

CHALMERS TEKNISKA HÖGSKOLA

KANDIDATARBETE KBTX01-15-09

INSTITUTIONEN FÖR BIOLOGI OCH BIOTEKNIK

---

Utveckling av ett system för  
enzyme evolution *in vivo*

---

Författare:

Ida ALEXANDERSSON  
Martin BOSTRÖM  
Linnea JOHANSSON  
Kajsa MALMBERG  
Sebastian VALENZUELA  
Jonathan ÅKESSON

Handledare:

Verena SIEWERS  
Florian DAVID

Examinator:

Eva ALBERS

19 maj 2015



CHALMERS

## **Förord**

Denna studie har genomförts under våren 2015 vid gruppen för Systembiologi och institutionen för Biologi och Bioteknik, Chalmers Tekniska Högskola. Vi vill rikta ett stort tack till våra handledare Verena Siewers och Florian David för deras goda handledning och brinnande engagemang. Sist men inte minst vill vi också tacka alla anställda på gruppen för Systembiologi för deras tålamod och svar på våra otaliga frågor.

## Abstract

### Developing an *in vivo* enzyme evolution system

By creating mutations in the DNA sequence coding for a protein, the activity of the protein can be modified. If there is insufficient information for the use of rational design, random mutagenesis can be used. If this is done *in vitro*, mutations are easily targeted at sequences, but the many required cycles of transformation and DNA extraction are time-consuming. If random mutagenesis is done *in vivo*, the process can be continuous, but the mutations often affect the entire genome. It would be beneficial to be able to combine the advantages of both *in vivo* and *in vitro* methods.

The purpose of this study is to construct a system capable of performing random but region-specific mutations, using the CRISPR/Cas system. To test the mutating system, it was targeted in yeast at the wax ester synthase WS2, which synthesises fatty acid alkyl esters from fatty acids and alcohols. Mainly alcohols with long carbon chains are used as substrate by WS2. In the presence of oleic acid, which is toxic for yeast, mutations that favour the use of alcohols with shorter carbon chains could be selected for. Together with ethanol, produced naturally by yeast, oleic acid can be used to synthesise oleic acid ethyl esters, which are less toxic than oleic acid.

Successful construction of vectors expressing WS2 and a dCas9-Mag1 fusion protein was achieved. The latter was to be targeted at a specific part of *WS2* with a gRNA molecule. The vectors were transformed to the yeast strain *Saccharomyces cerevisiae* JV04, which showed increased resistance to oleic acid in some colonies after induction of the mutating system with galactose. Unfortunately, no mutations were observed in said colonies after sequencing of *WS2*.

## Sammanfattning

Genom att skapa mutationer i DNA-sekvensen som kodar för ett protein kan proteinets aktivitet modifieras. Om tillräcklig information om enzymets struktur för rationell design saknas kan slumpmässig mutagenes användas. Om denna utförs *in vitro* kan mutationerna lätt riktas, men de många cykler av transformation och DNA-extraktion som behövs är tidskrävande. Utförs slumpmässig mutagenes istället *in vivo* kan det utföras kontinuerligt, men mutationerna påverkar ofta hela genomet. Det hade varit önskvärt att kunna kombinera fördelar från både *in vivo*- och *in vitro*-metoder.

Syftet med studien är att konstruera ett system som kan utföra slumpvisa men till en specifik sekvens riktade mutationer, med utgångspunkt i CRISPR/Cas-systemet. Som test av det muterande systemet riktades det för att i jäst mutera vaxestersyntaset WS2, som syntetiserar fettsyraalkylestrar av fettsyror och alkoholer. Det är huvudsakligen alkoholer med långa kolkedjor som WS2 använder som substrat. Vid närvaro av oljesyra, som är toxiskt för jäst, skulle mutationer som gynnar användandet av kortare alkoholer, till exempel etanol, kunna selekteras. Oljesyra tillsammans med etanol, som jästen själv producerar, kan med muterat WS2 bilda oljesyraetylestrar, som är mindre giftiga än oljesyra.

Lyckad konstruktion utfördes av vektorer som uttrycker WS2 och ett dCas9-Mag1-fusionsprotein, som med hjälp av en gRNA-molekyl skulle riktas till en specifik del av *WS2*. Dessa vektorer transformerades till jäststammen *Saccharomyces cerevisiae* JV04, som efter induktion med galaktos av det muterade systemet visade högre resistens mot oljesyra i vissa kolonier. Dessvärre kunde inga mutationer hos kolonierna urskiljas efter sekvensering av *WS2*.

## Innehåll

<b>1</b>	<b>Inledning</b>	<b>1</b>
1.1	Bakgrund . . . . .	1
1.2	Syfte . . . . .	2
<b>2</b>	<b>Teori</b>	<b>2</b>
2.1	CRISPR/Cas-systemet . . . . .	2
2.2	Test av det muterande systemet . . . . .	3
2.3	Design av den variabla delen i gRNA-molekylen . . . . .	4
2.4	Konstruktion av system . . . . .	5
2.5	Gibsonsammanfogning . . . . .	6
2.6	Transformation med hjälp av LiAc-metoden . . . . .	8
<b>3</b>	<b>Projektets avgränsningar</b>	<b>8</b>
<b>4</b>	<b>Metod</b>	<b>9</b>
4.1	Plasmidkonstruktion . . . . .	9
4.2	Transformation till <i>E. coli</i> samt selektion . . . . .	10
4.3	Transformation till jäst samt selektion . . . . .	11
4.4	Selektion och sekvensering . . . . .	12
<b>5</b>	<b>Resultat och diskussion</b>	<b>14</b>
5.1	Restriktionsanalys för p413TEF samt p416TEF . . . . .	14
5.2	PCR-amplifiering av fragment . . . . .	14
5.3	Restriktionsanalys för p413ins samt p416ins . . . . .	15
5.4	Sekvensering av p413ins samt p416ins . . . . .	16
5.5	Effektivitetskontroll av transformation till jäst . . . . .	17
5.6	Koloni-PCR som kontroll för närvaro av p413ins samt p416ins	18
5.7	Test av kritisk koncentration oljesyra . . . . .	20
5.8	Odling av kolonier där systemet inducerats . . . . .	21
5.9	Koloni-PCR för amplifiering av <i>WS2</i> till slutgiltig sekvensering	22
5.10	Droppanalys med inducerade, enskilda kloner . . . . .	23
5.11	Koloni-PCR för <i>dCas9</i> för enskilda kloner . . . . .	24
5.12	Sekvensering av <i>WS2</i> efter induktion av systemet . . . . .	24
5.13	Utökade experiment och kontroller . . . . .	27
<b>6</b>	<b>Slutsatser</b>	<b>28</b>
	<b>Referenser</b>	<b>29</b>
<b>A</b>	<b>Processchema</b>	
<b>B</b>	<b>Originalplasmider med restriktionsfragment för <i>SacI</i> och <i>XhoI</i></b>	

**C Plasmidkonstruktioner**

**D Förberedelse av och transformation till kompetenta jäst-celler**

**E PCR**

**F Laborationsinstruktioner för koloni-PCR**

**G Klyvningsställen *PvuII*, p413 och p416**

**H DNA-sekvenser för primrar**

H.1 Primrar för förlängning av segment inför Gibsonsammanfogning

H.2 Primrar för sekvensering

H.3 Primrar ytterligare kontroll av insert

**I DNA-sekvenser för konstruerade plasmider**

I.1 p413ins

I.2 p416ins

# 1 Inledning

Idag har mikroorganismer ett flertal applikationer inom olika industrier, exempelvis för produktion av enzymer och andra proteiner [1]. Den ökade användningen av mikroorganismer inom industrin beror på flera faktorer, bland annat att genmodifiering av metabola vägar är möjlig men också att det är möjligt att odla upp mikroorganismer i storskalig skala. På grund av ökad kunskap om mikroorganismer och deras metabola vägar så har vissa organismer manipulerats med så kallad "metabolic engineering". Exempelvis har jäst modifierats till att omvandla enklare sockermolekyler till olika typer av biobränsle [2]. Metabolic engineering kan utnyttjas för att tillverka enzymer, proteiner och metaboliter i önskade mängder. Då manipuleras de metabola vägarna, exempelvis för att en önskad produkt ska bildas, eller för att optimera hur mycket av den som bildas. Ett vanligt tillvägagångssätt för att förbättra produktion är att optimera aktiviteten hos ett enzym i de metabola vägarna [3].

## 1.1 Bakgrund

Ett enzyms aktivitet kan ändras genom att skapa mutationer i DNA-sekvensen som kodar för enzymet i fråga. Mutationerna som introduceras kan ge önskade egenskaper hos enzymet såsom förändring av temperaturoptimum eller produkt- och substratspecificitet. Om det inte finns tillräckligt med information om enzyms struktur för rationell design kan slumpmässig mutagenes användas. Då skapas ett flertal olika mutationer som kan ge upphov till olika förändringar i enzymets struktur. I processen för slumpmässig mutagenes *in vitro* krävs efter varje mutationssteg transformation av plasmider och screening av enzymernas nya egenskaper eller selektion för de enzymen med önskad aktivitet. De enzym som sedan selekterats går därefter igenom mutagenesen igen för att förbättras ytterligare [4]. Detta gör att det blir en tidskrävande process. En vanlig metod inom slumpmässig mutagenes *in vitro* är error-prone PCR [5]. Då utnyttjas Taq-polymeras, som har högre felfrekvens än andra polymeras och kan därför användas för att utföra slumpmässiga mutationer på en specifik sekvens. Dock har Taq-polymeras en tendens att mutera nukleotiderna adenin och tymin oftare än guanin och cytosin, varför vissa mutationer är svårare att åstadkomma än andra [6].

Om slumpmässig mutagenes istället utförs *in vivo* behöver inte processen nödvändigtvis ske i cykler av transformation och DNA-extraktion, utan kan istället utföras kontinuerligt. Dock har flertalet metoder för slumpmässig mutagenes *in vivo* nackdelen att de inte är specifikt riktade till en viss gen, utan istället påverkar hela genomet. Exempel på metoder för slumpmässig mutagenes *in vivo* är behandling med kemikalier vid inkubering och att ut-

nyttja stammar med mutationsframkallande gener. En nackdel med dessa metoder är att hur mycket en viss gen kan muteras begränsas av att cellen inte kan överleva om genomet i sin helhet får för många eller för allvarliga ogynnsamma mutationer [7].

I detta projekt vill vi skapa ett system som kombinerar fördelar med *in vivo* och *in vitro* slumpmässig mutagenes. Det ska fungera kontinuerligt *in vivo* och därmed inte behöva utföras i tidskrävande cykler, men samtidigt kunna riktas att mutera kring en specifik DNA-sekvens likt error-prone PCR.

## 1.2 Syfte

Syftet med rapporten är att redogöra för utvecklandet av ett system som slumpvis ska kunna mutera DNA *in vivo* på en valbar och därmed specifik del av ett genom.

## 2 Teori

Det system som ska konstrueras för att rikta mutationer *in vivo* behöver dels en muterande del som kan utföra slumpvisa mutationer, dels en lokaliserande del som kan riktas mot specifika DNA-sekvenser. Den lokaliserande delen baseras på CRISPR<sup>1</sup>/Cas<sup>2</sup>-systemet. De delar som krävs för att systemet ska fungera kommer att sammanfogas i plasmider med Gibsonsammanfogning. För transformation av dessa plasmider till jäst, som är den organism systemet kommer att testas i, används LiAc-metoden.

