellulase gene expression. PLoS One 9(11):e112799 doi:10.1371/journal.pone.0112799.g001

Alper H, Fischer C, Nevoigt E, Stephanopoulos G (2005) Tuning genetic control through promoter engineering. Proc Natl Acad Sci USA 102(36):12678-12683

Ashry E, Aly M (2007) Synthesis and biological relevance of N-acetylglucosamine-containing oligosaccharides. Pure Appl Chem 12:2229-2242

Aye KN, Stevens WF (2004) Improved chitin production by pretreatment of shrimp shells. J ChemTechnol Biotechnol 79(4):421-425 doi:10.1002/jctb.990

Bai F, Dai L, Fan J, Truong N, Rao B, Zhang L, Shen Y (2015) Engineered Serratia marcescens for efficient (3R)-acetoin and (2R,3R)-2,3-butanediol production. J Ind Microbiol Biotechnol 42(5):779-86 doi:10.1007/s10295-015-1598-5

Benedik MJ, Strych U (1998) *Serratia marcescens* and its extracellular nuclease. FEMS Microbiol Lett 165:1-13

Blazeck J, Gang R, Reed B, Alper HS (2012) Controlling promoter strength and regualtion in *Saccharomyces cerevisiae* using synthetic hybrid promoters. Biotechnol Bioeng 109:2884-2895

Bondioli L, Ruozi B, Belleti D, Forni F, Vandelli M, Tosi G (2011) Sialic acid as a potential approach for the protection and targeting of nanocarriers. Expert Opin Drug Deliv 8:921-937

Bormann S, Baer ZC, Sreekumar S, Kuchenreuther JM, Dean Toste F, Blanch HW, Clark DS (2014) Engineering *Clostridium acetobutylicum* for production of kerosene

and diesel blendstock precursors. Metab Eng 25:124-130

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utlizing the principle of protein-dye binding. Anal Biochem 72(1):248-254

Brooks JP, Burns W, Fong SS, Gowen CM, Roberts SB (2012) Gap detection for genome-scale constriant-based models. Adv Bioinformatics 2012:323472

Brurberg MB, Eijsink VG, Haandrikman AJ, Venema G, Nes IF (1995) Chitinase B from *Serratia marcescens* BJL200 is exported to the periplasm without processing. Microbiology 141:123-131

Brurberg MB, Eijsink VG, Nes IF (1994) Characterization of a chitinase gene (*chiA*) from *Serratia marcescen*s BJL200 and one-step purification of the gene product. FEMS Microbiol Lett 124(3):399-404

Burgard A, Pharkya P, Maranas C (2003) Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol Bioeng 84:647-657

Burn J, Hamilton W, Wootton J, Johnston A (1989) Singe and multiple mutations affecting properties of the regulatory gene *nodD* of *Rhizobium*. Mol Microbiol 3(11):1567-1577

Chagoyen M, Pazos F (2011) MBRole: enrichment analysis of metabolomic data. Bioinformatics 27:730-731

Chen J, Shen C, Liu C (2010) N-acetylglucosamine: production and applications. Mar Drugs 8:2493-2516

Chen X, Varki A (2010) Advances in the biology and chemistry of sialic acids. ACS Chem Biol 5:163-176

Chien L, Lee C (2007) Hyaluronic acid production by recombinant Lactococcus lactis. Appl Microbiol Biotechnol 77:339-346

Consortium TU (2008) The Universal Protein Resource (UniProt). Nucleic Acids Res 36:D190-D195

Dahiya N, Tewari R, Hoondal G (2006) Biotechnological aspects of chitinolytic enzymes: a review. Appl Microbiol Biotechnol 71(6):773-82

Danishefsky I, Seteiner H, Bella A, Friedlander A (1969) Investigations on the chemistry of heparin: VI position of the sulfate ester groups. J Biol Chem 244:1741-1745

Deng Y, Fong SS (2010) Development and application of a PCR-targeted gene disruption method for studying CelR function in *Thermobifida fusca*. Appl Environ Microbiol 76(7):2098-106 doi:10.1128/AEM.02626-09

Deng Y, Fong SS (2011a) Laboratory evolution and multi-platform genome re-sequencing of the cellulolytic actinobacterium Thermobifida fusca. J Biol Chem 286(46):39958-66 doi:10.1074/jbc.M111.239616

Deng Y, Fong SS (2011b) Metabolic engineering of *Thermobifida fusca* for direct aerobic bioconversion of untreated lignocellulosic biomass to 1-propanol. Metab Eng 13(5):570-577

Flyg C, Kenne K, Boman H (1980) Insect pathogenic properties of *Serratia marcescens*: phage-resistant mutants with a decreased resistance to *Cecropia* immunity and a decreased virulence to *Drosophila*. J Gen Microbiol 120:173-181

Friedholff P, Gimadutdinow O, Pingoud A (1994) Identification of catalytically relevant amino acids of the extracellular *Serratia marcescens* endonuclease by alignment-guided mutagenesis. Nucleic Acid Res 22(16):3280-3287

Gerc AJ, Song L, Challis GL, Stanley-Wall NR, Coulthurst SJ (2012) The insect pathogen *Serratia marcescens* Db10 uses a hybrid non-ribosomal peptide synthetase-polyketide synthase to produce the antibiotic althiomycin. PLoS One 7(9):e44673 doi:10.1371/journal.pone.0044673

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison III CA, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6:343-5

Gibson KE, Silhavy T (1999) The LysR homolog LrhA promotes RpoS degradation by modulating activity of the response regulator SprE. J Bacteriol 181:563-571

Gowen CM, Fong SS (2010) Genome-scale metabolic model integrated with RNAseq data to identify metabolic states of Clostridium thermocellum. Biotechnol J 5:759-767

Häkkinen M, Valkonen MJ, Westerholm-Parvinen A, Aro N, Arvas M, Vitikainen M,

Penttilä M, Saloheimo M, Pakula TM (2014) Screening of candidate regulators for cellulase and hemicellulase production in *Trichoderma reesei* and identification of a factor essential for cellulase production. Biotechnol Biofuels 7(1):14 doi:10.1186/1754-6834-7-14

Hamilton JJ, Marlow VL, Owen RA, Costa Mde A, Guo M, Buchanan G, Chandra G, Trost M, Coulthurst SJ, Palmer T, Stanley-Wall NR, Sargent F (2014) A holin and an endopeptidase are essential for chitinolytic protein secretion in *Serratia marcescens*. J Cell Biol 207(5):615-626

Harti L, Zach S, Seidl-Seiboth V (2012) Fungal chitinase: diversity, mechanistic properties and biotechnological potential. Appl Microbiol Biotechnol 93:533-543

Hasunuma T, Okazaki F, Okai N, Hara K, Ishii J, Kondo A (2013) A review of enzymes and microbes for lignocellulosic biorefinery and the possibility of their application to consolidated bioprocessing technology. Bioresour Technol 135:513-522

Horn SJ, Sorbotten A, Synstad B, Sikorski P, Sorlie M, Varum KM, Eijsink VG (2006) Endo/exo mechanism and processivity of family 18 chitinases produced by *Serratia marcescens*. FEBS J 273(3):491-503 doi:10.1111/j.1742-4658.2005.05079.x

Hover T, Maya T, Ron S, Sandovsky H, Shadkchan Y, Kijner N, Mitiagin Y, Fichtman B, Harel A, Shanks RM, Bruna RE, Garcia-Vescovi E, Osherov N (2016) Mechanisms of Bacterial (Serratia marcescens) Attachment to, Migration along, and Killing of Fungal Hyphae. Appl Environ Microbiol 82(9):2585-94 doi:10.1128/AEM.04070-15

Hult E-L, Katouno F, Uchiyama T, Watanabe T, Sugiyama J (2005) Molecular directionality in crystalline *β*-chitin: hydrolysis by chitinases A and B from *Serratia marcescens* 2170. Biochem J 388:851-856

Husson E, Hadad C, Huet G, Laclef S, Lesur D, Lambertyn V, Jamali A, Gottis S, Sarazin C, Nguyen Van Nhien A (2017) The effect of room temperature ionic liquids on the selective biocatalytic hydrolysis of chitin via sequential or simultaneous strategies. Green Chem 19:4122

Igarashi K, Uchihashi T, Uchiyama T, Sugimoto H, Wada M, Suzuki K, Sakuda S, Ando T, Watanabe T, Samejima M (2014) Two-way traffic of glycoside hydrolase family 18 processive chitinases on crystalline chitin. Nat Commun 5:3975

doi:10.1038/ncomms4975

Iguchi A, Nagaya Y, Pradel E, Ooka T, Ogura Y, Katsura K, Kurokawa K, Oshima K, Hatton M, Parkhill J, Sebaihia M, Coulthurst SJ, Gotoh N, Thomson NR, Ewbank JJ, Hayashi T (2014) Genome evolution and plasticity of Serratia marcescens, an important multidrug-resistant nosocomial pathogen. Genome Biol Evol 6(8):2096-2110

Itzstein Mv, Wu W, Kok GB, Pegg MS, Dyason JC, Jin B, Phan Tv, Smythe ML, White HF, Oliver SW, Colman PM, Varghese JN, Ryan DM, Woods JM, Bethell RC, Hotham VJ, Cameron JM, Penn CR (1993) Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature 363:418-423

Jee J, Ikegami T, Hashimoto M, Kawabata T, Ikeguchi M, Watanabe T, Shirakawa M (2002) Solution structure of the fibronectin type III domain from Bacillus circulans WL-12 chitinase A1. J Biol Chem 277:1388-1397

Ji XJ, Huang H, Ouyang PK (2011) Microbial 2,3-butanediol production: a state-of-the-art review. Biotechnol Adv 29(3):351-64 doi:10.1016/j.biotechadv.2011.01.007

Jiang LQ, Fang Z, Guo F, Yang LB (2012) Production of 2,3-butanediol from acid hydrolysates of *Jatropha hulls* with *Klebsiella oxytoca*. Bioresour Technol 107:405-10 doi:10.1016/j.biortech.2011.12.083

Johansen I, Bryn K, Stormer P (1975) Physiological and biochemical role of the butanediol pathway in Aerobacter (Enterobacter) aerogenes. J Bacteriol 13:1124-1130

Jones JA, Toparlak OD, Koffas MA (2015) Metabolic pathway balancing and its role in the production of biofuels and chemicals. Curr Opin Biotechnol 33:52-9 doi:10.1016/j.copbio.2014.11.013

Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of gees and genomes. Nucleic Acids Res 28(1):27-30

Kang J, Gu P, Wang Y, Li Y, Yang F, Wang Q, Qi Q (2012) Engineering of an *N*-acetylneuraminic acid synthetic pathway in *Escherichia coli*. Metab Eng 14(6):623-9 doi:10.1016/j.ymben.2012.09.002

Kim K, Creagh AL, Haynes CA (1998) Effective production of

N-Acetyl-β-D-glucosamine by *Serratia marcescens* using chitinaceous waste. Biotechnol Bioprocess Eng 3:71-77

Kim TI, Ki KS, Lim DH, Vijayakumar M, Park SM, Choi SH, Kim KY, Im SK, Park BY (2017) Novel Acinetobacter parvus HANDI 309 microbial biomass for the production of N-acetyl-beta-d-glucosamine (GlcNAc) using swollen chitin substrate in submerged fermentation. Biotechnol Biofuels 10:59 doi:10.1186/s13068-017-0740-1

Kless H, Sitrit Y, Chet I, Oppenheim AB (1989) Cloning of the gene coding for chitobiase of *Serratia marcescens*. Mol Gen Genet 217:471-473

Lee YC, Chien HC, Hsu WH (2007) Production of *N*-acetyl-D-neuraminic acid by recombinant whole cells expressing *Anabaena* sp. CH1 *N*-acetyl-D-glucosamine 2-epimerase and *Escherichia coli N*-acetyl-D-neuraminic acid lyase. J Biotechnol 129(3):453-60 doi:10.1016/j.jbiotec.2007.01.027

Li S, Wang J, Li X, Yin S, Wang W, Yang K (2015) Genome-wide identification and evaluation of constitutive promoters in *streptomycetes*. Microb Cell Fact 14:172 doi:10.1186/s12934-015-0351-0

Liao JC, Mi L, Pontrelli S, Luo S (2016) Fuelling the future: microbial engineering for the production of sustainable biofuels. Nat Rev Microbiol 14:288-304