### 2.1 CRISPR/Cas-systemet

CRISPR/Cas-systemet finns i många bakterier och används naturligt som ett försvar mot virus. Korta RNA-bitar uttryckta från CRISPR-regioner bildar komplex med Cas-proteiner, som är endonukleaser. RNA-bitarna kombineras på ett sådant vis att de kan baspara till någon sekvens i virusets genetiska material, och positionerar på så sätt Cas-proteinet för klyvning [8]. Det har nyligen visats att CRISPR/Cas-systemet kan utnyttjas för att på ett regions-specifikt vis påverka uttrycket av gener i olika organismer. Då nukleasaktivitet ej är önskvärd kan en katalytiskt inaktiv variant av endonukleaset Cas9, kallat dCas9, paras ihop med gRNA<sup>3</sup>, en konstruktion som liknar CRISPR-relaterat RNA, och sedan riktas mot specifika bassekvenser. gRNA har en

---

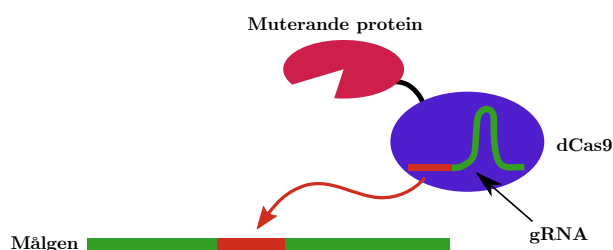
<sup>1</sup>Clustered Regularly Interspaced Short Palindromic Repeats

<sup>2</sup>CRISPR-associated protein

<sup>3</sup>guide-RNA



tydligt definierad del som binder till dCas9, och en variabel del som basparar till en DNA-sekvens på målgenen. Genom att designa den variabla delen kan gRNA-molekylen, och därmed det medföljande dCas9, riktas mot specifika sekvenser [9]. Beroende på om dCas9 är länkat till något annat protein eller inte, och i så fall vilket, kan olika funktioner uppnås, till exempel geninaktivering [9] och transkriptionsreglering [8]. I detta arbete länkas dCas9 till ett muterande protein, se figur 1, för att uppnå slumpvisa mutationer på målsekvensen.

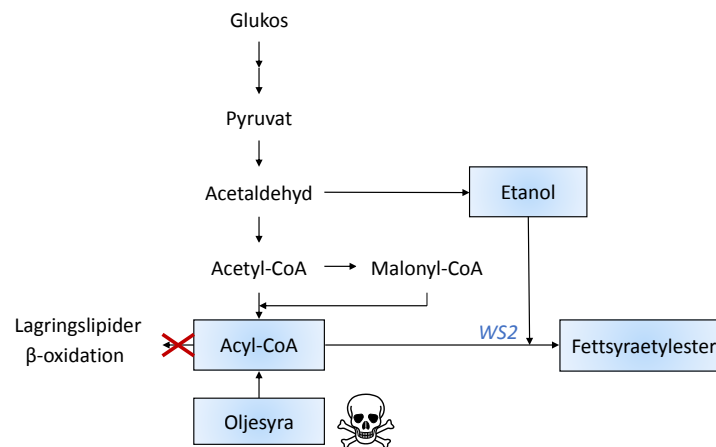


**Figur 1:** dCas9, här sammanfogat med ett muterande enzym, riktas till en specifik del av en gen. Detta sker med hjälp av en gRNA-molekyl, vars variabla del (rödmarkerad i figuren) är komplementär till målsekvensen. Det muterande proteinet kan sedan utföra slumpvisa mutationer på denna.

## 2.2 Test av det muterande systemet

Målgenen som valts för att testa det muterande systemet *in vivo* kodar för vaxestersyntas, och benämns hädanefter *WS*. Vaxestersyntaser är involverade i syntesen av fettsyraalkylestrar, även kallade vaxestrar, från alkoholer och acyl-coenzym A. Om etanol skulle användas som alkohol skulle den bildade fettsyraetylestern kunna användas som dieselbränsle. Genom att då utnyttja jästens egna fermentativa process, där etanol är en produkt, kan man skapa ett biobränsle om *WS* är närvarande. Anledningen till att *Saccharomyces cerevisiae* utnyttjas som värdorganism är att organismen har flera fördelar att producera vaxestrar, till exempel dess korta generationstid samt den omfattande kunskapen om dess fettsyrametabolism. *WS* tenderar att syntetisera vaxestrar från alkoholer med längre kolkedjor än etanol. Dock kan *WS* fortfarande bilda dessa produkter med etanol, men inte i lika stor mängd [10]. Detta problem skulle kunna åtgärdas genom att introducera mutationer så att *WS* blir mer substratspecifikt mot etanol. Fem vaxestersyntaser jämfördes i en studie av Shi et al., och där visades att vaxestersyntaset från *Marinobacter hydrocarbonoclasticus*, hädanefter benämnt *WS2*, har högst substratspecificitet mot etanol [10]. Därför väljs *WS2* i denna studie.

I jäststammen *JV04*, som används under projektet och härstammar från *S.*



**Figur 2:** Ett muterat vaxestersyntas gör att JV04, som saknar förmåga till  $\beta$ -oxidation samt att bilda lagringslipider, kan växa vid tillsats av höga koncentrationer av oljesyra. Detta möjliggör bildandet av fettsyraetylestrar, som är mindre giftiga än oljesyra, av acyl-CoA och etanol. Etanolen kommer från jästens egna fermentativa process.

*cerevisiae* CEN.PK 113-7D, har fem gener eliminerats. Fyra av dessa gener är *ARE1*, *ARE2*, *DGA1*, och *LRO1*, vilka ger jästen egenskapen att naturligt kunna inkorporera acyl-CoA i lagringslipider, se figur 2. Avsaknaden av dessa gener leder till att egenskapen att kunna inkorporera acyl-CoA i lagringslipider elimineras. Den femte genen som eliminerats, *POX1*, kodar för ett enzym som bestämmer flödet av  $\beta$ -oxidationen. I  $\beta$ -oxidationen genereras energi genom att bryta ner fettsyror [11]. För att analysera om det muterande systemet har fungerat och om eventuella mutationer har uppstått på *WS2* odlas jästen i närvaro av oljesyra,  $C_{18}H_{34}O_2$ . Detta leder till att mängden acyl-CoA i jästen kommer att öka, vilket är giftigt för cellen, detta på grund av att acyl-CoA ej längre kan inkorporeras i lagringslipider [12], som illustreras i figur 2. För att jästen ska kunna växa i dessa förhållanden krävs att den kan omvandla acyl-CoA tillsammans med etanol till mindre giftiga fettsyraetylestrar, vilket katalyseras av *WS2*. Om jästen erhållit mutationer som i tillräckligt hög grad ökar användandet av etanol som substrat, tillsammans med acyl-CoA, för bildning av fettsyraetylestrar så kommer cellerna att växa vid närvaro av oljesyra.

### 2.3 Design av den variabla delen i gRNA-molekylen

Den variabla delen av gRNA måste designas så att dCas9-Mag1-komplexet, där Mag1 är det muterande proteinet, muterar *WS2*, se figur 1. Precis bredvid den sekvens till vilken gRNA-molekylen ska binda behöver en PAM<sup>4</sup>-

<sup>4</sup>Protospacer-Adjacent Motif

sekvens finnas, vilket är antingen de tre baserna NGG eller NAG, där N är en godtycklig bas. PAM-sekvensen utnyttjas av Cas9 (eller dCas9) för att öppna DNA-spiralen, så att gRNA-molekylen kan binda in [13]. NGG är att föredra framför NAG, då den förstnämnda resulterar i högre Cas9-aktivitet [14], varför en sekvens bredvid en NGG-PAM-sekvens valdes.

Troligtvis är det mutationer i det aktiva sätet för WS2 som påverkar dess aktivitet mest, och därför bör gRNA-molekylen ej binda där. Aminosyrorna i position 118, 140, 141, 145, 270 och 271 har visat sig vara viktiga för enzymets aktivitet [15]. Mutationer av aminosyra 360 kan resultera i att WS2 i större utsträckning kan utnyttja kortare alkoholer som substrat [16]. Närhet till kodonen för dessa aminosyror undveks därför då den variabla sekvensen valdes, eftersom gRNA-molekylen utgör ett steriskt hinder för det muterande proteinet vid inbindningssekvensen.

För att undvika att dCas9-Mag1-komplexet blockerar promotorn valdes en sekvens på 3'-änden av *WS2* som mall för den variabla sekvensen på gRNA-molekylen. Sekvensen som valdes är TAAGGGAGCCAGCGTAAGAC.

## 2.4 Konstruktion av system

Systemet konstrueras av två vektorer, innehållande nödvändiga sekvenser för att det muterande systemet ska fungera.

Den första vektorn, se figur 20 i appendix C, har som syfte att skapa ett proteinkomplex som kan ta sig tillbaka in i cellkärnan och skapa riktade mutationer. Vektorn består av fragmenten *pGAL1*, *dCas9-NLS*<sup>5</sup>, samt *MAG1*. *pGAL1* är en inducerbar promotor som kan induceras vid närvaro av galaktos. *dCas9-NLS* uttrycker dCas9 som med hjälp av gRNA kan riktas mot *WS2*. dCas9 uttrycks tillsammans med NLS som möjliggör att det muterande systemet, som uttryckt protein, kan ta sig in i cellkärnan. *MAG1* kodar för det muterande enzymet 3-metyladenin-DNA-glykosylas, som beskrivs i en studie av Finney-Manchester och Maheshri [17]. Mag1 är ett naturligt förekommande protein i jäst som normalt är involverat i repareringssystemet för DNA. När Mag1 är överuttryckt kan det skapa mutationer [17]. För att Mag1 och dCas9 ska kunna veckas korrekt infogas en linker mellan dem bestående av fem glycinrester.

Den andra vektorn, se figur 21 i appendix C, innehåller sekvensen som kodar för *WS2* och promotorn *pKEX2*. *pKEX2* är en svag promotor som hindrar att *WS2* överuttrycks, vilket är viktigt då det är lättare att utskilja mutanter vid lägre oljesyrakoncentrationer. Om *WS2* överuttrycks kan jästen få en högre tillväxt i närvaro av oljesyra utan att mutationer introducerats på *WS2*.

<sup>5</sup>Nuclear Localization Sequence

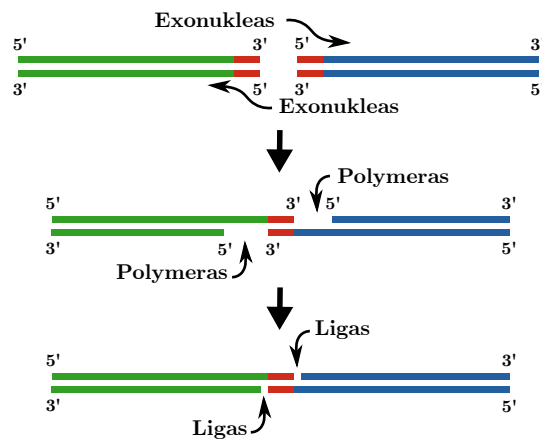
Nivån av uttryckt WS2 har utretts av Nacken et al. [18]. På vektorn finns även genen som kodar för gRNA, vars variabla del är designad att baspara med *WS2*. gRNA-kassetten har en SNR52-promotor och en SUP4-terminator, vilket gör att gRNA:t ej modifieras och därmed inte lämnar cellkärnan efter att det uttryckts [19]. Detta gör att det muterande systemet innehållande gRNA, dCas9 och Mag1 bör kunna binda till *WS2* och skapa slumpvis, riktade mutationer på genen.