Lin BX, Zhang ZJ, Liu WF, Dong ZY, Tao Y (2013) Enhanced production of *N*-acetyl-D-neuraminic acid by multi-approach whole-cell biocatalyst. Appl Microbiol Biotechnol 97(11):4775-84 doi:10.1007/s00253-013-4754-8

Lin H, Travisano M, Kazlauskas RJ (2016) Experimental evolution of Trichoderma citrinoviride for fast deconstruction of cellulose. PLoS One 11(1):e0147024

Lundgren BR, Boddy CN (2007) Sialic acid and *N*-acyl sialic acid analog production by fermentation of metabolically and genetically engineered *Escherichia coli*. Org Biomol Chem 5:1903-1909

Luo Y, Zhang L, Barton KW, Zhao H (2015) Systematic identification of a panel of strong constitutive promoters from *Streptomyces albus*. ACS Synth Biol 4(9):1001-10 doi:10.1021/acssynbio.5b00016

Lynd LR, van Zyl WH, McBride JE, Laser M (2005) Consolidated bioprocessing of

cellulosic biomass: an update. Curr Opin Biotechnol 16(5):577-583

Ma SM, Garcia DE, Redding-Johanson AM, Friedland GD, Chan R, Batth TS, Haliburton JR, Chivian D, Keasling JD, Petzold CJ, Lee TS, Chhabra SR (2011) Optimization of a heterologous mevalonate pathway through the use of variant HMG-CoA reductases. Metab Eng 13(5):588-97 doi:10.1016/j.ymben.2011.07.001

Maddocks SE, Oyston PC (2008) Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology 154(12):3609-23 doi:10.1099/mic.0.2008/022772-0

Madhuprakash J, Singh A, Kumar S, Sinha M, Kaur P, Sharma S, Podile AR, Singh TP (2013) Structure of chitinase D from *Serratia proteamaculans* reveals the structural basis of its dual action of hydrolysis and transglycosylation. Int J Biochem Mol Biol 4(4):166-178

Magee R, Kosarie N (1987) The microbial production of 2,3-butanediol. Adv Appl Microbiol 32:89-161

Mahmoudian M, Noble D, Drake C, Middleton R, Montgomery D, Piercey J, Ramlakhan D, Todd M, Dawson M (1997) An efficient process for production of *N*-acetylneuraminic acid using *N*-acetylneuraminic aid aldolase. Enzyme and Microbial Technology 20:393-400

Markowitz V, Szeto E, Palaniappan K, Grechkin Y, Chu K, Chen I, Dubchak I, Anderson I, Lykidis A, Mavromatis K, Ivanova N, Kyrpides N (2008) The integrated microbial genomes (IMG) system in 2007: data content and analysis tool extensions. Nucleic Acids Res 36:D528-533

McEwen JT, Kanno M, Atsumi S (2016) 2,3 Butanediol production in an obligate photoautotrophic cyanobacterium in dark conditions via diverse sugar consumption. Metab Eng 36:28-36 doi:10.1016/j.ymben.2016.03.004

Mekasha S, Byman IR, Lynch C, Toupalova H, Andera L, Nas T, Vaaje-Kolstad G, Eijsink V (2017) Development of enzyme cocktail for complete saccharification of chitin using mono-component enzymes from *Serratia marcescens*. Process Biochemistry 56:132-138

Miller MD, Tanner J, Alpaugh M, Benedik MJ, Krause KL (1994) 2.1 Å structure of

Serratia endonuclease suggests a mechanism for binding to double-stranded DNA. Nature Structural Biology 1:461-468

Monreal J, Reese ET (1969) The chitinase of *Serratia marcescens*. Can J Microbiol 15:689-696

Mosier N, Wyman C, Dale B, Elander R, Lee Y, Holtlzaple M, Ladisch M (2005) Features of promosing technologies for pretreatment of lignocellulosic biomass. Bioresour Technol 96:673-686

Murarka A, Dharmadi Y, Yazdani SS, Gonzalez R (2008) Fermentation utilization of glycerol by *Escherichia coli* and its implications for the production of fuels and chemicals. Appl Environ Microbiol 74(4):1124-1135

Nowroozi FF, Baidoo EE, Ermakov S, Redding-Johanson AM, Batth TS, Petzold CJ, Keasling JD (2014) Metabolic pathway optimization using ribosome binding site variants and combinatorial gene assembly. Appl Microbiol Biotechnol 98(4):1567-81 doi:10.1007/s00253-013-5361-4

Nozzi NE, Atsumi S (2015) Genome engineering of the 2,3-butanediol biosynthetic pathway for tight regulation in Cyanobacteria. ACS Synth Biol 4(11):1197-204 doi:10.1021/acssynbio.5b00057

Oh Y, Palsson BO, Park S, Schilling C, Mahadevan R (2007) Genome-scale reconstruction of metabolic network in *Bacillus subtilis* based on high-throughput phenotyping and gene essentiality data. J Biol Chem 282(39):28791-28799

Olson DG, McBride JE, Shaw AJ, Lynd LR (2012) Recent progress in consolidated bioprocessing. Curr Opin Biotechnol 23(3):396-406

Palsson BO (2015) Systems Biology: constraint-based reconstruction and analysis. Cambridge University Press, New York

Patyshakuliyeva A, Arentshorst M, Allijn IE, Ram AF, de Vries RP, Gelber IB (2016) Improving cellulase production by Aspergillus niger using adaptive evolution. Biotechnol Lett 38:969-974

Percot A, Viton C, Domard A (2003) Optimization of chitin extraction from shrimp shells. Biomacromolecules 4:12-18

Piacente F, Bernardi C, Marin M, Blanc G, Abergel C, Tonetti M (2013)

Characterization of a UDP-N-acetylglucosamine biosynthetic pathway encoded by the giant DNA virus Mimicirus. Glycobiology

Pitera DJ, Paddon CJ, Newman JD, Keasling JD (2007) Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. Metab Eng 9(2):193-207 doi:10.1016/j.ymben.2006.11.002

Polley M, Phillips M, Wayner E, Singhal A, Hakomori S, Paulson J (1991) CD62 and endothelial cell-leukocyte adhesion molecule 1 (ELAM-1) recognize the same carbohydrate ligand, sialyl-Lewis x. Proc Natl Acad Sci USA 88(14):6382-6386

Portnoy T, Margeot A, Seidl-Seiboth V, Le Crom S, Ben Chaabane F, Linke R, Seiboth B, Kubicek CP (2011) Differential regulation of the cellulase transcription factors XYR1, ACE2, and ACE1 in *Trichoderma reesei* strains producing high and low levels of cellulase. Eukaryot Cell 10(2):262-71 doi:10.1128/EC.00208-10

Quade N, Dieckmann M, Haffke M, Heroven AK, Dersch P, Heinz DW (2011) Structure of the effector-binding domain of the LysR-type transcription factor RovM from Yersinia pseudotuberculosis. Biological Crystallography 67:81-90

Ranganathan S, Suthers P, Maranas C (2010) OptForce: an optimization procedure for identifying all genetic manipulations leading to targeted overproductions. PloS Comput Biol 6:e1000744

Rinaudo M (2006) Chitin and chitosan: properties and applications. Prog in Polym Sci 31:603-632

Roberts SB, Gowen CM, Brooks JP, Fong SS (2010) Genome-scale metabolic analysis of Clostridium thermocellum for bioethanol production. BMC Syst Biol 4:31-47

Roberts W, Selitrennikoff C (1988) Plant and bacterial chitinases differ in antifungal activity. J Gen Appl Microbiol 134:169-176

Samain E (2008) High yield production of sialic acid (Neu5Ac) by fermentation.

Schauer R (2000) Achievements and challenges of sialic acid research. Glycoconj J 17:485-499

Schellenberger J, Park J, Conrad T, Palsson B (2010) BiGG: a biochemical genetic and genomic knowledgebase of large scale metabolic reconstructions. BMC

Bioinformatics 11:213

Shen CR, Lan EI, Dekishima Y, Baez A, Cho KM, Liao JC (2011) Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. Appl Environ Microbiol 77:2905-2915

Shin HD, Yoon SH, Wu J, Rutter C, Kim SW, Chen RR (2012) High-yield production of meso-2,3-butanediol from cellodextrin by engineered *E. coli* biocatalysts. Bioresour Technol 118:367-73 doi:10.1016/j.biortech.2012.04.100

Shlomi T, Cabili MN, Herrgard MJ, Passon BO, Ruppin E (2008) Network-based prediction of human tissue-specific metabolism. Nat Biotech 26:1003-1010

Solovyev V, Salamov A (2011) Automatic annotation of microbial genomes and metagenomic sequences. In: Li R (ed) In metagenomics and its applications in agriculture, biomedicine and environmental studies. Nova science publishers, pp 61-78

Song M, Kim E, Ban Y, Yoo Y, Kim E, Park S, Pandey R, Sohng J, Yoon Y (2013) Achievements and impacts of glycosylation reactions involved in natural products biosynthesis in prokaryotes. Appl Microbiol Biotechnol

Steiger MG, Mach-Aigner AR, Gorsche R, Rosenberg EE, Mihovilovic MD, Mach RL (2011) Synthesis of an antiviral drug precursor from chitin using a saprophyte as a whole-cell catalyst. Microb Cell Fact 12:102-110

Sun J, Shao Z, Zhao H, Nair N, Wen F, Xu JH (2012a) Cloning and characterization of a panel of constitutive promoters for applications in pathway engineering in *Saccharomyces cerevisiae*. Biotechnol Bioeng 109(8):2082-92 doi:10.1002/bit.24481

Sun JA, Zhang LY, Rao B, Shen YL, Wei DZ (2012b) Enhanced acetoin production by *Serratia marcescens* H32 with expression of a water-forming NADH oxidase. Bioresour Technol 119:94-8 doi:10.1016/j.biortech.2012.05.108

Sundaram AK, Pitts L, Muhammad K, Wu J, Betenbaugh M, Woodard RW, Vann WF (2004) Characterization of *N*-acetylneuraminic acid synthase isozyme 1 from *Campylobacter jejuni*. Biochem J 383:83-89

Suzuki K, Shimizu M, Sasaki N, Ogawa C, Minami H, Sugimoto H, Watanabe T (2016) Regulation of the chitin degradation and utilization system by the ChiX small

RNA in *Serratia marcescens* 2170. Biosci Biotechnol Biochem 80(2):376-385

Suzuki K, Sugawara N, Suzuki M, Uchiyama T, Katouno F, Nikaidou N, Watanabe T (2002) Chitinases A, B, and C1 of *Serratia marcescens* 2170 produced by recombinant *Escherichia coli*: enzymatic properties and synergism on chitin degradation. Biosci Biotechnol Biochem 66(5):1075-1083

Suzuki K, Suzuki M, Taiyoji M, Nikaidou N, Watanabe T (1998) Chitin binding protein (CBP21) in the culture supernatant of *Serratia marcescens* 2170. Biosci Biotechnol Biochem 62(1):128-35 doi:10.1271/bbb.62.128

Suzuki K, Taiyoji M, Sugawara N, Nikaidou N, Henrissat B, Watanabe T (1999) The third chitinase gene (*chiC*) of *Serratia marcescens* 2170 and the relationship of its product to other bacterial chitinases. Biochem J 343:587-596

Suzuki K, Uchiyama T, Suzuki M, Nikaidou N, Regue M, Watanabe T (2001) LysR-type transcriptional regulator ChiR is essential for production of all chitinases and a chitin-binding protein, CBP21, in *Serratia marcescens* 2170. Biosci Biotechnol Biochem 65(2):338-47 doi:10.1271/bbb.65.338

Synstad B, Vaaje-Kolstad G, Cederkvist FH, Saua SF, Horn SJ, Eijsink VG, Sorlie M (2008) Expression and characterization of endochitinase C from *Serratia marcescens* BJL200 and its purification by a one-step general chitinase purification method. Biosci Biotechnol Biochem 72(3):715-23 doi:10.1271/bbb.70594

Tabata K, Koizumi S, Endo T, Ozaki A (2002) Production of *N*-acetyl-D-neuraminic acid by coupling bacteria expressing *N*-acetyl-D-glucosamine 2-epimerase and *N*-acetyl-D-neuraminic acid synthetase. Enzyme and Microbial Technology 30:327-333

Tao F, Zhang Y, Ma C, Xu P (2011) One-pot bio-synthesis: *N*-acetyl-D-neuraminic acid production by a powerful engineered whole-cell catalyst. Sci Rep 1:142 doi:10.1038/srep00142

Tews I, Perrakis A, Oppenheim A, Dauter Z, Wilson KS, Vorgias CE (1996) Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. Nature Structural Biology 3:636-648

Thorleifsson S, Thiele I (2011) rBioNet: A COBRA toolbox extension for

reconstructing high-quality biochemical networks. Bioinformatics 27:2009-2010

Toratani T, Kezuka Y, Nonaka T, Hiragi Y, Watanabe T (2006) Structure of full-length bacterial chitinase containing two fibronectin type III domains revealed by small angle X-ray scattering. Biochem Biophys Res Commun 348:814-818

Tronsmo A, Harman GE (1993) Detection and quantification of *N*-acetyl-*β*-D-glucosaminidase, chitobiosidase, and endochitinase in solutions and on gels. Anal Biochem 208:74-79

Tsai M, Takeishi T, Thompson H, Langley K, Zsebo K, Metcalfe D, Geissler E, Galli S (2001) Induction of mast cell proliferation, maturation, and protein synthesis by the rat c-kit ligand, stem cell factor. Proc Natl Acad Sci USA 88(14):6382-6386

Tuveng TR, Hagen LH, Mekasha S, Frank J, Arntzen MO, Vaaje-Kolstad G, Eijsink VGH (2017) Genomic, proteomic and biochemical analysis of the chitinolytic machinery of Serratia marcescens BJL200. Biochim Biophys Acta 1865(4):414-421 doi:10.1016/j.bbapap.2017.01.007

Uchihashi T, Katouno F, Nikaidou N, Nonaka T, Sugiyama J, Watanabe T (2001) Roles of the exposed aromatic residues in crystalline chitin hydrolysis by chitinase A from Serratia marcescens 2170. J Biol Chem 276:41343-41349

Vaaje-Kolstad G, Horn SJ, Sorlie M, Eijsink VG (2013) The chitinolytic machinery of *Serratia marcescens*--a model system for enzymatic degradation of recalcitrant polysaccharides. FEBS J 280(13):3028-49

Vaaje-Kolstad G, Houston DR, Riemen AH, Eijsink VG, van Aalten DM (2005) Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. J Biol Chem 280(12):11313-9 doi:10.1074/jbc.M407175200

Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sorlie M, Eijsink VG (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. Science 330(6001):219-222 doi:10.1126/science.1192231

Vaikuntapu PR, Rambabu S, Madhuprakash J, Podile AR (2016) A new chitinase-D from a plant growth promoting Serratia marcescens GPS5 for enzymatic conversion of chitin. Bioresouce Technology 220:200-207

Van Houdt R, Givskov M, Michiels CW (2007) Quorum sensing in *Serratia*. FEMS

Microbiol Rev 31:407-424

Vanee N, Brooks JP, Fong SS (2017) Metabolic Profile of the Cellulolytic Industrial Actinomycete Thermobifida fusca. Metabolites 7(4) doi:10.3390/metabo7040057

Vanee N, Brooks JP, Spicer V, Shamshuin D, Krokhin O, Wilkins JA, Deng Y, Fong SS (2014) Proteomics-based metabolic modeling and characterization of the cellulolytic bacterium Thermobifida fusca. BMC Syst Biol 8:86--99

Vanee N, Roberts SB, Fong SS, Manque P, Buck G (2010) A genome-scale metabolic model of *Cryptosporidium hominis*. Chem Biodivers 7:1026-1039

Varki A (1992) Diversity in the sialic acids. Glycobiology 2:25-40

Vimr E, Lichtenteiger C (2002) To sialylate, or not to sialylate: that is the question. Trends Microbiol 10:254-257

Wang B (2009) Sialic acid is an essential nutrient for brain development and cognition. Annu Rev Nutr 29:177-222

Wang C, Ehrhardt CJ, Yadavalli VK (2015) Single cell profiling of surface carbonhydrates on Bacillus cereus. Interface 12:20141109

Wang S, Liu G, Wang J, Yu J, Huang B, Xing M (2013) Enhancing cellulase production in *Trichoderma reesei* RUT C30 through combined manipulation of activating and repressing genes. J Ind Microbiol Biotechnol 40(6):633-41 doi:10.1007/s10295-013-1253-y

Watanabe T, Ariga Y, Sato U, Toratami T, Hashimoto M, Nikaidou N, Kezuka Y, Nonaka T, Sugiyama J (2002) Aromatic residues within the substrate-binding cleft of Bacillus circulans chitinase A1 are essential for hydrolysis of crystalline chitin. Biochem J 376:37-44

Watanabe T, Ishibashi A, Ariga Y, Hashimoto M, Nikaidou N, Sugiyama J, Matsumoto T, Nonaka T (2001) Trp122 and Trp134 on the surface of the catalytic domain are essential for crystalline-chitin hydrolysis by Bacillus circulans chitinase A1. FEBS Lett 494:74-78

Watanabe T, Ito Y, Yamada T, Hashimoto M, Sekine S, Tanaka H (1994) The roles of the C-terminal domain and type III domains of chitinase A1 from Bacillus circulans WL-12 in chitin degradation. J Bacteriol 176:4465-4472

Watanabe T, Kimura K, Sumiya T, Nikaidou N, Suzuki K, Suzuki M, Taiyoji M, Ferrer S, Regue M (1997) Genetic analysis of the chitinase system of *Serratia marcescens* 2170. J Bacteriol 179(22):7111-7117

Watanabe T, Suzuki K, Oyanagi W, Ohnishi K, Tanaka H (1990) Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed its evolutionary relationship to Serratia chitinase and to the type III homology units of fibronectin. J Biol Chem 265:15659-15665

Wu G, Yan Q, Jones JA, Tang YJ, Fong SS, Koffas MA (2016) Metabolic burden: cornerstones in synthetic biology and metabolic engineering applications. Trends Biotechnol 34(8):652-64 doi:10.1016/j.tibtech.2016.02.010

Xu P, Ranganathan S, Fowler ZL, Maranas CD, Koffas MA (2011) Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA. Metab Eng 13(5):578-87 doi:10.1016/j.ymben.2011.06.008

Yan N, Chen X (2015) Sustainability: Don't waste seafood waste. Nature 524:155-157

Yan Q, Fong SS (2015) Bacterial chitinase: nature and perspectives for sustainable bioproduction. Bioresources and Bioprocessing 2:31-39

Yan Q, Fong SS (2016) Biosensors for metabolic engineering. In: Singh S (ed) Systems Biology Application in Synthetic Biology. Springer, New Delhi, pp 53-70

Yan Q, Fong SS (2017a) Challenges and advances for genetic engineering of non-model bacteria and uses in consolidated bioprocessing. Front Microbiol 8(2060)

Yan Q, Fong SS (2017b) Study of *in vitro* transcriptional binding effects and noise using constitutive promoters combined with UP element sequences in *Escherichia coli*. J Biol Eng 101(20):7567-7578

Yan Q, Hong E, Fong SS (2017) Study of ChiR function in *Serratia marcescens* and its application for improving 2,3-butanediol from crystal chitin. Appl Microbiol Biotechnol 101(20):7567-7578 doi:org/10.1007/s00253-017-8488-x

Yang L, Cluett W, Mahadevan R (2011) EMILiO: a fast algorithm for genome-scale strain design. Metab Eng 13:272-281

Yang P, Wang J, Pang Q, Zhang F, Wang J, Wang Q, Qi Q (2017) Pathway

optimization and key enzyme evolution of *N*-acetylneuraminate biosynthesis using an *in vivo* aptazyme-based biosensor. Metab Eng 43:21-28

Yao G, Li Z, Gao L, Wu R, Kan Q, Liu G, Qu Y (2015) Redesigning the regulatory pathway to enhance cellulase production in *Penicillium oxalicum*. Biotechnol Biofuels 8:71 doi:10.1186/s13068-015-0253-8

Zhang L, Guo Z, Chen J, Xu Q, Lin H, Hu K, Guan X, Shen Y (2016) Mechanism of 2,3-butanediol stereoisomers formation in a newly isolated *Serratia* sp. T241. Sci Rep 6:19257 doi:10.1038/srep19257

Zhang L, Sun Ja, Hao Y, Zhu J, Chu J, Wei D, Shen Y (2010a) Microbial production of 2,3-butanediol by a surfactant (serrawettin)-deficient mutant of *Serratia marcescens* H30. J Ind Microbiol Biotechnol 37(8):857-862 doi:10.1007/s10295-010-0733-6

Zhang L, Yang Y, Sun J, Shen Y, Wei D, Zhu J, Chu J (2010b) Microbial production of 2,3-butanediol by a mutagenized strain of *Serratia marcescens* H30. Bioresour Technol 101(6):1961-7 doi:10.1016/j.biortech.2009.10.052

Appendix A Tables

Table S1 Primers used in this study

Primers	DNA sequences (5′-3′)
pUC19bb-f	CCTCGTCTCAACCAAAGCAATCAACCCATCAACCACCTGGGGCGTAATCATGGTCATAGC
pUC19bb-r	CGAGTAGTTCAGTAGCGGAAATGTCAGAGCCAGCGTCTTGACTGGCCGTCGTTTTACAA
UNS7-f	CAAGACGCTGGCTCTGA
UNS5-f	GAGCCAACTCCCTTTACAAC
ha1-f	CAAGACGCTGGCTCTGACATTTCCGCTACTGAACTACTCGGATTATTTGCTCAGGTTGAC
ha1-r	AATGACACTAAATTCCCAAGG
ha2-f	GGAATAGAGACGAAGGGAC
ha2-r	GAGCCAACTCCCTTTACAACCTCACTCAAGTCCGTTAGAGTTTCTTCCTGTCACGCTATT
kanoverlap-f	CGATTTTTGTGCGCCACCACCTTGGGAATTTAGTGTCATTTTTCTTATTTTATTTTAAAGAGGAATTTGTAACAAGGGGTGTTATGAGCC
Kanoverlap-r	GCCGAATGCTGCGGGGTTTTCATTGAGTTAACCGTTTGATTTTCGCGTCCCTTCGTCTCTATTCCATTAGAAAAACTCATCGAGCATCAA
kan-f	ATGAGCCATATTCAACGGG
kan-r	TTA GAA AAACTCATCGAGCATC

[a] underline in red is overlap.

Table S2 Real-time PCR primers used in this study

gene	Forward primers (5 ’-3 ’)	Reverse primer (5’-3’)
luxS	CCCTGAGCTGAACGAGTACC	ATCCAGAATGTGCTTGGCGA
chiA	CATTTGGCGCCGTTGAAAGA	GGTGAAATTGCGCCCGTAAA
chiB	GAAAGGGCTGTTCGTCACCT	AGATGCCAGAACATCACGCC
chiC	AGCAAGGCCAGTTCACCAAT	GGATGCCCTGGCCTTTCATA
chiR	GCGACACGCATAGCACAAAG	ATCATCGCCATCGGCCAAAT
cbp21	TTCGAACTGGATCAGCAAACG	GCGGTCAGCTTCCAGGTAAA
chb	TCGATAGCGTCAACGACTGG	GTCCGGGTTCACCTCGTAAG
SMDB11_1083	AACATGCGTCAACCAAACCG	GCCACCAGACCATTCCAGAC
SMDB11_1994	AACGTCAAAAGCCCACAGGA	CGTTGAATCGTACAGGTTGGC
SMDB11_4602	CAGGACGGCTATCAGGGCAT	CGGCAGAATTTCATTGCGGT
SMDB11_1190	GCTACGAGCTGGATGCAGAA	TCGTGGAAGTTGCGGGTAAA

Table S3 Potential binding sites of 10 chitinase-related genes in *S. marcescens* predicted using a conserved LTTR family motif