Till den första vektorn används en plasmidryggrad tillhörande p413TEF, se figur 18 i appendix B. Plasmiden p413TEF innehåller en gen, *ampR*, som ger upphov till ampicillinresistens, samt en gen som uttrycker ett enzym som är involverad i biosyntesen av histidin, *HIS3*. Till den andra vektorn används en plasmidryggrad tillhörande p416TEF, se figur 19 i appendix B. Plasmiden p416TEF innehåller, precis som p413TEF, genen *ampR*, samt en gen som uttrycker ett enzym som är involverat i biosyntesen av uracil, *URA3*. Dessa två plasmider är centromeriska plasmider, vilket resulterar i en kopia per transformerad cell [20].

Ampicillinresistens används för att selektera för *Escherichia coli* som tagit upp plasmiderna genom att odla *E. coli* på medium innehållandes ampicillin. För att senare kunna selektera för de jästkolonier som inorporerat båda plasmiderna utnyttjades en jäststam som ej kunde producera uracil eller histidin. Genom att odla transformerad jäst på medium med avsaknad av uracil och histidin kan de jästkolonier som inorporerat bägge plasmiderna selekteras fram.

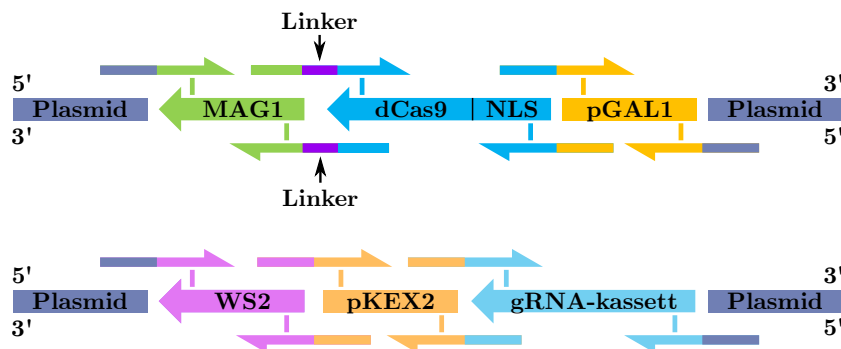
## 2.5 Gibsonsammanfogning

För att konstruera de två plasmiderna ovan utnyttjas Gibsonsammanfogning, se figur 3, som är en så kallad “chew back and anneal”-metod. Ett exonukleas som verkar från 5'-ändan angriper de dubbelsträngade segmenten som ska sammanfogas, tar bort ett antal baser på den ena strängen, och exponerar på så vis ett stycke enkelsträngat DNA. Segmentens sekvenser är valda så att varje ände är homolog till den ände den ska sammanfogas med. Segmenten basparar spontant så att komplementära ändar ansluts, varpå ett polymeras ersätter de baser som exonukleaset tog bort innanför de matchande regionerna. Till sist reparerar ligaser alla jack. Eftersom exonukleaset bara kan angripa från ändarna är sammanfogade segment säkra från oönskad klyvning. Om Gibsonsammanfogning används för att skapa cirkulära produkter, till exempel plasmider, kommer dessa därför att kunna ansamlas utan risk för angrepp, eftersom de saknar ändar. Flera segment kan sammanfogas samtidigt om varje ände endast är komplementär till den ände med vilken den ska sammanfogas [21].



**Figur 3:** I Gibsonsammanfogning utnyttjas tre olika enzymer. Ett exonukleas avlägsnar baser från 5'-ändan av segmenten. Ändar som ska sammanfogas är homologa och basparar därför till varandra. Ett polymeras ersätter de baser som tagits bort utanför de homologa regionerna. Slutligen reparerar ett ligas jacken som återstår. Figuren är modifierad efter en bild av New England Biolabs® [22].

För att erhålla de homologa ändrar som är nödvändiga på segmenten som ska anslutas med Gibsonsammanfogning används PCR. Segmenten förlängs med hjälp av primrar som basparar längst ut på varje segment, och har en del som är komplementär till den ände som segmentet ska kopplas samman med. För att erhålla den linker bestående av fem glycinrester som ska finnas mellan Mag1 och dCas9 infogas den lämpliga sekvensen i de primrar som förlänger de relevanta ändarna, se figur 4.



**Figur 4:** Primrar för konstruktion av de homologa ändrar som är nödvändiga i Gibsonsammanfogning för p413 (överst) samt p416 (underst). Vilket segment som primrarna hör till indikeras med streck mellan primer och segment, som är färgkodade efter homologi. För primrarna mellan *MAG1* och *dCas9* indikeras även sekvensen som behövs för att infoga en linker av fem glycinrester.

## 2.6 Transformation med hjälp av LiAc-metoden

Transformation av en cell innebär upptag och inkorporering av för organismen främmande genetisk material. Transformationsmetoden som används i projektet är LiAc<sup>6</sup>-transformation. I metoden används litiumkatjoner, tillsammans med PEG<sup>7</sup>, för att DNA ska tas upp i jästcellerna. Alternativt hade andra katjoner från alkalimetaller kunnat användas, men litium har visats vara mest effektiv för denna metod [23]. Cellerna utsätts för en värmechock för att stimulera upptaget av DNA. Metoden har visats vara mycket effektiv, kräver ingen specialutrustning, samt är ekonomiskt förmånlig att använda [24]. I projektet används denna metod för att transformera in plasmidkonstruktionerna i jästcellerna.

## 3 Projektets avgränsningar

Följande avgränsningar gjordes under projektet på grund av begränsad tid. Jäst med ett muterat vaxestersyntas har, som tidigare nämnts, potential att kunna användas inom produktion av biobränsle. Detta var dock ett långsiktigt mål med projektet som det ej fanns tid till att undersöka. I projektet användes *WS2* enbart som en försöksgen som med fördel skulle kunna muteras *in vivo*. Vår fokus har därför ej lagts på att ta fram önskade egenskaper hos *WS2* utan på att ta fram ett system som kan användas för att genomföra riktade mutationer *in vivo*. Den första avgränsningen med projektet var därför att endast försöka påvisa att *WS2* hade muterats av det muterande systemet, men inte att karaktärisera själva enzymet *WS2* efter att det muterats.

En andra avgränsning för projektet var att jästens genom ej modifierades. För att förenkla inkorporeringen av gener användes plasmider istället för till exempel homolog rekombination på jästens kromosomer.

Under projektet användes enbart jästen *S. cerevisiae* som värdorganism för det muterande systemet och systemet riktades endast mot *WS2*. Tanken är dock att systemet även ska kunna riktas mot andra gener och användas i andra organismer. Då gRNA binder till en specifik sekvens blockeras det närliggande området från modifiering, på grund av steriska hinder. För att man ska kunna mutera varje del av en gen behöver därför minst två gRNA-molekyler med olika sekvenser användas. Detta skulle leda till en förlängning av det laborativa arbetet. Då tidsaspekterna för det laborativa arbetet var begränsade under projektet, resulterade det i att bara en gRNA-sekvens användes. gRNA:t antogs även ha bundit in till målgenen *WS2* och därför

---

<sup>6</sup>Litiumacetat

<sup>7</sup>Polyetylenglykol

lades ingen fokus på att undersöka eventuella ospecifika mutationer på andra platser i genomet. Då den variabla delen av gRNA designas väljs med fördel en sekvens som inte återkommer för ofta i organismens genom; detta för att inbindningen ska bli så specifik som möjligt [25]. Eftersom målet med detta projekt främst är att visa att systemet fungerar har inte stor vikt lagts vid att hitta en sekvens som är unik.

För att avgöra om mutagenessystemet fungerar som avsett borde en kontroll av huruvida dCas9 med hjälp av gRNA binder in till målsekvensen ha utförts. Utan en sådan kontroll är det möjligt att Mag1 har muterande aktivitet som ej är riktad, vilket inte var målet med detta projekt. Kontrollen skulle till exempel kunnat ha varit att testa systemet utan en gRNA-molekyl, eller med fel gRNA-molekyl. På grund av tidsbrist utfördes ej en sådan kontroll.

## 4 Metod

Följande avsnitt beskriver de metoder och tillvägagångssätt som användes i studien. Avsnittet beskriver först hur plasmider konstruerades samt hur fragment sammanfogades med hjälp av Gibsonsammanfogning. Därefter följer en beskrivning av hur plasmiderna transformerades till samt amplifierades i *E. coli* och hur kontroll samt selektion för detta utfördes. Avsnittet beskriver också hur utvalda plasmider transformerades till jäst med hjälp av LiAc-transformation samt de selektionssteg som gjordes. Sista stycket behandlar de avslutande stegen som utfördes i studien. Dessa steg innefattar inducering av det muterande systemet, odling av kolonier samt selektering med selektionstryck från oljesyra. För processchema över den laborativa delen se figur 17 i appendix A.

### 4.1 Plasmidkonstruktion

Segmenten för p*GAL1*, *MAG1* och p*KEX2* amplifierades från genomiskt DNA. gRNA-kassetten beställdes från leverantören GenScript medan *WS2* erhöles från tidigare gjorda experiment. Vektorerna p413TEF och p416TEF, samt en plasmid innehållandes *dCas9* och NLS, erhöles från *E. coli*, som inokulerades med LB-medium och tillsatt ampicillin. Cellerna inkuberades vid 37 °C över natten i en skakinkubator. Plasmiderna extraherades med GeneJET™ Plasmid Miniprep Kit enligt givna instruktioner, med undantaget att vatten användes istället för elueringsbufferten. Plasmidkoncentrationerna samt deras renhet mättes med en biofotometer.

För att linjärisera vektorn och isolera ryggraden av p413TEF och p416TEF användes restriktionsenzymerna *SacI* och *XhoI*. 700 ng av respektive vektor

tillsattes till vardera reaktionsblandningen (0,5  $\mu$ l *SacI* och *XhoI* samt 2  $\mu$ l nedbrytningsbuffert) som späddes till en total volym av 20  $\mu$ l med vatten. Gelelektroforesen (80 V, 50 min) kördes, varefter de större fragmenten skars ut ur gelen, se figur 5b, och renades fram med GeneJET<sup>TM</sup> Gel Extraction Kit. Nanodrop 2000c användes för att mäta koncentrationerna av de klippta vektorerna.