Gene ID	Upstream sequence
SMDB11_4243	accattcgaaaaacctgctttttaaagcaggttttttttcgcccgtagtttatgaggatgagtacctccgcaggaggttcgagcctcgcgaagcgagacaacgttgctttagcaacggcccgcagggcgcgcatcaaagatg cgcgtaatcctctctcgccgaccaccattcgaaaaacctgctttttaaagcaggttttttttcgcctgcagtttatgaggatgagtacctccttcgctacgaccggcttcgctttatggcgtgcctcaatctgccgtcctcgtgaaat atatttatcctttacgcttaattaattcacattccttattcccattaggttcttatttttcatgtttaaataaattcatgcttgccaaataaaacccgattgatagcgctcttgttttcacgttttttttacctatagtctaaatggacaacgcggg aactcttattcccgtcgcgtgggaaataccgtgattatttattaacgttaatcttcgtggattattgcggaatttttcgcttcggcaatgcgtagcgacgattaactcttttatgtttatcctctcggaataaaggaatcagttATG
SMDB11_2875	tagctcatctccttttttattttaattaaatgccaatcacggccggtgggacatcggatattcagaatctaattaaggatgctattacaaagcaaatgaataaatacagcccctatggtcgaattcattcatgctgagcgctggcac aactcatcagcgccaatttgtgcagcagtgaataaaaattcattgtcattgtgatttatttcgacttttgttttcactaaaaataaacattaatggcaacgggaaatatttccccatcataaaaacatccactctggagaaatgccAT G
SMDB11_0468	atgcggccatggcggccgcccggccgccattatttccttctccctcagcgtcagtttgaattaattcgttcatggcagcaaaaggtttcagcctgcctgcattaaaaatcctcattataactttacgccccgccaatagctgatat tgccggcgagcggaaaacccttacccctatcaatgaggctaccATG
SMDB11_2876	taagtcactcctgactgaaataatgttgtaagttgagcgtatttcaccgacaaaatgagagagtaataacgcaattggaaataacccttcttgtcggcttgttttatttttatcattctggtcatgcgatagagcggaatgttttaacg gctattaacgaaaattattatgatgacgatctttgatcgttaggctgaatttatttaactgtcgctgtgtaaaagtagggcgctatgatgaattttatatcaggtcgtgatgaatagtatttaagatagcgctctcatttatgaagtatatt gcttgttgctatgagtacccttgttcgatttgatcgatttttgtgcgccaccaccttgggaatttagtgtcatttttcttatttttatttttaaagaggaatttgtATG
SMDB11_2877	tacaaattcctctttaaaataaaataagaaaaatgacactaaattcccaaggtggtggcgcacaaaaatcgatcaaatcgaacaagggtactcatagcaacaagcaatatacttcataaatgagagcgctatcttaaatactatt catcacgacctgatataaaattcatcatagcgccctacttttacacagcgacagttaaataaattcagcctaacgatcaaagatcgtcatcataataatttttcgttaatagccgttaaaacattccgctctatcgcatgaccagaat gataaaaataaaacaagccgacaagaagggttatttccaattgcgttattactctctcattttgtcggtgaaatacgctcaacttacaacattatttcagtcaggagtgacttATG
SMDB11_0477	cacccagtatgacaaccacaccaacatcccgagctggggcggcggttacggcaacatcttccaggacgagaaagacgtcaagttcatggtcatcgcgccattcaccatcttctgatgcgttccccggcctcgccgccgg ggcagttaagggaattcaacgatgaaaaaaatcatgctgatgttggcggccgctgcggccctgagcgcctgcgcgcaacccaccgcgccgccggaagacgccaagttgaagcaggcttacagcgcctgcatcaacac tgccgaaggctcgccggagcgtctgcagccctgtaaggcggtgctgaacgtgctgaaacaagagaagcagcaccagcagttcgccgcgcaggaaacggtgcgggtgatggattatcagaactgcattatggcggtgc acagcggtaacggtcaggcttacgacgccaagtgcggcaagctgtggcaagaaattcgcgataacaataattaagaaggataagaagATG
SMDB11_1083	atccctcttaagcccaccgttttcgggcgcgctctgcggttgccggtctgcggcagagcgtgccttgcgtttattgccatctcagtgtgagctgctacgattttagctgacgtttttatgccacccgggtcaccggccttgcgc atattcttcatctcagccagaagaggggcagctATG
SMDB11_1994	aatccgctcgggcgccaccggcgccctctctgcgcggcggcgaaattatcgacgctgctcacaataccgcctttcttcttccccaacggcttgtgaccggcaggcgcttttgctgttatgactgcgacgcttttggcgtcaat cggttatacgcatccggggcccactgagcccggctttattgtgcccgcccctggcggtacggtcggccttgccgaccacaaggagactgacATG
SMDB11_4602	acccgttcgccccctccggccgggaggggggcaactgtgatcccgtccgctttttcaccgcaacggttggacatctgttctgtgctgagtctttatctacagcgttggctgtttttattgatttttttcctctgaggaatttccgAT

	G
SMDB11_1190	cggttggatctgcggcggcagttccgccgctactggcagcagctggatacgcggcgtttaactcctggctggctgcgcctgcaacatcacttttgcaaaacacgccaccctcaccgctgaagctggcgccgctgcacat ggacatccatccgggcaatttgctcgcgaccggcgaggggctgcggctgatcgactgggagtatgccgccgacggcgatgtggcgttggatatcgccgcgctgtttcgcggcaacggctgggcggcggaggcgcag cagcgctttttgcagcactacgcgcgccagggctatcacgacgtcgcccggctgcaggcgcaggtgcagcgctggctgccgtgggtggattatctgatgctgctgtggtttgaagtgcgctggcagcagagcggcgac gccgaatttttgcgctggggcgcggcgctgcgccggcgattctgtttatcatcatccgaattctgaacgaacaataatgaagtgag**ATG**

Green highlight is predicted promoter -35 box, yellow highlight was predicted promoter -10 box; Conserved RBS motif was highlighted in blue; each chitinase-related gene start codon (ATG) was marked in bold capital; potential binding sites were either highlighted in grey or marked by underline.

Table S4 Production of 2,3-BD and byproducts after 72 h fermentation from 2% chitin using *S. marcescens* Db11 and *chiROE*.

	S. marcescens Db11	*S. marcescens chiROE*
Biomass (g/L)	1.08 ±0.11	0.90 ±0.13
2,3-BD (g/L)	0.4 ±0.04	1.13 ±0.08
Acetic acid (g/L)	0.26 ±0.03	0.71 ±0.04
Lactic acid (g/L)	0.15 ±0.01	0.38 ±0.03
Ethanol (g/L)	0.03 ±0.00	0.08 ±0.03
Succinic acid (g/L)	0.17 ±0.02	0.40 ±0.03

Table S5 Ct & Data Analysis for *Serratia marcescens* Gene Expression

Gene		wild-type	chiROE	ΔchiR
luxS	Rep1	20.26	20.97	19.64
	Rep2	20.39	21.02	19.43
	Rep3	20.34	20.97	19.46
	Ave	20.33	20.99	19.51
	Stdv	0.07	0.03	0.11
	CV%	0.32%	0.14%	0.58%
chiA	Rep1	26.72	24	26.89
	Rep2	26.87	24.02	26.84
	Rep3	26.97	24.16	27.06
	Ave	26.85	24.06	26.93
	Stdv	0.13	0.09	0.12
	CV%	0.47%	0.36%	0.43%
chiB	Rep1	27.77	26.16	28.45
	Rep2	27.79	26.13	28.47
	Rep3	28.13	26.17	28.73
	Ave	27.90	26.15	28.55
	Stdv	0.20	0.02	0.16
	CV%	0.73%	0.08%	0.55%
chiC	Rep1	26.26	23.73	26.47
	Rep2	26.32	23.87	26.45
	Rep3	26.29	23.58	25.98
	Ave	26.29	23.73	26.30
	Stdv	0.03	0.15	0.28
	CV%	0.11%	0.61%	1.05%
chiR	Rep1	28.53	20.03	38.19
	Rep2	28.37	20.02	38.12
	Rep3	28.68	19.86	38.17
	Ave	28.53	19.97	38.16
	Stdv	0.16	0.10	0.04
	CV%	0.54%	0.48%	0.09%
cbp21	Rep1	26.73	24.43	28.18

	Rep2	27.07	24.22	28.2
	Rep3	26.92	24.42	28.25
	Ave	26.91	24.36	28.21
	Stdv	0.17	0.12	0.04
	CV%	0.63%	0.49%	0.13%
chb	Rep1	25.91	24.81	25.62
	Rep2	26.13	24.74	25.47
	Rep3	25.97	24.68	25.65
	Ave	26.00	24.74	25.58
	Stdv	0.11	0.07	0.10
	CV%	0.44%	0.26%	0.38%
SMDB11_1083	Rep1	23.94	23.6	23.39
	Rep2	23.88	23.51	23.55
	Rep3	24.05	23.47	23.02
	Ave	23.96	23.53	23.32
	Stdv	0.09	0.07	0.27
	CV%	0.36%	0.28%	1.17%
SMDB11_1994	Rep1	25.55	23.3	24.87
	Rep2	25.36	23.38	24.82
	Rep3	25.36	23.3	24.98
	Ave	25.42	23.33	24.89
	Stdv	0.11	0.05	0.08
	CV%	0.43%	0.20%	0.33%
SMDB11_4602	Rep1	22.9	22.15	22.35
	Rep2	23.12	21.93	22.38
	Rep3	23.02	22.05	22.68
	Ave	23.01	22.04	22.47
	Stdv	0.11	0.11	0.18
	CV%	0.48%	0.50%	0.81%
SMDB11_1190	Rep1	21.79	21.16	21.12
	Rep2	21.77	21.07	21.32
	Rep3	21.74	21.32	21.24
	Ave	21.77	21.18	21.23
	Stdv	0.03	0.13	0.10
	CV%	0.12%	0.60%	0.47%

Table S6 Data Analysis for *S. marcescens* chitinase gene of interest (GOI) expression level using *LuxS* as control

GOI	Sample	GOI, Ave Ct	GOI, STDV	luxS, Ave Ct	luxS, STDV	Δ Ct, GOI-luxS	Δ Δ Ct	Δ Δ Ct, STDV	Fold Change, AVG	Fold change, STDV
chiA	wild-type	26.85	0.13	20.33	0.07	6.52	0.00	0.14	1.00	0.14
	chiROE	24.06	0.09	20.99	0.03	3.07	-3.45	0.09	10.93	0.09
	ΔchiR	26.93	0.12	19.51	0.11	7.42	0.90	0.17	0.54	0.08
chiB	wild-type	27.90	0.20	20.33	0.07	7.57	0.00	0.21	1.00	0.21
	chiROE	26.15	0.02	20.99	0.03	5.17	-2.40	0.04	5.28	0.04
	ΔchiR	28.55	0.16	19.51	0.11	9.04	-0.93	0.20	0.87	0.10
chiC	wild-type	26.29	0.03	20.33	0.07	5.96	0.00	0.07	1.00	0.07
	chiROE	23.73	0.15	20.99	0.03	2.74	-3.22	0.15	9.32	0.15
	ΔchiR	26.30	0.28	19.51	0.11	6.79	0.83	0.30	0.56	0.16
chiR	wild-type	28.53	0.16	20.33	0.07	8.20	0.00	0.17	1.00	0.17
	chiROE	19.97	0.10	20.99	0.03	-1.02	-9.21	0.10	593.59	0.10
	ΔchiR	38.16	0.04	19.51	0.11	18.65	10.45	0.12	0.00	0.00
cbp21	wild-type	26.91	0.17	20.33	0.07	6.58	0.00	0.18	1.00	0.18
	chiROE	24.36	0.12	20.99	0.03	3.37	-3.21	0.12	9.23	0.12
	ΔchiR	28.21	0.04	19.51	0.11	8.70	2.12	0.12	0.23	0.03
chb	wild-type	26.00	0.11	20.33	0.07	5.67	0.00	0.13	1.00	0.13
	chiROE	24.74	0.07	20.99	0.03	3.76	-1.92	0.07	3.78	0.07
	ΔchiR	25.58	0.10	19.51	0.11	6.07	0.40	0.15	0.76	0.11
SMDB11_1083	wild-type	23.96	0.09	20.33	0.07	3.63	0	0.11	1.00	0.11
	chiROE	23.53	0.07	20.99	0.03	2.54	-1.09	0.07	2.12	0.07
	ΔchiR	23.32	0.27	19.51	0.11	3.81	0.18	0.29	0.88	0.25
SMDB11_1994	wild-type	25.42	0.11	20.33	0.07	5.09	0.00	0.13	1.00	0.13
	chiROE	23.33	0.05	20.99	0.03	2.34	-2.75	0.05	6.74	0.05
	ΔchiR	24.89	0.08	19.51	0.11	5.38	0.29	0.14	0.82	0.11
SMDB11_4602	wild-type	23.01	0.11	20.33	0.07	2.68	0.00	0.13	1.00	0.13
	chiROE	22.04	0.11	20.99	0.03	1.06	-1.63	0.11	3.09	0.11

	ΔchiR	22.47	0.18	19.51	0.11	2.96	0.28	0.21	0.83	0.17
SMDB11_1190	wild-type	21.77	0.03	20.33	0.07	1.44	0.00	0.07	1.00	0.07
	chiROE	21.18	0.13	20.99	0.03	0.20	-1.24	0.13	2.36	0.13
	ΔchiR	21.23	0.10	19.51	0.11	1.72	0.28	0.15	0.82	0.12

Table S7 Genetic engineering by regulatory protein as a strategy for increasing feedstock biomass utilization.

Regulatory protein	Microorganisms	Strategies	Significance	References
ClrB, Bgl2, and CreA	*Penicilliumoxalicum*	Overexpression *clrB* Deletion *bgl2* and *creA*	Cellulase activity increased from 10- to 50-fold	(Yao et al. 2015)
CelR	*Thermobifida fusca*	Deletion *celR*	Cellulase activity increased 16.69-fold	(Deng and Fong 2010)
Xyr1 and Ace1	*Trichoderma reesi*	Overexpression *xyr1* Downregulation *ace1*	Cellulose activity increased 2- to 4-fold	(Wang et al. 2013)
VEL1	*Trichoderma reesi*	Overexpression *VEL1*	Cellulase activity increased around 4-fold	(Aghcheh et al. 2014)
Ace3	*Trichoderma reesi*	Overexpression *Ace3*	Cellulase activity increased 3- to 5-fold	(Häkkinen et al. 2014)
XYR1, ACE2, and ACE1	*Trichoderma reesi*	Overexpression *XYR1*, *ACE1* and *ACE2*	Cellulase production rate increased around 4.5-fold	(Portnoy et al. 2011)
ChiR	*S. marcescens*	Overexpression *chiR*	Chitinase activity increased from 2.14 to 6.31-fold	This study

Table S8 Gene and promoter sequences used in this study.

Gene	Sequences
slr1975	5’-ATGATTGCCCATCGCCGTCAGGAGTTAGCCCAGCAATATTACCAGGCTTTACACCAGGACGTATTGCCCTTTTGGGAAAAATATTCCCTCGATCGCCAGGGGGGCGGTTACTTTAC CTGCTTAGACCGTAAAGGCCAGGTTTTTGACACAGATAAATTCATTTGGTTACAAAACCGTCAGGTATGGCAGTTTGCCGTTTTCTACAACCGTTTGGAACCAAAACCCCAATGGTT AGAAATTGCCCGCCATGGTGCTGATTTTTTAGCTCGCCACGGCCGAGATCAAGACGGTAATTGGTATTTTGCTTTGGATCAGGAAGGCAAACCCCTGCGTCAACCCTATAACGTTTT TTCCGATTGCTTCGCCGCCATGGCCTTTAGTCAATATGCCTTAGCCAGTGGGGCGCAGGAAGCTAAAGCCATTGCCCTGCAGGCCTACAATAACGTCCTACGCCGTCAGCACAATCC CAAAGGTCAATACGAGAAGTCCTATCCAGGTACTAGACCCCTCAAATCCCTGGCGGTGCCGATGATTTTAGCCAACCTCACCCTGGAGATGGAATGGTTATTACCGCCTACTACCGT GGAAGAGGTGTTGGCCCAAACCGTCAGAGAAGTGATGACGGATTTCCTCGACCCAGAAATAGGATTAATGCGGGAAGCGGTGACCCCCACAGGAGAATTTGTTGATAGTTTTGAA GGGCGGTTGCTCAACCCAGGACACGGCATTGAAGCCATGTGGTTCATGATGGACATTGCCCAACGCTCCGGCGATCGCCAGTTACAGGAGCAAGCCATTGCAGTGGTGTTGAACA CCCTGGAATATGCCTGGGATGAAGAATTTGGTGGCATATTTTATTTCCTTGATCGCCAGGGCCACCCTCCCCAACAACTGGAATGGGACCAAAAGCTCTGGTGGGTACATTTGGAAA CCCTGGTTGCCCTAGCCAAGGGCCACCAAGCCACTGGCCAAGAAAAATGTTGGCAATGGTTTGAGCGGGTCCATGATTACGCCTGGAGTCATTTCGCCGATCCTGAGTATGGGGAA TGGTTTGGCTACCTGAATCGCCGGGGAGAGGTGTTACTCAACCTAAAAGGGGGGAAATGGAAAGGGTGCTTCCACGTGCCCCGAGCTCTGTGGCTCTGTGCGGAAACTCTCCAAC TTCCGGTTAGTTAA-3’
AGE	5’-ATGGGTAAGAACCTCCAAGCCCTGGCGCAATTGTATAAGAATGCTTTGCTGAACGATGTATTGCCGTTCTGGGAAAACCACTCACTTGACTCTGAGGGAGGCTATTTCACCTGTT TGGATCGCCAAGGCAAGGTCTACGATACGGACAAATTCATCTGGCTGCAAAATCGGCAAGTCTGGACTTTTTCAATGCTCTGCAACCAACTTGAAAAACGTGAAAATTGGTTGAA AATAGCCCGGAATGGGGCTAAATTCTTGGCGCAACATGGTCGTGATGACGAAGGAAATTGGTATTTTGCCTTGACTCGGGGTGGCGAGCCGTTGGTCCAGCCATATAATATATTCAG CGATTGTTTCGCGGCCATGGCCTTCAGCCAATACGCTCTGGCCAGTGGGGAGGAATGGGCCAAGGATGTTGCTATGCAAGCCTATAATAACGTTCTGCGTCGGAAAGACAATCCTA AGGGGAAGTACACGAAAACCTACCCTGGTACGCGCCCGATGAAGGCCCTGGCAGTCCCCATGATCCTTGCAAATCTCACGCTTGAAATGGAGTGGCTTTTGCCGCAAGAGACGCT GGAAAATGTCCTTGCTGCAACGGTGCAAGAGGTTATGGGGGATTTCCTGGACCAGGAGCAGGGTCTTATGTATGAGAATGTAGCGCCAGACGGATCGCACATAGATTGTTTTGAAG GACGGCTGATAAATCCGGGACACGGAATCGAAGCCATGTGGTTTATCATGGACATCGCCCGGCGGAAGAATGATTCCAAGACTATAAATCAAGCGGTCGATGTTGTACTCAATATCT TGAACTTCGCATGGGATAATGAATATGGGGGTCTGTATTACTTCATGGATGCAGCCGGGCATCCTCCGCAACAGTTGGAGTGGGACCAGAAATTGTGGTGGGTCCACCTCGAATCCC TCGTCGCCCTGGCCATGGGATACCGTCTGACTGGTCGGGACGCCTGTTGGGCCTGGTACCAGAAAATGCATGATTACAGTTGGCAACATTTTGCAGACCCCGAATACGGTGAATGG TTCGGCTACTTGAATCGCCGGGGTGAGGTCCTGCTTAATTTGAAAGGAGGAAAGTGGAAAGGTTGTTTTCACGTTCCTCGGGCGATGTATTTGTGCTGGCAGCAATTTGAGGCTCT GAGCTAA-3’
nanA	5’-ATGGCAACGAATTTACGTGGCGTAATGGCTGCACTCCTGACTCCTTTTGACCAACAACAAGCACTGGATAAAGCGAGTCTGCGTCGCCTGGTTCAGTTCAATATTCAGCAGGGC

	ATCGACGGTTTATACGTGGGTGGTTCGACCGGCGAGGCCTTTGTACAAAGCCTTTCCGAGCGTGAACAGGTACTGGAAATCGTCGCCGAAGAGGCGAAAGGTAAGATTAAACTCA TCGCCCACGTCGGTTGCGTCAGCACCGCCGAAAGCCAACAACTTGCGGCATCGGCTAAACGTTATGGCTTCGATGCCGTCTCCGCCGTCACGCCGTTCTACTATCCTTTCAGCTTTG AAGAACACTGCGATCACTATCGGGCAATTATTGATTCGGCGGATGGTTTGCCGATGGTGGTGTACAACATTCCAGCCCTGAGTGGGGTAAAACTGACCCTGGATCAGATCAACACA CTTGTTACATTGCCTGGCGTAGGTGCGCTGAAACAGACCTCTGGCGATCTCTATCAGATGGAGCAGATCCGTCGTGAACATCCTGATCTTGTGCTCTATAACGGTTACGACGAAATC TTCGCCTCTGGTCTGCTGGCGGGCGCTGATGGTGGTATCGGCAGTACCTACAACATCATGGGCTGGCGCTATCAGGGGATCGTTAAGGCGCTGAAAGAAGGCGATATCCAGACCGC GCAGAAACTGCAAACTGAATGCAATAAAGTCATTGATTTACTGATCAAAACGGGCGTATTCCGCGGCCTGAAAACTGTCCTCCATTATATGGATGTCGTTTCTGTGCCGCTGTGCCG CAAACCGTTTGGACCGGTAGATGAAAAATATCTGCCAGAACTGAAGGCGCTGGCCCAGCAGTTGATGCAAGAGCGCGGGTGA-3’
neuB	5’-ATGCAAATAAAAATCGACAAGCTTACCATCTCGCAGAAGAACCCACTTATAATCCCTGAGATCGGAATAAACCATAATGGCAGCCTCGAGATCGCGAAACTTATGGTTGACGCGG CGAAACGGGCGGGGGCAAAGATTATTAAACACCAAACTCACATAGTCGAAGACGAAATGAGCCAGGAGGCTAAGAACGTGATCCCCGGCAATGCTAATATTAGTATCTATGAGATC ATGGAACAGTGCGCCCTGAACTATAAGGACGAGCTTGCTCTGAAAGAATACGTTGAAAAGCAAGGACTTGTATATCTGTCAACCCCGTTTTCTCGTGCGGCTGCGAACCGTCTGGA GGATATGGGGGTTTCGGCATATAAAATCGGTAGCGGCGAGTGCAATAATTACCCCCTTATCAAGCACATAGCCCAGTTTAAAAAACCAATGATCATAAGTACTGGGATGAATTCTATC GAAAGTATAAAACCAACCGTCAAAATTCTTCGGGACTATGAAATACCATTTGTCTTGCTGCACACTACCAACTTGTATCCGACCCCATCACATCTCGTGCGTTTGCAGGCAATGCTG GAACTTTACAAGGAGTTTAACTGTCTTTACGGACTCTCTGATCACACGACTAATAATCTTGCTTGTATTGGCGCCATCGCCTTGGGAGCTTCTGTGCTGGAGCGTCATTTCACTGACA CGATGGACCGTAAAGGCCCCGACATTGTCTGCAGTATGGATGAGTCAACCTTGAAGGACCTTATAAATCAAACCCAGGAGATGGTGCTCCTGCGCGGAGATAACAATAAAAACCCG CTCAAGGAAGAACAAGTTACTATCGACTTCGCCTTCGCCTCCGTAGTGTCGATCAAGGATATAAAAAAGGGGGAGATTCTGAGTATGGATAACATCTGGGTGAAGCGCCCGTCAAA AGGTGGTATCAGCGCGAAGGACTTTGAGGCCATCCTTGGGAAACGGGCCAAGAAGGACATAAAAAACAATATACAACTGACATGGGATGACTTTGAGTAA-3’
P_{T5}	5’-TGTGGAATTGTGAGCGGATAACAATTACGAGCTTCATGCACAGTGAAATCATGAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATATGTGGAATTGTGAGCGCTCA CAATTCCACAACG-3’
P_{budAB}	5’-TTTTTGCATTATATGCAAAAGCGTTTCTTATGAATACCCGCTTCTCTGTTGGCGACGCGGCGGCTACTCTGATTGCAGGTTGACGACGGGCTTCAGCGTCGCCGCATCCCGGCCGT ATAGAGAGAAAACGCATAAACCTCGCGGCGTTTAGCACCAAATAGCAGATATTTTTCCCAGGCTGCGTCAGCCGCCTGTGGCGTTGATTGTCGCTCCAACACCAGATTAACCGTTC AATTTGATATCCGACTCAACCTTAATGATTAGTGAGTCTTTACTTTCACTTATCTTCTTGTTGTTTCTGGATATCGGCCGGCCAGCCGTGCCGGTTTTTTTGTCGCCTCAGGCTGAATC GTTCCAACTGATGAACGCGCCTGAGGGGCATCTTGGCCCGCCGTCGGCGGGTCTTTTGAACGCGTGGAACCGCAGTTCCGCGCTGGTGGAGGAGTCAGCA-3’
P_{ampR}	5’-TTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAAT-3’
P_{ompA}	5’-GTAAATTTAGGATTTTTCCCATCATTTTCGCACCCCCTAAGCGGAACCTTTGCGTGGGAAGGCACAAAGCTAGTCCGGCTGCGGTTTGGCAAGCAAAAACGCTCGATGAATCGC CGTTTTTTTAGGCAAAGAATGTTAGCAAATATGTACGATTCCGCGTTTTTTTTTAGAGCCTTATCACATCACACTTGTAACTTTCGCGCCACGTTGTAGACTTTACATCGCCAAGGTTG