Primrar tillhörande varje fragment späddes till en koncentration av 100  $\mu$ M, se tabell 6 och 7 i appendix H.1. Därefter späddes proverna 1:10. Komponenterna, förutom Phusion-polymeras, adderades enligt tabell 2 i appendix E till en slutvolym av 50  $\mu$ l. I sista steget innan PCR adderades Phusion-polymeras till alla prover. Två olika PCR-program krävdes, ett för *dCas9* samt ett för resterande fragment, då *dCas9* hade en längre förlängningstid. Se tabell 4 i appendix E för PCR-schemat för *dCas9*. Det andra PCR-schemat kördes för de övriga fragmenten, se tabell 3 i appendix E. Som en kontroll av fragmentlängderna efter PCR utfördes en gelelektrofores (80 V, 50 min) för alla prover. Banden i gelen för varje fragment skars ut och gelen fotograferades, se figur 5a och 5b. Fragmenten renades fram med GeneJET<sup>TM</sup> Gel Extraction Kit och koncentrationerna mättes med Nanodrop 2000c.

Fragmenten sattes ihop i två separata plasmider, p413TEF samt p416TEF, med Gibsonsammanfogningsmetoden, se figur 4. Två reaktionsblandningar gjordes med *WS2*, *pKEX2* samt gRNA-kassetten i p416TEF respektive *MAG1*, *dCas9* samt *pGAL1* i p413TEF. Volymen av respektive fragment baserades på tidigare mätta koncentrationer med Nanodrop 2000c, samt enligt Gibson Assembly<sup>®</sup> Master Mix Instruction Manual. 10  $\mu$ l 2XGibson Assembly Master Mix tillsattes till vardera av reaktionsblandningarna till en slutvolym av 20  $\mu$ l och inkuberades vid 50 °C i 60 minuter.

## 4.2 Transformation till *E. coli* samt selektion

De erhållna plasmiderna från Gibsonsammanfogningen transformerades därefter till kompetenta *E. coli*-celler. Till de kompetenta *E. coli*-cellerna tillsattes var för sig 4  $\mu$ l konstruerad plasmid till två olika tuber. En negativ kontroll utfördes genom att transformera klippta p413TEF till *E. coli*. Transformationerna utfördes med heat shock. De transformerade cellerna spreds sedan ut på en enskild agarplatta innehållandes ampicillin och inkuberades vid 37 °C över natten.

Därefter gjordes en inokulering av de erhållna kolonierna. Tolv tuber med LB-medium och ampicillin förbereddes och kolonier från p413 med insert (p413ins) samt p416 med insert (p416ins) placerades på en platta. Därefter tillsattes en koloni till varje tub, p413ins till tub 1-6 och p416ins till tub 7-12. Tuberna sattes därefter i en skakinkubator vid 37 °C över natten.



Dagen efter användes ett GeneJET™ Plasmid Miniprep Kit för extraktion enligt manualen på tub 1-12, med undantaget att vatten användes istället för elueringsbufferten. Koncentrationerna och renheten för alla plasmider uppskattades därefter med Nanodrop 2000c.

För p413ins samt p416ins gjordes en restriktionsanalys med enzymet *PvuII*. 1 µl av de renade p413ins plasmiderna tillsattes till en korresponderande tub tillsammans med 2 µl buffer, 0,5 µl av enzymet *PvuII*, samt vatten till en total volym av 20 µl. På samma sätt förbereddes p416ins, men med undantaget att 2 µl av plasmiden tillsattes. De tolv tuberna inkuberades därefter i tio minuter vid 37 °C för att restriktionsenzymet skulle verka. Därefter kördes en gelelektrofores (80 V, 90 min) för alla de tolv plasmiderna, se figur 6 på sida 16. Av de plasmider som uppvisade rätt antal restriktionsfragment skickades de med högst koncentration, en av p413ins och en av p416ins, in för sekvensering hos företaget Eurofins. För primrar som användes till sekvensering, se tabell 8 och 9 i appendix H.2.

### 4.3 Transformation till jäst samt selektion

Plasmiderna p413ins och p416ins skulle tillsammans transformeras till jäststammen JV04. Utifrån resultatet från restriktionsanalysen och sekvenseringen valdes vilka två av de tolv plasmidkonstruktionerna som skulle användas för transformationen. Skapandet av kompetenta celler och själva transformationen gjordes enligt appendix D. För att odla fram jästceller som kunde uttrycka det muterande systemet transformerades kombinationen "p413ins och p416ins" till cellerna. För att studera transformationens effektivitet gjordes tre kontroller där olika kombinationer mellan de rekombinanta plasmiderna och de ursprungliga plasmiderna transformerades in. Kombinationerna för de tre kontrollerna var "p413ins och p416TEF", "p413TEF och p416ins", samt "p413TEF och p416TEF". Även en kontroll gjordes där transformationsprotokollet följdes men inget DNA transformerades in. Cellerna odlades på plattor med medium med avsaknad av uracil och histidin för att selektera fram de celler som tagit upp plasmiderna. Jästen kultiverades sedan i två dagar 30 °C. För analys av transformationen till jäst se avsnitt 5.5 samt tabell 1.

Fyra kolonier av varje kombination av plasmider spreds ut i var sin sektor på nya plattor innehållandes medium med avsaknad av uracil och histidin. För kolonierna med "p413ins och p416ins" ströks det totalt ut åtta stycken kolonier på plattor. De nya plattorna inkuberades sedan i två dagar vid 30 °C. Efter detta togs enskilda kolonier från varje sektor och överfördes till nya plattor. Dessa två överföringar gjordes för att minimera risken att odla grupper av kolonier istället för enskilda kolonier.

Efter det utfördes en koloni-PCR för att se om *WS2* samt *dCas9* var närvarande i plasmiderna med insert (och frånvarande i de utan insert). Detta gjordes på hälften av kolonierna som erhöles. Det valdes alltså fyra kolonier innehållandes båda plasmiderna med insert, två kolonier med "p413ins och p416TEF", två kolonier av "p416ins och p413TEF" samt två kolonier med "p413TEF och p416TEF". Eftersom både *WS2* och *dCas9* studerades genom koloni-PCR för de utvalda kolonierna behövdes 20 olika prov, varav 10 av dem innehöll flankerande primrar för *WS2* och de 10 andra flankerande primrar för *dCas9*. Två positiva kontroller utfördes även med plasmiderna p416ins, innehållande *WS2*, och p413ins innehållande *dCas9*. För de positiva kontrollerna användes samma PCR-reaktionslösning som för kolonierna, med undantaget att 0,25 µl templat användes istället för 0,5 µl templat. För förberedning av de 20 tuberna samt framställning av PCR-reaktionslösning se appendix F.

De elva tuberna innehållandes flankerande primrar till *dCas9* kördes i PCR enligt schema, se tabell 4 i appendix E. De resterande tuberna innehållandes flankerande primrar för *WS2* sattes i en annan PCR-maskin enligt tabell 3 i appendix E. Efter PCR kördes en gelelektrofores (80 V, 50 min), se figur 8 och 9.

#### 4.4 Selektion och sekvensering

Baserat på resultatet från den koloni-PCR som gjordes, valdes kloner ut till förkulturer. Två kloner av följande kombination valdes; "p413ins och p416ins" och "p413ins och p416TEF". Utöver dessa så valdes även två kloner ut från CEN.PK 113-7D, hädanefter benämnt "wild type", som fanns att tillgå från tidigare gjorda experiment. Förkulturerna, en för klon 1 samt en för klon 2, inkuberades under omrörning i 30 °C över natten i medium innehållande glukos. Huvudkulturer förbereddes från förkulturerna baserat på ett önskat OD på 0,1 i huvudkulturer. De två förkulturerna med "p413ins och p416TEF" samt de två förkulturerna av wild type överfördes till nytt medium innehållande glukos. De två förkulturerna med "p413ins och p416ins" delades vardera upp i tre huvudkulturer som innehöll olika media. Det första mediet innehöll glukos, det andra innehöll galaktos och det tredje innehöll både galaktos och oljesyra. De två sistnämnda media innehållande galaktos användes för att inducera systemet via promotorn *pGAL1*. Huvudkulturerna inokulerades och inkuberades sedan i 30 °C under omrörning. Ett prov togs efter 5 timmar och ett togs efter 24 timmar.

För att testa tillväxt av de olika kulturerna på oljesyra förbereddes plattor innehållande sex olika koncentrationer av oljesyra samt plattor med avsaknad av oljesyra. Koncentrationerna som användes var 0,3125 mM, 0,625 mM, 1,25 mM, 2,5 mM, 5 mM, samt 10 mM. Förutom oljesyra innehöll mediet

i plattorna 0,5 % tergitol, för att lösliggöra oljesyra samt mediumlösning. För medium på 200 ml blandas 4 g glukos, 4 g agar, 0,15 g CSM med avsaknad av uracil och histidin samt 1,38 g kvävebaser utan aminosyror. För att åskådliggöra tillväxten på oljesyra utfördes droppanalys för alla koncentrationer oljesyra med alla tio huvudkulturer för de två induktionstiderna (5 respektive 24 timmar). De fyra huvudkulturer som inducerats med galaktos ströks även ut på plattor, med ca 10 000 celler per platta, för de två olika induktionstiderna för alla olika koncentrationer av oljesyra. Plattorna inkuberades i 30 °C över natten.

Kloner från dessa plattor fördes över till nya plattor med 10 mM koncentration av oljesyra samt avsaknad av uracil och histidin, detta för att få fram ensamma kloner som kunde skickas in för sekvensering. Tolv kolonier plockades från klon 1 respektive tolv från klon 2, varav sex stycken var odlade på oljesyra och sex stycken utan oljesyra. Detta upprepades för plattorna med de två olika induktionstiderna från dagen innan. De åtta nya plattorna placeras i 30 °C över natten.

Tolv stycken kolonier från plattorna valdes ut för att utföra koloni-PCR, detta för att påvisa närvaro av samt amplifiera *WS2* för att kunna sekvensera genen och förhoppningsvis visa att mutationer skett. En positiv samt negativ kontroll gjordes parallellt för att kunna bedöma pålitligheten av PCR-resultatet. Den positiva kontrollen bestod av P416ins och den negativa kontrollen av vatten. En gelelektrofores utfördes och åtta av de starkaste banden för *WS2* skars ut, se figur 13 och 14 på sida 23. DNA:t renades fram med hjälp av GeneJET™ Gel Extraction Kit. Nanodrop 2000c användes för att mäta koncentrationen av DNA i proverna. Proverna späddes så att slutkoncentrationen var 10 ng/µl. Sekvenseringsprimrar designades, se tabell 10 i appendix H.2, och proverna skickades in till företaget Eurofins för sekvensering.

De åtta kolonierna vars band var starkast på gelen togs från plattan och inkuberades i glukosmedium över natten. Som en negativ kontroll inkuberades även en koloni innehållande "p413ins och p416ins" respektive en koloni med "p413ins och p416TEF". Varje prov späddes till fyra nya koncentrationer; OD 0,1, 0,01, 0,001, 0,0001. Dessa koncentrationer från varje koloni droppades på åtta plattor med oljesyrakoncentrationer från 0-15 mM, se figur 15 på sida 25.

För de åtta kolonierna som skickades in på sekvensering och påvisat närvaro av *WS2* gjordes ytterligare en koloni-PCR. Denna gång gjordes det med flankerande primrar för *dCas9*, se appendix F och tabell 4 i appendix E, med undantaget att polymerisering kördes i 130 sekunder istället för 65 sekunder. En positiv kontroll med p413ins och en negativ kontroll med vatten laddades tillsammans med PCR-produkterna på en gel (80 V, 45 min), se figur 16 på sida 26. Detta gjordes för att kontrollera att de kolonier som skickats in på

sekvensering innehöll båda plasmiderna.

## 5 Resultat och diskussion

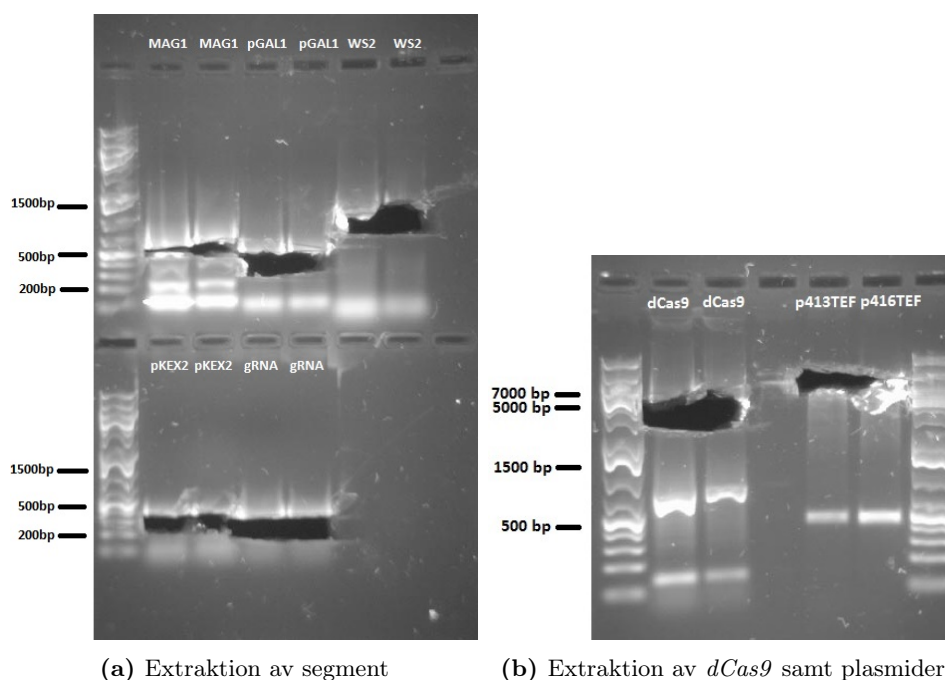
Följande avsnitt beskriver de resultat, samt diskussion kring dessa, som erhöles under studien. Studien innefattade ett flertal olika kontroller och verifieringar, beskrivna i avsnitt 4, som gav indikationer på hur ett visst steg lyckats och som därmed låg till grund för de val som behövde göras innan nästa steg kunde utföras.

### 5.1 Restriktionsanalys för p413TEF samt p416TEF

Restriktionsenzymen *SacI* och *XhoI* användes för att isolera plasmidernas ryggrader av p413TEF och p416TEF. Deras klippsekvenser var bäst lämpade för att klippa ut plasmidernas polylinkers som innehåller klippsekvenser för andra restriktionsenzym [26], samt TEF1-promotorn som inte var nödvändig i denna konstruktion, se figur 18 och 19 i appendix B. Gelelektroforesen som utfördes bekräftade att rätt längd på de klippta plasmiderna hade erhållits. I figur 5b ses resultatet av gelelektroforesen till höger. Det översta bandet motsvarar ryggraden av plasmiderna som har en längd på 5123 bp för p413TEF och 5051 bp för p416TEF vilket bekräftas av stegen till höger. Det understa bandet är polylinkern samt TEF1-promotorn som ska ha en längd på 478 bp vilket även här bekräftas av stegen. Eftersom längderna var korrekta kunde banden för plasmidernas ryggrad användas vidare i projektet.

### 5.2 PCR-amplifiering av fragment

De fragment som behövdes för att konstruera systemet amplifierades enligt PCR-scheman, se appendix E, tabell 3 och tabell 4. För att verifiera att både fragment och plasmider var av korrekta längder utfördes en gelelektrofores. Från figur 5a samt delar av figur 5b urskiljs att fragmenten klippts ut. Då fragmenten var nära de förväntade längderna (*MAG1* 891 bp, *pGAL1* 665 bp, *WS2* 1422 bp, *pKEX2* 488 bp, gRNA 99 bp, *dCAS9* 4101 bp, p413 ryggrad 5123 bp, p416 ryggrad 5050 bp) kunde fragmenten skäras ut, extraheras och användas vidare i projektet. Se appendix I.1, I.2 samt B för de förväntade längderna på fragmenten samt plasmidryggraderna.

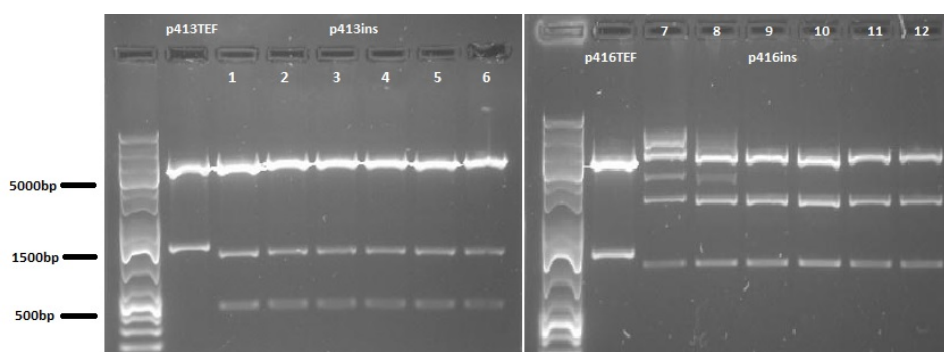


**Figure 5:** I (a) ses resultatet av PCR-amplifieringen och de fragment som skars ut efter gelelektrofores av *MAG1*, *pGAL1*, *WS2*, *pKEX2* samt *gRNA*. I (b) ses resultatet av PCR-amplifieringen samt de fragment som skars ut av *dCas9*. Figur (b) visar även resultatet av isoleringen av plasmidryggraden av *p413TEF* och *p416TEF* med restriktionsenzymerna *SacI* samt *XhoI*. För korrekta längder på fragment, se appendix I.1 och I.2. För korrekta längder på plasmider, se appendix B

### 5.3 Restriktionsanalys för p413ins samt p416ins

Efter konstruktionen av plasmiderna med Gibsonsammanfogning transformerades plasmiderna till *E. coli*. En kontroll av klippta plasmider utfördes för att mäta eventuellt bakgrundsbrus av kolonier som kan ha tagit upp oklippta, hela plasmider utan insert. Bakgrundsbruset av oklippta plasmider var såpass lågt att det inte antas påverka resultatet märkbart.

Sexton stycken transformerade *E. coli*-kolonier valdes från båda plasmiderna med insert. Dessa plasmider isolerades varpå en restriktionsanalys av plasmiderna utfördes med restriktionsenzymet *PvuII*, se figur 6. Restriktionsenzymet *PvuII* valdes då enzymet har specifika klippsekvenser som ger olika stora band beroende på om plasmiden tagit upp insertet eller ej. I appendix G figur 22 kan man se att p413ins kommer ge upphov till fyra fragment (4895, 4522, 977 och 456 bp) medan p416ins kommer ge upphov till tre (4449, 2018 och 890 bp) då de klippts med *PvuII*. Figur 23 visar att p413TEF (4523 och 1078 bp) och p416TEF (4451 och 1078 bp) ger upphov till två fragment då de klipps med *PvuII*. I figur 6 kan man urskilja att p416ins har klippts



**Figur 6:** Resultatet av restriktionsanalysen med *PvuII* för plasmiderna med insert där 1-6 är p413ins och 7-12 är p416ins. I brunnen närmast de båda stegarna är p413TEF respektive p416TEF.

till tre fragment och att p416TEF endast har klippts till två. p413ins har klippts till fyra fragment, där det översta och starkaste bandet på gelen antagligen är två band, medan p413TEF har klippts till två fragment. Då de flesta fragmenten var av förväntade längder verifierade det att Gibsonsammanfogningen av fragmenten skett korrekt, att inserten har infogats korrekt samt att plasmiderna transformerats till *E. coli*. De extra banden i brunn 7 och 8 kan bero på ofullständig klyvning med restriktionsenzymet, och därför valdes dessa prover ej att gå vidare med i projektet.

#### 5.4 Sekvensering av p413ins samt p416ins

Baserat på vilka prov som hade högst plasmidkoncentration valdes prov 1 för p413ins och prov 10 för p416ins för att sekvenseras, se figur 6. Detta gjordes för att kontrollera om Gibsonsammanfogningen lyckats. Det önskvärda resultatet av sekvenseringen var att sekvenserna skulle överensstämma med de teoretiskt designade plasmiderna, se appendix I.

Resultatet för p416ins var tillfredställande medan det för p413ins var bristande på grund av nukleotidvariationer och läsramsskiften, hädanefter benämmt "frameshift". Gällande p413ins upptäcktes ett flertal avvikelser i promotorsekvensen för *GAL1*, vilket kunde förklaras som skillnader mellan jäststammen JV04 som användes i projektet och den publicerade sekvensen för p*GAL1* hos CEN.PK [27]. En frågeställning gällande huruvida p*GAL1* i JV04 fungerade likvärdigt som p*GAL1* i den publicerade CEN.PK-sekvensen uppstod. Då tidsbrist rådde var ej vidare undersökningar ett alternativ, så det förutsattes att promotorn skulle fungera väl, alltså att promotorn endast skulle uttryckas vid närvaro av galaktos.

I *dCas9-NLS*-fragmentet identifierades en avvikelse i början av genen vilket hade lett till ett utbyte av en aminosyra. Detta var icke önskvärt då detta

kan leda till ändrad struktur och funktion hos proteinet. Sekvensövergången mellan *dCas9* och *MAG1* innehöll två avvikelser. En av dessa ledde till ett frameshift vilket gjorde plasmiden oanvändbar eftersom detta skulle innebära en felaktig avläsning av sekvensen. Värt att nämna är att båda dessa mutationer identifierades i primerregionen, se figur 7, och var därför ett förmodat fel som uppstått vid primersyntetisering hos företaget Eurofins. Med alla de olika avvikelser, mutationer och frameshift för p413ins-provet fattades beslutet att sekvensera de övriga proverna för p413ins, prov 2-6, se figur 6. Prov 6 för p413ins visade inga mutationer och valdes därför att gå vidare med, medan prov 10 valdes för p416ins.

	900	910	920	930
Målplasmid	CCACCTCCACCTCCATCACCACCTAATTGAG			
Sekvensering	.....-.....g.....			

**Figur 7:** p413ins som teoretisk konstruktion (ovan) jämförs med sekvenseringen av den konstruerade plasmiden vid den sekvenseringsprimer som ligger vid övergången mellan *MAG1* och *dCas9*. Prickar indikerar matchande sekvens, medan streck indikerar utebliven bas. Numrering indikerar antal baser från den första sekvenseringsprimern i sekvensen för den teoretiska konstruktionen.