	CTCTATAACGTCAGAAAAACTGGCGAGTAACAAACGAGGGCTTAAACCTTGGCGAAGGAATTTAACCAAGGGCTTAAACAGCTTTAAAGCTCA-3’
P_{rplJ}	5’-TCGCACTTGCGATTATCGCTTTGACTTGGGGCCGAGATTTGTCTAGAATCTTATGCCCCAAGGTTTTGCTACTGGCTATGCCGTAGGCAAAACAAAGATTTTTCGGTTGGAGCCTG GCCTATCCAGGCCTCCGTCCAAGACCGCAGGTGTATCGCAAGATACTTAATCCTTCCTGCGTAGACGGTGACAGAGCCTAAGAAAATTTTTCTTTTTTATAAAAGAATGTCTTTGCTG GATTCTGCTCACCGTGTTAAGCCGCTCGGCGTGGCGTTTTACAAAGCGCCGAGTGAAGTGAGTTCCGGGGGATTTTCCCCGGCTAATCCAGGAGCAAAAAGCTA-3’
P_{rpsG}	5’-ATGGTTCTCCGTTAAGTAAGGCCAAACGTTTTAACTTTAATGTCAAAATAAACTCGTAGAGTTTTGGACAATCCTGAATTAACAACGGAGTATTTCC-3’
P_{rpsL}	5’-ATGAGAGTGATGTTGTATATTTCTTGACACCCATACTGGTCAGCCCTAAAATTCTGCGTCCTCGTACTCTGCCAATAGTGCATACGAGGGGATTTATCACATGTTTACGAAGCTAGT ACGAAGCAAAAACCAGGAGCTTTTTTA-3’
P_{gpmA}	5’-ATTTCCCCCACGCCACGGCGCGCCCGGAGCGATTTCGGCGCGCCGTTTCCCCCTCACAGCGCCAGCCTGAACGATTCAGCTTTTTCCTTTCGCCAGCCCTTACTTTGCGGCAAGA AATTGGCTATGGTTAGACACTATCGATTGCCGAGCCCCGGCCCGCGGCAATATAATGATTTTCGTTATCATTGAAATGTCAGATTATTAAGGAGTTAAGCT-3’
P_{metF}	5’-TCCTCTGCGGGCGCGGGCTCACCGCGCCCCTTTGCTTTTCCTGCCATACTTATCTTTGCCGCGATCATGAAATTCCCTCGTTGAGCGGGTGAGGATTAATCATCTTATTCCCCTCAT TTTCCTGTTGACTCTTTAGCCAACATGCGGCATTTTCTATCTAGACGTCTAAACGTATAGACGCTCATCAAGATGAAATAAAAACAGTGATCGATAAAGAGGTGAGCAGGGT-3’
P_{cysP}	5’-CGCCGTTGCCGAAGCCGGTGCGCCCGGCTTCGGCCCGTTTTTCCGTCCCGAAAACGCCTCGCATTCCCTGCCTTATAGCGTTTGATGCTTGTAAATCACCAAACGGTATATAAAA GCGTTACAGCTCCGCACCGAGTTATAAATACACTGGTCGTCAATAAGGCGAAAGCCTGAAGGAATGACTCAGTGATAAGGTGCAGA-3’
P_{phlA}	5’-TTTGCGGCGCAGTCAATCAGGGAGCTTCGGCTCCCTTTTCTGCGTTTGGTGGCCGAAAACCGAACGCGGATCACATTCTGTACAAAGATAAGCATTTCTAATACAGAACTCATCC GACCTGCCGATAGCTAAATCAGCACCTATTCAGGTGCTCAATAAAAAGTCTATCGACAAGGAGTCGGC-3’
P_{SMDB11_4059}	5’-TTTGATGCCGGCCGGCGGCACTTTTTTAACCGCTGGCGGCCCCTGTGGGGCAACTGCTATACTATGCGGGCCGCTTCTGCGGCCCGATAGCTTCCCTTCTGGTTGTTATCGCTGCT TTTTTTACTATGTACTACCCTTCCTCAAGAGCGCCGTCATCCCGTCGGCACGTATGGATAGATT-3’

Table S9 Primer sequences used for cloning and sequencing in this study.

Primers	Sequences
UNS7-F	5’-CAAGACGCTGGCTCTGA-3’
UNS7-R	5’-CGAGTAGTTCAGTAGCGGA-3’
UNS5-F	5’-GAGCCAACTCCCTTTACAAC-3’
UNS5-R	5’-CTCTAACGGACTTGAGTGAGG-3’
UNS3-F	5’-GCACTGAAGGTCCTCAATC-3’
UNS3-R	5’-CGACCTTGATGTTTCCAGTG-3’

Table S10 Primer sequences used for real-time PCR in this study.

Primers	Sequences
luxS-f	5’-CCCTGAGCTGAACGAGTACC-3’
luxS-r	5’-ATCCAGAATGTGCTTGGCGA-3’
slr1975-f	5’-GGCCGAGATCAAGACGGTAA-3’
slr1975-r	5’-GGGCAATGGCTTTAGCTTCC-3’
nanA-f	5’-CTGCGATCACTATCGGGCAA-3’
nanA-r	5’-CACCTACGCCAGGCAATGTA-3’
gfp-f	5’-GGGTGAAGGTGACGCAACTA-3’
gfp-r	5’-CGAGCAAAGCACTGAACACC-3’

Table S11 Fermentation parameters of the *S. marcescens* recombinant strains on M9 medium with 0.1% yeast extract and 20 g/L GlcNAc at 30 ℃ and 200 rpm.

S. marcescens strain	Biomass (g/L)	Consumed GlcNAc (g/L)	Neu5Ac (g/L)	2,3-BD (g/L)	Acetic acid (g/L)	Lactic acid (g/L)
P_{T5}-*slr1975*-P_{rplJ}-*nanA*	1.35 ±0.12	18.4 ±0.58	0.30 ±0.03	0.81 ±0.06	0.34 ±0.02	0.53 ±0.09
P_{T5}-*AGE*-P_{rplJ}-*neuB*	1.34 ±0.10	19.4 ±0.32	0.45 ±0.05	0.75 ±0.07	0.21 ±0.03	0.45 ±0.10
P_{T5}-*AGE*-P_{rplJ}-*nanA*	1.72 ±0.10	18.4 ±0.37	0.21 ±0.04	0.92 ±0.24	0.31 ±0.02	0.62 ±0.10
P_{T5}-*slr1975*-P_{rplJ}-*neuB*	1.5 ±0.08	16.9 ±0.48	0.47 ±0.06	0.77 ±0.11	0.22 ±0.01	0.47 ±0.02
P_{rplJ}-*slr1975*-P_{T5}-*nanA*	1.44 ±0.11	19.0 ±0.60	0.1 ±0.02	1.12 ±0.56	0.67 ±0.02	1.03 ±0.33
P_{T5}-*slr1975*-P_{T5}-*nanA*	0.58 ±0.09	5.7 ±0.11	0.04 ±0.00	0.18 ±0.01	0.09 ±0.00	0.06 ±0.00
P_{rplJ}-*slr1975*-P_{rplJ}-*nanA*	1.85 ±0.10	19.3 ±0.78	0.10 ±0.02	1.15 ±0.22	0.68 ±0.02	1.3 ±0.23
P_{cysP}-*slr1975*-P_{cysP}-*nanA*	2.42 ±0.11	20.2 ±0.86	0.04 ±0.00	1.21 ±0.28	0.73 ±0.03	1.12 ±0.34
wild-type	2.41 ±0.12	20.0 ±0.99	0.00 ±0.00	3.38 ±0.52	0.84 ±0.03	0.84 ±0.03

Table S12 Number of genes excluded/included after the lower/upper bound cutoff.

Number of genes	M9 glucose	M9 *N*-acetylglucosamine	M9 glycerol
Lower bound 3.00	500	459	196
Upper bound 850.56	251	271	314