## 5.5 Effektivitetskontroll av transformation till jäst

För att studera hur effektiv transformationen till jäst var, gjordes tre kontroller där olika kombinationer av plasmiderna med och utan insert transformerades in. Det utfördes även en kontroll utan DNA där transformationsprotokollet följdes, se appendix D. Dessa kolonier odlades på plattor och antalet kolonier räknades, se tabell 1. Plattorna bestod av medium med avsaknad av uracil och histidin, vilket resulterar i att de kolonier som har tagit upp plasmider, med eller utan insert, kan överleva på plattan. Detta bekräftades utav kontrollen utan DNA där inga kolonier växte på plattan. Flest antal kolonier observerades på plattan med “p413TEF och p416TEF”. Detta kan bero på att dessa plasmider är mindre jämfört med de med insert, vilket kan resultera i en mer effektiv transformation. I en artikel av Hanahan visades att effektiviteten av transformationen minskar linjärt med ökad plasmidstorlek; dock skedde här transformationen till *E. coli* [28]. På plattan med “p413TEF och p416ins” observerades fler kolonier jämfört med plattan med “p413ins och p416TEF”. Då p416ins är 7357 bp jämfört med p413ins som är 10850 bp kan samma resonemang för en effektiv transformation baserat på plasmidens storlek föras även här. På plattan med “p416ins och p413ins” fanns minst antal kolonier, vilket stärker hypotesen om att transformationens effektivitet minskar med ökad plasmidstorlek. En alternativ hypotes

till resultaten i tabell 1 är att inserten ger en skadlig påverkan på cellerna och därmed minskar antalet överlevande kolonier som har plasmider med insert.

**Tabell 1:** Visar antalet kolonier efter transformation till jäst på plattorna med; “p413ins och p416ins”, “p413ins och p416TEF”, “p413TEF och p416ins”, “p413TEF och p416TEF” samt kontroll utan DNA. Plattorna innehöll medium med avsaknad av uracil och histidin.

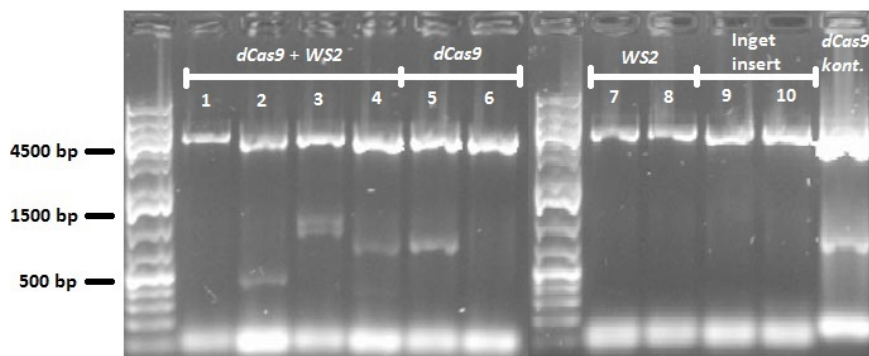
Platta	Antal kolonier
p413ins + p416ins	9
p413ins + p416TEF	35
p413TEF + p416ins	ca 170
p413TEF + p416TEF	ca 250
utan DNA	0

## 5.6 Koloni-PCR som kontroll för närvaro av p413ins samt p416ins

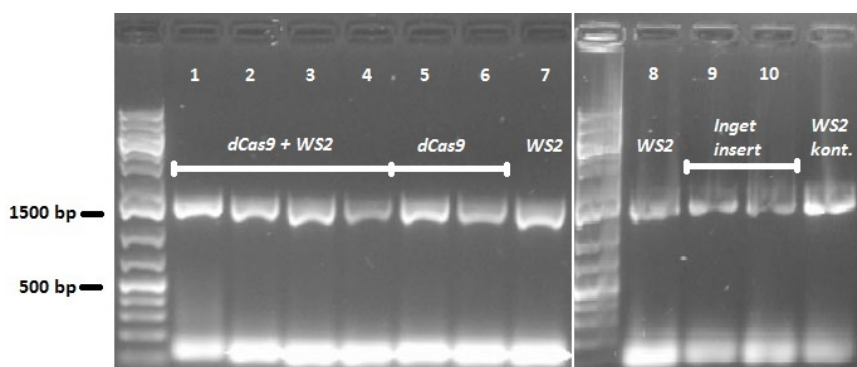
Efter att kolonier samt kontroller hade odlats upp på plattor utfördes koloni-PCR. Detta gjordes för att kontrollera att *WS2* samt *dCas9* fanns med i kolonier där de förväntades att finnas med. Då *WS2* är specifik för p416ins och *dCas9* för p413ins har närvaro av dessa använts som indikation att hela inserten finns med, då tidigare sekvensering har bekräftat att inserten är korrekta. I det första försöket med koloni-PCR användes Taq-polymeras då det var billigare än Phusion-polymeras. Taq-polymeras har dock vissa begränsningar jämfört med Phusion-polymeras, som att det saknar korrekturläsningförmåga, vilket leder till en högre mutationsfrekvens [6]. Då resultatet från första försöket med koloni-PCR gav bristande resultat för *dCas9* samt inget resultat för *WS2* fattades beslutet att göra en ny koloni-PCR med Phusion-polymeras. Resultatet av denna koloni-PCR ses i figur 8 och 9. I figur 8 ses *dCas9* som det översta bandet i den positiva kontrollen längst till höger med en längd på ca 4100 bp. Alla kolonier som valdes till koloni-PCR uppvisar ett band av samma längd. Detta är resultatet som endast förväntades för prov 1-6. I prov 7-10 förväntas inga band för *dCas9*. Att prov 7-10 uppvisar band för *dCas9* samt att prov 2-5 uppvisar extra band kan bero på kontaminering med DNA i PCR, se figur 8 och figur 9. I figur 9 ses *WS2* som det översta bandet i den positiva kontrollen längst till höger med en längd på ca 1400 bp. Även här urskiljs ett band för *WS2* för alla prover, vilket ej är vad som förväntades för proven som i figur 9 benämns “dCas9” samt “inget insert”, vilket även i detta fallet kan bero på kontaminering. På grund av begränsad tid valde vi, trots osäkerheten, att gå vidare med prov nummer 3 och 4 för kolonier med båda inserten. Prov 5 (-*WS2* + *dCas9*) och 10 (-*WS2* - *dCas9*), se figur 8 och 9, valdes för att senare fungera som kontroller. Då resultaten var tvivelaktiga beställdes primrar för inserten, se appendix H.3, för att göra



ytterligare en kontroll och att få bekräftat om inserten finns närvarande eller ej. Denna kontroll gav dessvärre inga tolkningsbara resultat.



**Figur 8:** Koloni-PCR gjord med flankerande primers till *dCas9*. I brunn 1-4 finns DNA från kolonier med “p413ins och p416ins”, i 5-6 DNA från kolonier med “p413ins och p416TEF”, i 7-8 kolonier med “p413TEF och p416ins” och i 9-10 kolonier med “p413TEF och p416TEF”. Längst till höger i figuren är en positiv kontroll för *dCas9*, bestående av framrenad p413ins. *dCas9* är totalt 4101 bp långt.

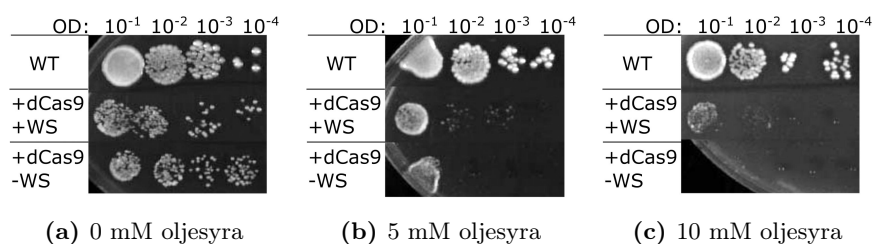


**Figur 9:** Koloni-PCR gjord med flankerande primers till *WS2*. I brunn 1-4 finns DNA från kolonier med “p413ins och p416ins”, i 5-6 DNA från kolonier med “p413ins och p416TEF”, i 7-8 kolonier med “p413TEF och p416ins” och i 9-10 kolonier med “p413TEF och p416TEF”. Längst till höger i figuren är en positiv kontroll för *WS2* bestående av framrenad p416ins. *WS2* är totalt 1422 bp långt.

De primrar som användes för koloni-PCR var samma primrar som tidigare använts för Gibsonsammanfogning. I försöket med koloni-PCR för *WS2* användes de primrar som tillhörde *WS2* och för *dCas9* användes likväl de primrar som tillhörde *dCas9*, se figur 4. Viktigt att notera är att dessa primrar ej är optimala för koloni-PCR eftersom de är långa på grund av överlappen till länkade fragment. Detta kan leda till att primrarna bildar sekundära strukturer.

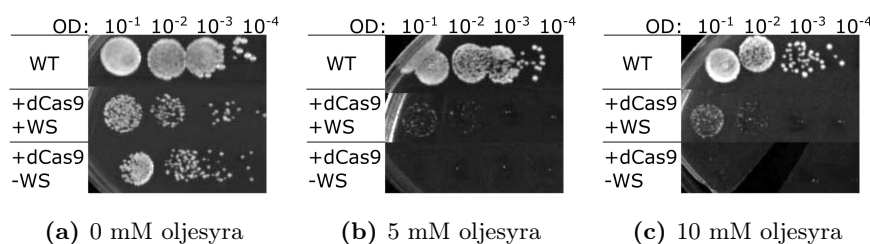
### 5.7 Test av kritisk koncentration oljesyra

Plasmiderna transformerades till jäst i två olika kombinationer; "p413ins och p416ins" samt "p413ins och p416TEF". Med dessa kombinationer av plasmider gjordes droppanalyser vid två tillfällena, med ett dygns mellanrum, på plattor med de valda koncentrationerna för oljesyra, 0-10 mM, se avsnitt 4.4. Detta gjordes för att kunna urskilja vilken koncentration som är kritisk för tillväxt av de celler som har "p413ins och p416ins". Jämförelse gjordes med de andra kombinationerna av plasmider samt wild type. Det förväntade resultatet var att wild type skulle ha den största tillväxten. Detta beror på att den, till skillnad från resten, kan lagra oljesyra i form av acyl-CoA i lagringslipider och oljesyran blir då ej giftig för cellen [12]. Denna teori kan bekräftas i figur 10 där wild type uppvisar högre tillväxt än de rekombinanta klonerna vid närvaro av oljesyra. Wild type växte även bättre än de andra klonerna utan oljesyra, men skillnaden var mindre. Då resterande celler som droppades på plattan är av jäststammen JV04, saknas egenskapen att kunna bilda lagringslipider samt generera energi med  $\beta$ -oxidation [11], se figur 2. Då oljesyran inte kan inkorporeras i lagringslipider samt att cellerna inte kan generera tillräckligt mycket energi kommer deras tillväxt att hämmas. Kombinationerna med avsaknad av *WS2*, "p413ins och p416TEF" samt "p413TEF och p416TEF", är de celler som förväntades ha den lägsta tillväxten. Denna hypotes baseras på att acyl-CoA och etanol inte kan omvandlas till fettsyraetylestrar utan *WS2* och acyl-CoA kan ej heller inkorporeras i lagringslipider. Dessa kolonier utan *WS2* växer sannolikt sämre än då *WS2* finns närvarande, då *WS2* utan mutationer möjligen kan katalysera processen för bildning av fettsyraetylestrar. I figur 10 och 11 ses resultatet från de två droppanalyser som gjordes. Tillväxten av kolonier med *WS2* är högre än kolonier utan *WS2* vid närvaro av oljesyra. Detta innebär sannolikt att närvaro av icke-muterat *WS2* ökar cellernas förmåga att växa vid närvaro av oljesyra.