Appendix B Figures

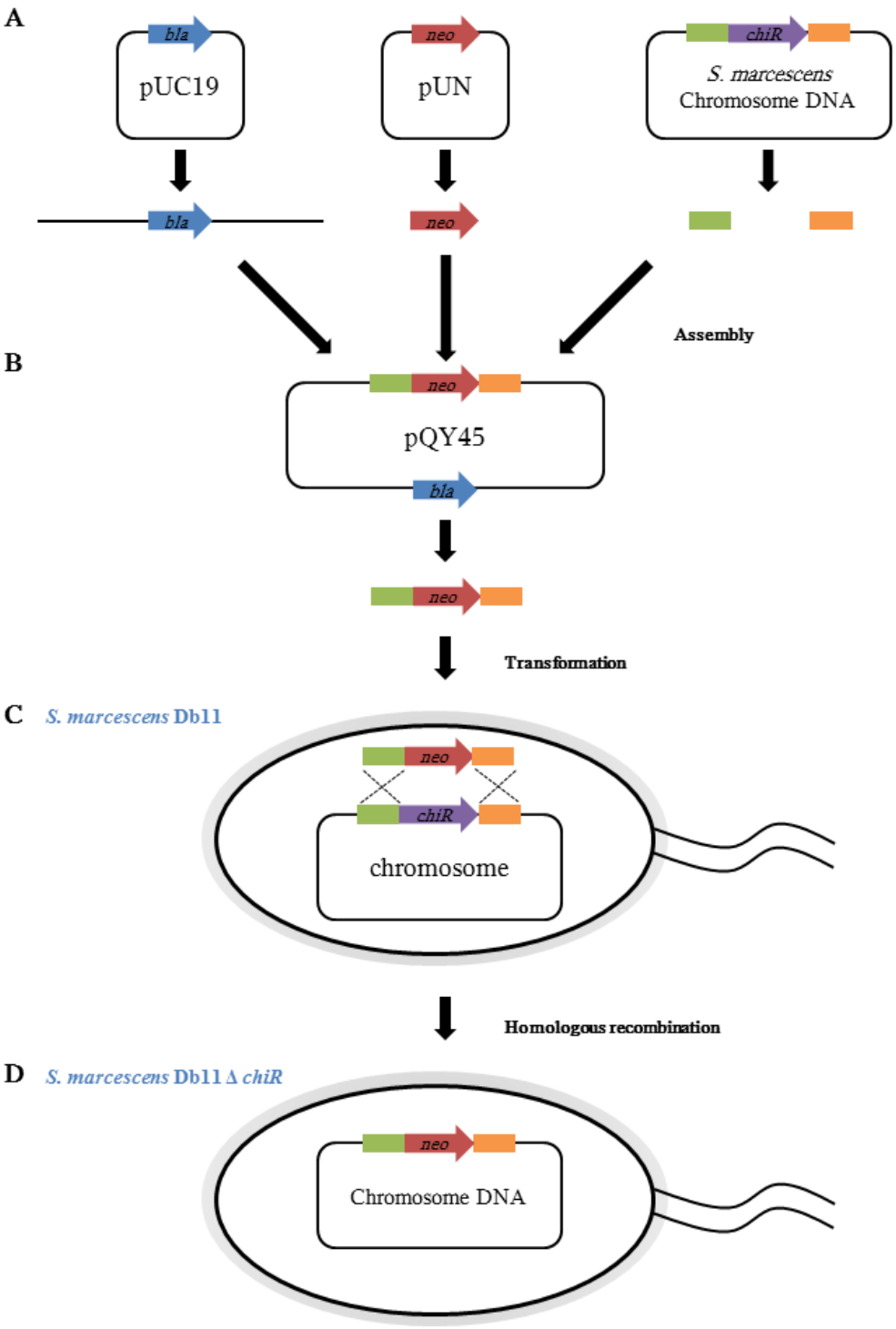


Fig S1A

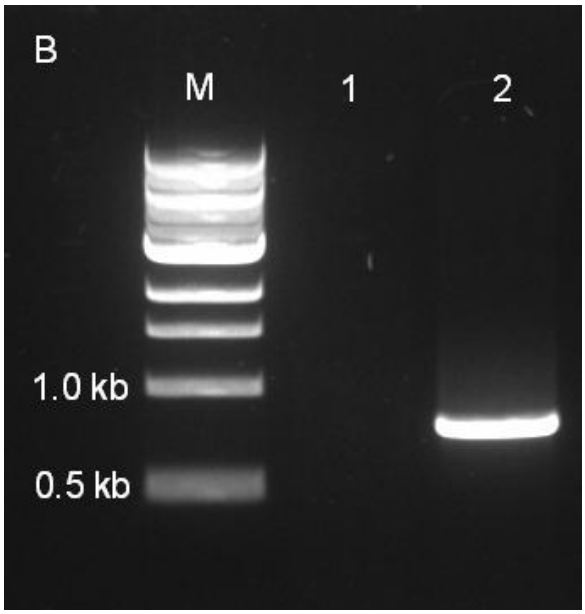


Fig S1B

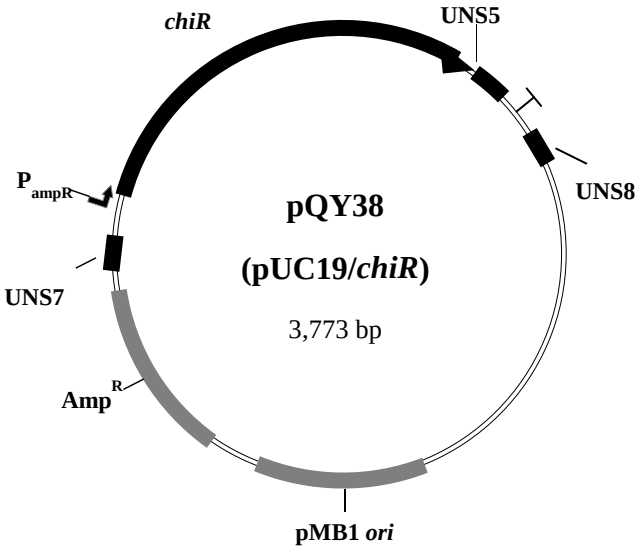


Fig S1C

Figure S1 (A) Overview of the strategy used for replacement of *chiR* in *S. marcescens*. **(**B**)** PCR verification of the deletion mutants. PCR was conducted using wild-type or *chiR* deletion strain genome as template and kan-f/kan-r as primers. Lanes from left to right are 1 kb marker, wild-type (no band) and *chiR* deletion strain (a single ~0.8 kb band). (C) Shuttle vector pQY38 map for overexpression *chiR* in *S. marcescens*.

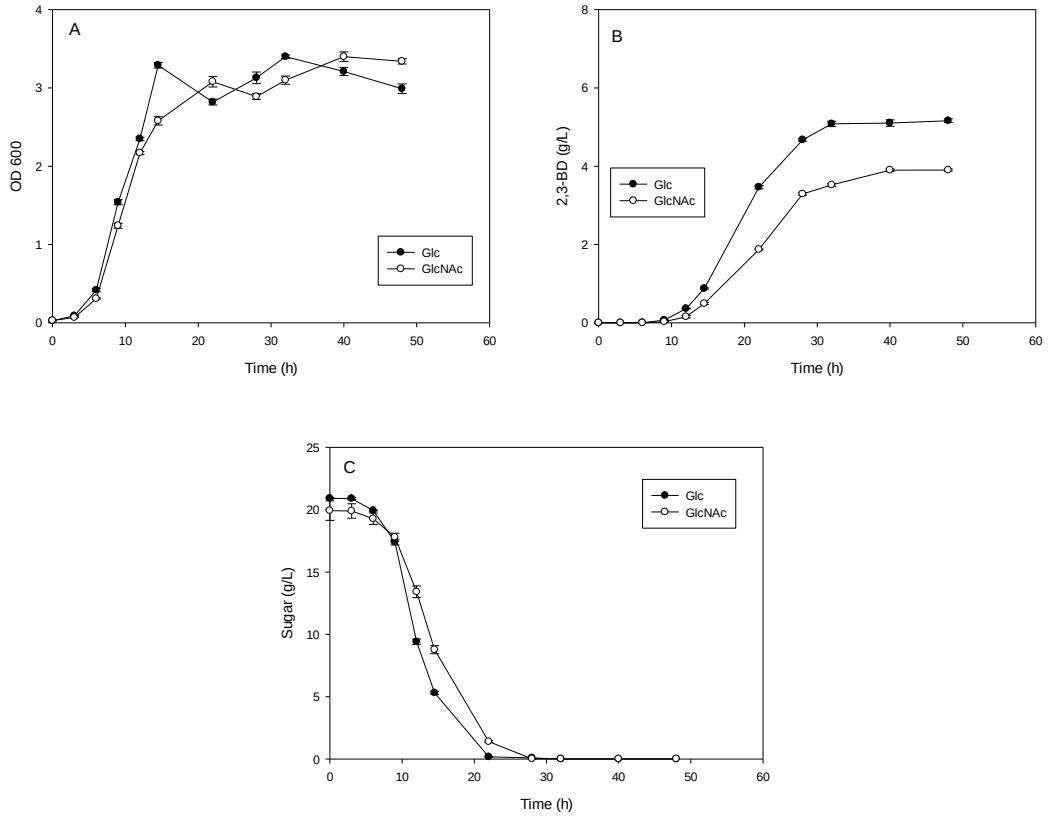


Figure S2 Production of 2,3-BD in *S. marcescens* Db11. (A) Cell growth, (B) 2,3-BD production and (C) sugar consumption was monitored using *S. marcescens* Db11 under M9 medium with 1% yeast extract and 20 g/L glucose (Glc, filled circle) or *N*-acetylglucosamine (GlcNAc, unfilled circle) at 30 ℃ and 100 rpm.

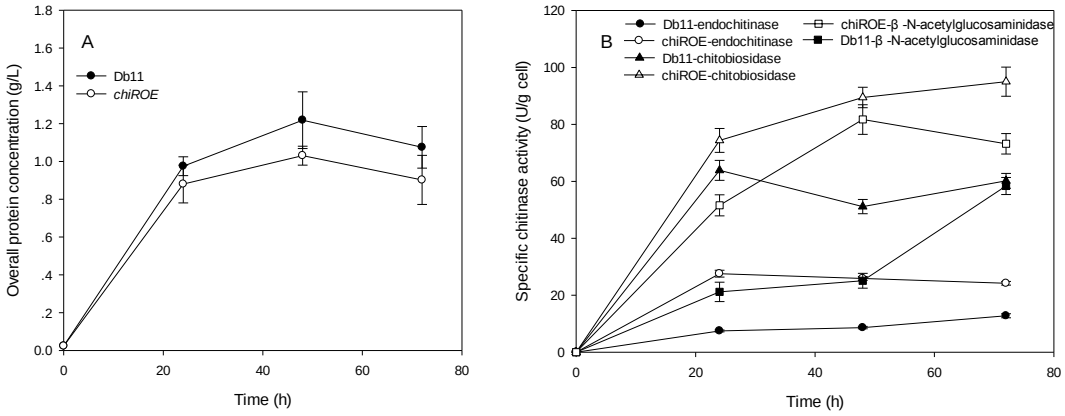


Figure S3 Production of 2,3-BD from *S. marcescens* using untreated chitin. (A) Cell growth using the *S. marcescens* Db11 (filled circle) and the *S. marcescens chiROE* strain (unfilled circle); (B) endochtinase activity (circle), chitobiosidase activity (triangle) and *β-N*-acetylglucosaminidase activity (square) using the *S. marcescens* Db11 (filled) and the *S. marcescens chiROE* strain (unfilled) under M9 medium supplemented with 1% yeast extract and 2% chitin at 30 ℃ and 100 rpm.

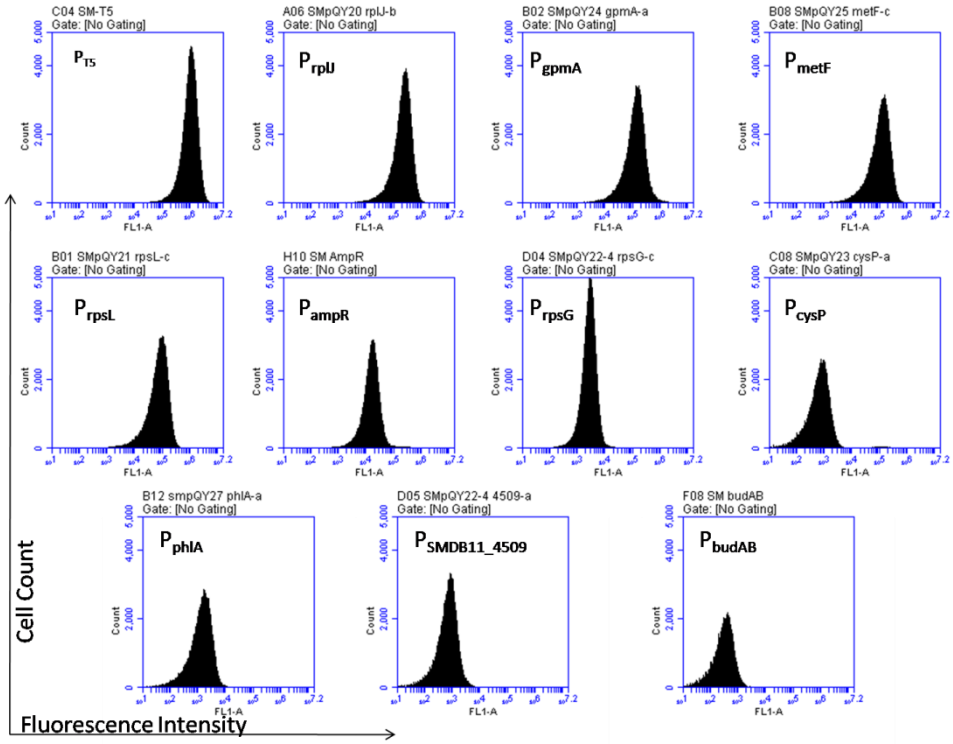


Figure S4 Raw data of fluorescence intensity of *S. marcescens* promoters using flow cytometry.

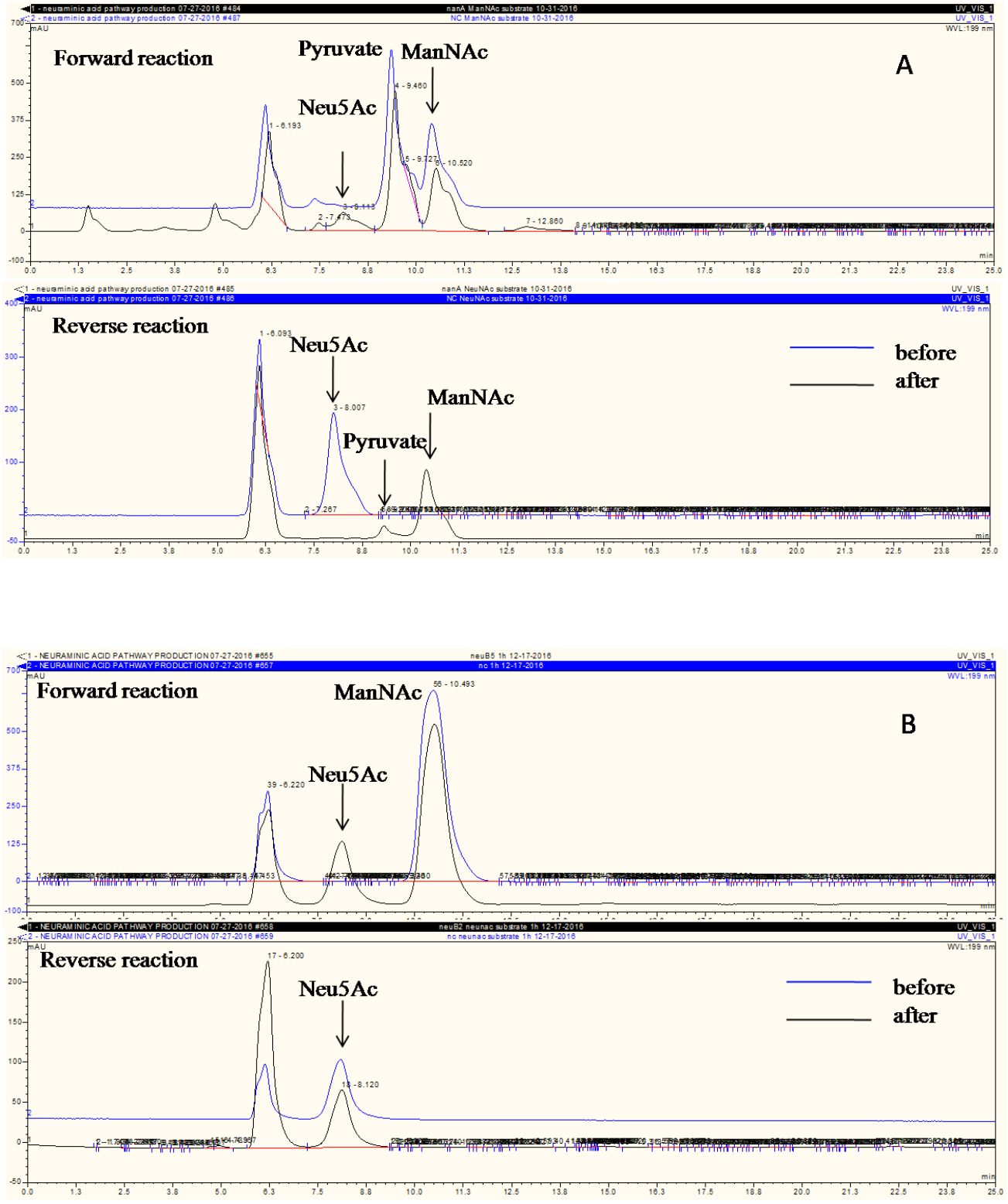


Figure S5 Characterization of (A) NanA and (B) NeuB enzymatic activities.