**Figur 10:** Droppanalys #1 av kolonier på olika koncentrationer av oljesyra. Wild type jämförs med oinducerade kolonier innehållandes i ena fallet både *dCas9* och *WS2*, i det andra fallet enbart *dCas9*.

Då det fanns tillväxt på plattor innehållandes 10 mM oljesyra fattades beslutet att till nästa droppanalys även odla kolonier på 15 mM för att säkerställa



**Figur 11:** Droppanalys #2 av kolonier på olika koncentrationer av oljesyra. Wild type jämförs med oinducerade kolonier innehållandes i ena fallet både *dCas9* och *WS2*, i det andra fallet enbart *dCas9*.

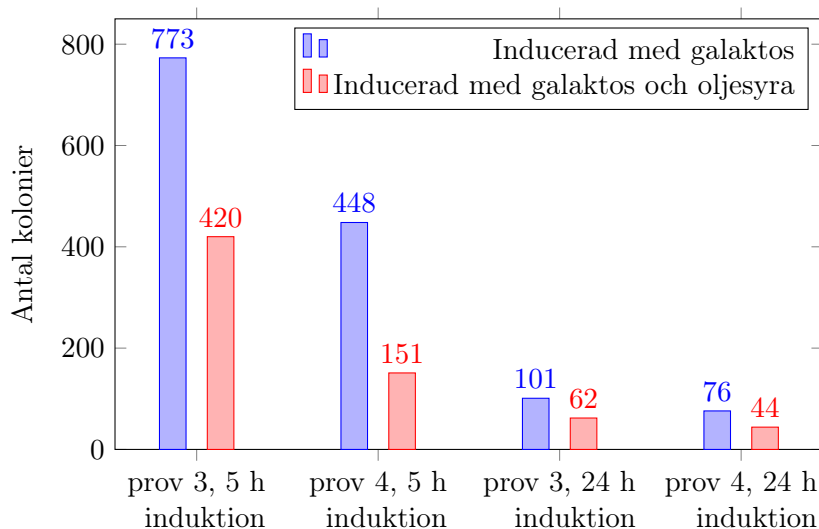
att kritisk koncentration av oljesyra uppnås. Detta var viktigt att säkerställa för att kolonier utan mutationer på *WS2* inte skulle kunna växa.

## 5.8 Odling av kolonier där systemet inducerats

För att det muterande systemet skulle kunna uttryckas behövde det induceras med galaktos. För att undersöka effekten av olika induktionstider testades inducering i både 5 och 24 timmar. Dessutom undersöktes huruvida selektionstryck från oljesyra under induktionen ger fler gynnsamma mutationer.

De inducerade klonerna, innehållandes p413ins och p416ins, odlades på plattor med 0-10 mM oljesyra, se avsnitt 4.4. De kloner som odlades var prov 3 och 4 från figur 8 och 9. Dessa valdes att gå vidare med baserat på resultat från koloni-PCR, se avsnitt 5.6. Här förväntades endast de kolonier som klarade selektionstrycket från oljesyran kunna växa, bland dessa de kolonier som möjligen fått en gynnsam mutation på *WS2*. För att hitta en kritisk koncentration av oljesyra, sett till tillväxt, skulle droppanalysen optimalt utförts innan uppodling av jäst med ett inducerat system. Då hade det inducerade systemet kunnat odlas endast vid den kritiska koncentrationen. På grund av tidsbrist gjordes detta parallellt, och det inducerade systemet odlades därför vid samma koncentrationer som droppanalysen. Genom att räkna kolonier som inducerats olika länge kunde en eventuell koppling mellan tillväxt och induceringstid utredas. De inducerade kolonier som räknades var de som fått växa på plattorna med högst koncentration, 10 mM, se figur 12.

Fler kolonier observerades då induktionstiden var kortare. En hypotes kring detta är att ju kortare tid kolonierna induceras, desto färre mutationer utförs av det muterande systemet. Fler mutationer kan ha resulterat i att *WS2* blivit mer ineffektivt. Ju längre tid som systemet inducerats, desto större risk är det också att mutationer har skett på andra delar av genomet som skulle kunna ge en negativ påverkan på tillväxt. Att färre kolonier bildades då systemet inducerades i närvaro av oljesyra beror sannolikt på att jästens

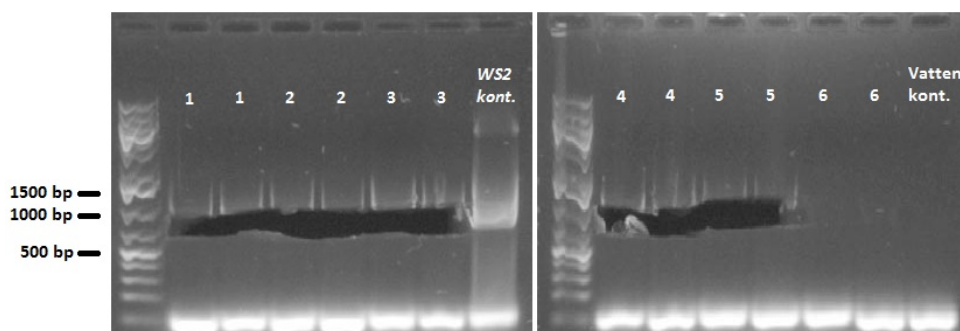


**Figur 12:** Antal kolonier på plattorna efter att systemet inducerats i 5 eller 24 timmar och med eller utan oljesyra.

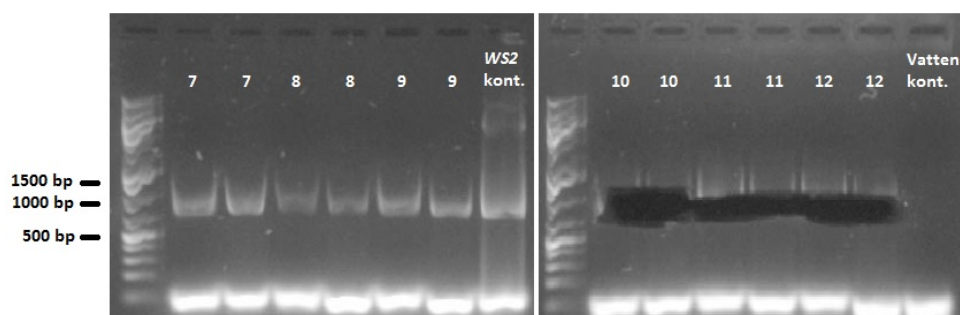
tillväxt har hämmats av oljesyran [12]. Kolonier från dessa plattor överfördes till nya plattor med 10 mM oljesyra för att erhålla enskilda kloner som möjligen innehöll mutationer på *WS2*.

### 5.9 Koloni-PCR för amplifiering av *WS2* till slutgiltig sekvensering

Då enskilda kloner erhöles på 10 mM oljesyraplattor valdes 12 av dessa ut för eventuell sekvensering med primrar för *WS2*. För att säkerställa att de kloner som valdes till sekvensering verkligen innehöll *WS2* utfördes en koloni-PCR för *WS2* för dessa kloner. De primrar som användes var de primrar som tidigare använts för Gibsonsammanfogning, se tabell 7 i appendix H.1, som korresponderade till *WS2*. Resultatet av denna koloni-PCR för *WS2* visas i figur 13 och 14. Alla de enskilda klonernas band jämfördes mot den positiva kontrollen, som visade ett tydligt band för *WS2* på ca 1500 bp. Band för alla kolonier utom en uppvisades och de prov med starkast band valdes och skars ut. Den negativa kontrollen med vatten uppvisade inga band, vilket gav oss belegg för att vi hade rena prov utan kontamination. Dessa åtta kloner skickades på sekvensering för *WS2*, för primrar se tabell 10 i appendix H.2.



**Figur 13:** Koloni-PCR på gel 1 gjord med flankerande primers till *WS2* för utvalda kolonier inför sekvensering. Till höger om prov 3 ses en positiv kontroll för *WS2*, bestående av framrenad p416ins, och bredvid prov 6 en negativ kontroll med vatten. *WS2* har en total längd på 1422 bp.



**Figur 14:** Koloni-PCR på gel 2 gjord med flankerande primers till *WS2* för utvalda kolonier inför sekvensering. Till höger om prov 9 ses en positiv kontroll för *WS2*, bestående av framrenad p413ins, och bredvid prov 12 en negativ kontroll med vatten. *WS2* har en total längd på 1422 bp.

### 5.10 Droppanalys med inducerade, enskilda kloner

Ännu en droppanalys utfördes för att undersöka hur väl de åtta kloner som valdes från de ovan nämnda plattorna, med 10 mM oljesyra, växte jämfört med icke inducerade varianter, se figur 15. Det förväntade resultatet var att kulturerna “p413ins och p416ins, icke inducerat” samt “p413ins och p416TEF” skulle ge upphov till mindre kolonier på plattorna för alla valda koncentrationer av oljesyra. Kolonierna med “p413ins och p416TEF” saknar *WS2* och kan därmed inte använda oljesyra för att syntetisera oljesyraety-lestrar. Därför bör dessa kolonier växa sämre än resten. Hos kolonierna med “p413ins och p416ins, icke inducerat” finns *WS2*, men då det muterande systemet ej har inducerats med galaktos borde inga mutationer, gynnsamma eller ogynnsamma, påverka enzymets aktivitet. Även utan gynnsamma mutationer är det sannolikt att närvaron av *WS2* leder till att en viss del av oljesyran kan omvandlas till oljesyraety-lestrar [10].

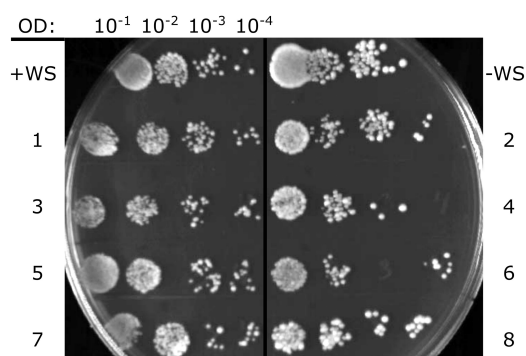
Som synes i figur 15 stämmer de erhållna resultaten mestadels väl överens med de förväntade. “p413ins och p416ins, icke inducerat” gav upphov till större kolonier än “p413ins och p416TEF” vid en oljasyrakoncentration av både 5 mM, se figur 15b, och 15 mM, se figur 15c. Dock växer kolonierna med de båda kombinationerna av plasmider sämre än de åtta inducerade kolonierna vid närvaro av oljesyra. Av de åtta kloner som innehöll “p413ins och p416ins, inducerat” visade många mer tillväxt än de icke inducerade, men samtidigt syntes ibland större tillväxt vid lägre OD-värde, vilket ej är rimligt. Detta kan bero på oförsiktighet då spädningarna utfördes. Klon 1 och 3 visade båda mer tillväxt utan felaktig spädning. Anledningen till att dessa har visat bättre tillväxt kan vara att de har fått fördelaktiga mutationer för att motstå oljesyrans toxicitet, till exempel på *WS2*. Alla åtta kloner sekvenserades med primrar för *WS2* och resultatet för denna droppanalys jämfördes mot sekvenseringsresultatet för att se vilka eventuella mutationer på *WS2* som skett i respektive klon.