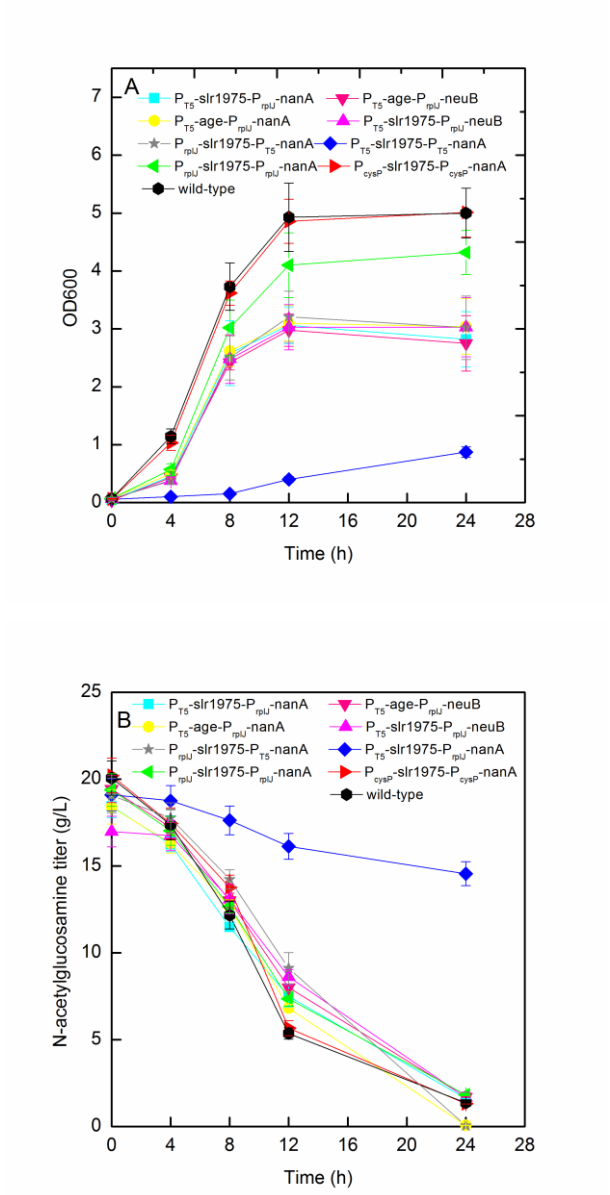


Figure S6 Growth of *S. marcescens* recombinant strains on M9 medium with 0.1% yeast extract and 20 g/L GlcNAc at 30℃ and 200 rpm. (A) OD600; (B) GlcNAc concentration.

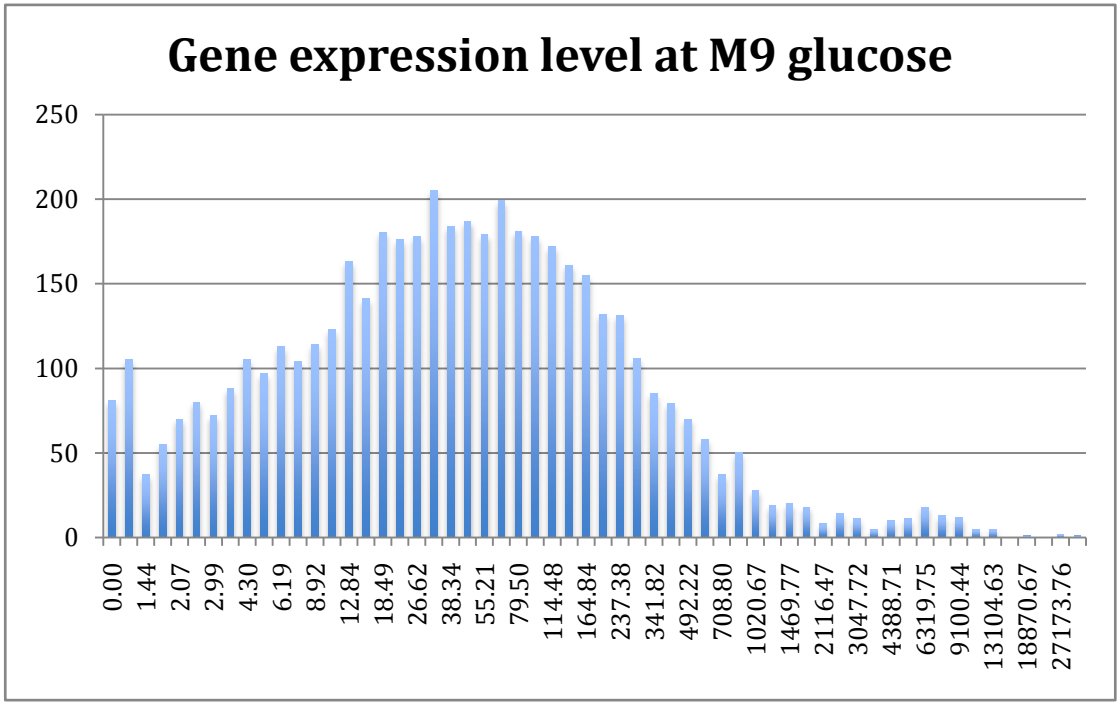


Figure S7 Representative figures of gene expression level distribution under M9 glucose medium growth conditions.

VITA

Qiang Yan was born on September 17th, 1988 in Dalian, China. He attended Dalian Polytechnic University, China from September 2007 to June 2011 earning the degree of Bachelor of Engineering in Bioengineering and Jiangnan University, China from September 2011 to June 2014 earning the degree of Master of Science in Engineering, majoring in Fermentation Engineering. He attended Virginia Commonwealth University earning Doctor of Philosophy in Engineering majoring in Chemical Engineering in April 2018.

PUBLICATIONS

[1] **Yan, Q.,** Fong, S.S. Evolutionary Engineering of *Escherichia coli* After Downregulation of Essential Genes to Improve Carbon Uptake Rates. (In preparation)

[2] **Yan, Q.,** Fong, S.S. Adaptive Evolution *S. marcescens* to Increase Chitin Utilization Efficiency and Genome Re-sequencing Platform Provides Insightful Mechanism. (In preparation)

[3] **Yan, Q.,** Fong, S.S. Design and modularized optimization of one-step production of N-acetylneuraminic acid from chitin in Serratia marcescens (In preparation)

[4] **Yan, Q.,** Fong S.S. A transcriptomics-based genome scale metabolic model of the chatinolytic bacterium Serratia marcescens reveals metabolic capabilities. **(**Submitted**)**

[5] **Yan, Q.,** Fong S.S. Alleviating Carbon Limitations for Metabolic Engineering. ***Current Opinion in Biotechnology*** (Under Revision)

[6] **Yan, Q.,** Fong S.S. Cloning and Characterization of a Chitinase from *Thermobifida fusca* reveals Tfu_0580 Acts as a Thermostable and Acidic Endochitinase. (Submitted)

[7] **Yan, Q.,** Fong S.S. 2017. Strategies and Challenges for Genetic Engineering of Non-Model Bacteria and its Use in Consolidated Bioprocessing. ***Frontiers in Microbiology*** 8:2060

[8] **Yan, Q.,** Fong S.S. 2017. Study of in vitro Transcriptional Binding Effects and Noise Using Constitutive Promoters Combined with UP Element Sequences in *Escherichia coli*. ***Journal of Biological Engineering*** 11:33

[9] **Yan, Q**., Fong, S.S. 2017. Study of ChiR Function in *Serratia marcescens* and its Application for Improving 2,3-Butanediol from Crystal Chitin. ***Applied Microbiology and Biotechnology*** 101(20): 7567-7578

[10] Wu, G., **Yan, Q.,** Jones, A.J., Tang, Y.J., Fong, S.S., Koffas, M.A.G. 2016 Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering. ***Trends in Biotechnology***. 34(8): 652-664

[11] **Yan, Q.,** Fong, S.S. 2015. Bacterial Chitinase: Nature and Perspectives for Sustainable Bioproduction. ***Bioresources and Bioprocessing***. 2:31

[12] **Yan, Q.,** Zheng P., Tao S.T., Dong J.J., 2014. Fermentation Process for Continuous Production of Succinic Acid in a Fibrous Bed Bioreactor. ***Biochemical Engineering Journal*** 91:92-98

[13] **Yan, Q.**, Zheng, P., 2013. Improving the Productivity of Succinic Acid Fermentation Using Fibrous Bed Bio-Reactor by *Actinobacillus succinogenes.* ***Journal of Chemical Technology and Biotechnology*** 40:831-840

[14] Gu, B.H., Zheng, P., **Yan, Q.**, Liu, W., 2014. Aqueous Two-phase System: An Alternative Process for Recovery of Succinic Acid from Fermentation Broth. ***Separation and Purification Technology*** 138:47-54

[15] Zheng, P., Zhang, K.K., **Yan, Q.**, Xu, Y., Sun, Z.H., 2013. Enhanced Succinic Acid Production by *Actinobacillus Succinogenes* after Genome Shuffling. ***Journal of Industrial Microbiology and Biotechnology*** 40:831-840

BOOK CHAPTER

[1] **Yan, Q.,** Fong, S.S. 2016. Biosensors for Metabolic Engineering. In ***Systems Biology Applications in Synthetic Biology***. Chapter 4. 53-70

PATENT

[1] Zheng, P., **Yan, Q.**, One Method for Succinic Acid Production Using a Fibrous-Bed Bioreactor. Chinese Patent **CN103436561B**

CONFERENCES AND PRESENTATIONS

[1] **Yan, Q.,** Fong, S.S. Designing and Creating a Modularized Synthetic Pathway in

*Serratia marcescen*s for Production of *N*-acetylneuraminic acid from chitin. June 24-28, 2018, Metabolic Engineering 12, Munich, Germany (Poster presentation)

[2] **Yan, Q.**, Fong, S.S. Enhance *N*-Acetylneuraminic Acid Production after Pathway Balancing in *Serratia marcescens*, July 2017 3rd International Conference on Systems and Synthetic Biology, Munich, Germany (Oral presentation)

[3] **Yan, Q.,** Fong, S.S. Direct Bioconversion of Chitin into *N*-Acetylneuraminic Acid Production Using *Serratia Marcescens*. May, 2016 Integrated Biological Engineering Annual Meeting, Greensville, South Caralina, USA (Oral presentation)

[4] **Yan, Q.**, Hong, E.S., Fong, S.S. Metabolic modeling of the vaginal microbiome: Methods and Progress Updates. August, 2017 The Human Microbiome: Emerging Themes at the Horizon of the 21st Century Workship, Bethesda, Maryland, USA (Poster presentation)

[5] **Yan, Q.**, Zheng, P., Dong, J.J., Wang, X.L., Succinic Acid Fermentation in a Fibrous-Bed Bioreactor. International Symposium on New Frontiers in Microbiology and Biotechnology. March 2013, Wuxi, China (Poster presentation)

AWARDS AND FELLOWSHIPS

Virginia Commonwealth University Travel Grant Award

Virginia Commonwealth University Dissertation Award Fellowship

Graduate Fellowship of Virginia Commonwealth University

Excellent Master Dissertation Award of Jiangsu Province