### 5.11 Koloni-PCR för *dCas9* för enskilda kloner

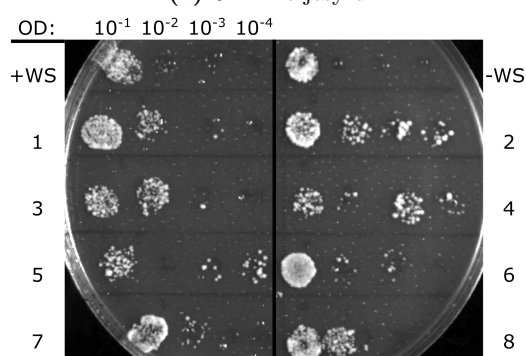
För att kunna säkerställa att *dCas9* fanns närvarande i alla åtta kloner, vilket inte tydligt kunnat visas med tidigare försök, utfördes en koloni-PCR för *dCas9* på dessa kloner. Det var viktigt att bekräfta att *dCas9* var närvarande, för om mutationer hade uppstått skulle det innebära att det muterande systemet sannolikt fungerar. Om *dCas9* däremot inte hade varit närvarande och mutationer ändå uppstått hade vi ej kunnat dra slutsatsen att det konstruerade systemet varit det som orsakat mutationerna. Resultaten för denna koloni-PCR visas i figur 16. Här framgår det tydligt att *dCas9* finns närvarande i sex utav åtta kloner. Det är trots detta fullt möjligt att även de två kloner som ej uppvisar band för *dCas9* innehåller detta fragment, med tanke på hur många tidigare försök med koloni-PCR som inte har visat resultat och behövt upprepas. I figur 16 åskådliggörs att den positiva kontrollen för *dCas9*, utförd med p413ins som templat, uppvisar ett något högre band än förväntat. Detta kan ha berott på att plasmidkoncentrationen för templatet var för hög. Det är också möjligt att PCR-reaktionen misslyckades, då bandets höjd motsvarar storleken på plasmiden p413ins i den positiva kontrollen. Även spår av kontaminering syns i form av tunna band i den negativa kontrollen med vatten, men dessa band är betydligt svagare än de för *dCas9* och därför ansågs inte *dCas9*-banden bero på kontaminationen.

### 5.12 Sekvensering av *WS2* efter induktion av systemet

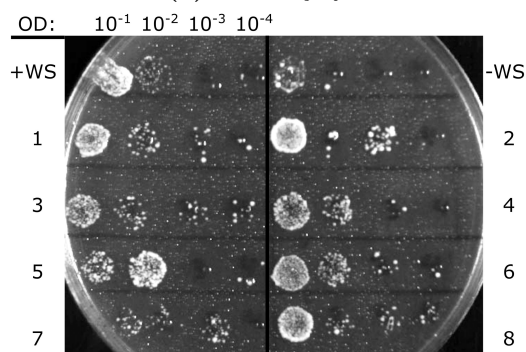
Sekvenseringen visade på att inga mutationer hade uppkommit på *WS2* för någon av de åtta klonerna. Båda sekvenseringsprimrarna för klon 8 och en av



(a) 0 mM oljesyra



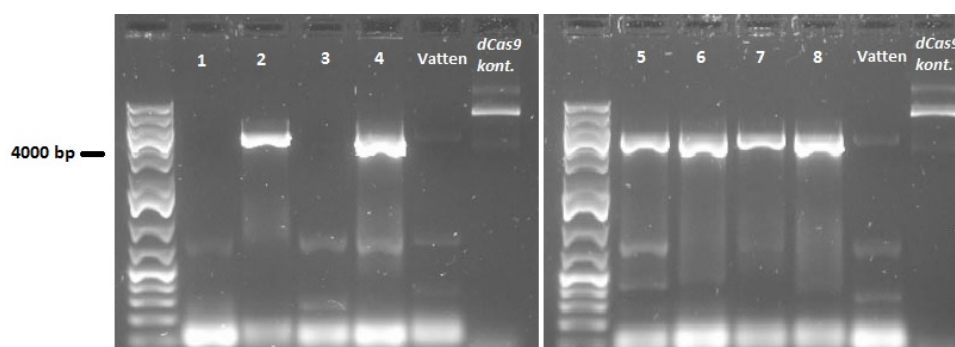
(b) 5 mM oljesyra



(c) 15 mM oljesyra

**Figur 15:** Droppanalys på 0, 5 respektive 15 mM oljesyra för de oinducerade “p413ins och p416ins” (+WS) och “p413ins och p416TEF” (-WS), samt de åtta inducerade kloner som valdes baserat på koloni-PCR för *WS2*. OD-beteckningen anger spädningen i dropparna i respektive kolonn.

primrarna för klon 7 gav resultat med låg tillförlitlighet. Det är inte omöjligt att det beror på mutationer har uppstått där primrarna binder in, men det är troligare att något fel har inträffat, till exempel att DNA-koncentrationen var för låg.



**Figur 16:** Koloni-PCR gjord med flankerande primers till *dCas9* för de åtta utvalda kolonierna för sekvensering av *WS2*. Bredvid prov 4 samt prov 8 ses en negativ kontroll med vatten samt en positiv kontroll för *dCas9*, bestående av framrenad p413ins. *dCas9* har en total längd på 4143 bp.

Att inga mutationer hittades betyder inte att inga finns. Då endast *WS2* sekvenserades skulle mutationer kunna finnas på andra sekvenser på plasmiden, till exempel på någon promotor. Sådana mutationer skulle kunna leda till ökat uttryck av *WS2* eller högre kopietal av plasmiden, vilket skulle kunna ge ökad tillväxt vid närvaro av oljesyra. Att gynnsamma mutationer skulle kunna finnas understöds av att många av de inducerade klonerna i droppanalysen med 5 mM och 15 mM, se figur 15b respektive 15c, visar mer tillväxt än de som inte hade inducerats. Mutationer skulle även kunna ha uppkommit på andra delar av genomet, framförallt om inbindning till *WS2* med gRNA inte fungerade. Dessa mutationer skulle också kunna vara förklaringen till resultatet i figur 15b och 15c.

Det är också möjligt att mutationer uppkom på *WS2*, men inte på de kloner som sekvenserades. Detta är särskilt troligt om 10 mM är en för låg koncentration av oljesyra för att stoppa tillväxt av jäst utan gynnsamma mutationer. Då skulle sannolikheten att välja muterade kloner från plattorna med inducerat system kunna vara låg. En annan möjlighet är att mutationsfrekvensen är för låg för att kunna observeras i en så liten population. Detta skulle i så fall också förstärka problemen med att använda för låg oljesyrakoncentration. Induktionstiden skulle också kunna vara för kort i förhållande till mutationsfrekvensen.

Enligt Finney-Manchester och Maheshri orsakar Mag1 både frameshift-mutationer och punktmutationer [17]. Om frameshift har uppstått är det osannolikt att tillväxt har fortsatt, och därför skulle sådana mutationer inte kunna observeras på plattorna.

Det är sannolikt att någon del av det muterande systemet inte fungerar eller kanske inte ens uttrycks. Till exempel kan skillnaderna på *pGAL1* jämfört med den publicerade CEN.PK-sekvensen, se avsnitt 5.4, innebära att pro-



motorn inte fungerar som den ska. För att kontrollera att alla steg under projektet har lyckats skulle ytterligare kontroller kunna ha utförts. Detta var dock ej möjligt inom tidsramen för detta projekt.

### 5.13 Utökade experiment och kontroller

Då plasmiderna med insert sekvenserades efter transformation till *E. coli* bekräftades närvaron av alla fragment i inserten. Baserat på detta resultat har det sedan under projektets gång tagits för givet att alla fragment i inserten är korrekta och närvarande i plasmiderna. Efter transformation av p416ins och p413ins till jäst så kan det dock ha uppstått komplikationer, som kan innebära att jästen har avlägsnat vissa fragment i inserten. De kontroller i form av koloni-PCR som har gjorts under projektet har baserats på att *WS2* är specifik för p416ins och *dCas9* är specifik för p413ins. Om dessa fragments närvaro har kunnat påvisas, så antogs resterande fragment på insertet finnas där. För att kontrollera huruvida detta antagande stämmer så hade en PCR med omvänd transkription kunnat utföras. En PCR med omvänd transkription [26] hade kunnat påvisa att alla fragment finns med i inserten samt att alla fragment transkriberas till RNA.

Ett andra antagande som gjordes under projektets gång var att det muterande systemet uttrycktes korrekt till proteiner. Detta antagande kontrollerades dock aldrig. För att undersöka huruvida sekvenser som kodar för systemet har uttryckts till protein hade metoden Western blot kunnat användas. Detta är en metod då proteiner separeras i en polyakrylamidgel, överförs till membran och exponeras för märkta antikroppar vilket kan påvisa ett proteins närvaro [29]. En alternativ metod för att undersöka om proteinet har uttryckts är att sammanfoga proteinet med ett fluorescerande protein. Om sekvensen för det fluorescerande proteinet inkorporeras direkt efter sekvensen för proteinet som undersöks, så kan man anta att om fluorescens kan detekteras så har proteinet uttryckts. Ett problem med detta kan dock vara att det muterande systemet, uttryckt som protein, får en felaktig veckning och eventuellt funktion om det kopplas till ett fluorescerande protein.

Den variabla delen för gRNA designades så att det muterande systemet skulle kunna binda specifikt till *WS2*. För att undersöka om det muterande systemet verkligen binder till *WS2*, kan metoden EMSA<sup>8</sup> användas. En blandning av protein och DNA och körs sedan i gelelektrofores. Protein som interagerar med DNA vandrar generellt långsammare i gelen än det motsvarande fria DNA:t [29]. På detta sätt skulle interaktionen mellan det muterande systemet och *WS2* kunna påvisas.

I projektet odlades jästen med det inducerade systemet på oljesyra. Detta

---

<sup>8</sup>Electrophoretic Mobility Shift Assay

antogs kunna påvisa vilka kolonier som hade fått fördelaktiga mutationer på *WS2* som möjliggör en syntes av fettsyraetylestrar, se figur 2. Detta antogs vara anledningen till att vissa kolonier hade en högre tillväxt än andra på plattorna med oljesyra. Dock visade sekvenseringen av *WS2*, för några av de kolonierna med högst tillväxt, att inga mutationer hade uppstått. Anledningen till att kolonierna ändå hade en hög tillväxt, kan bero på att vi under projektet har selekterat för kolonier med en hög resistens mot oljesyra, vilket inte måste innebära att mutationer på *WS2* har uppstått. Detta kan istället bero på att andra fördelaktiga mutationer i jästens genom har uppstått vilket resulterat i ökad tillväxt på oljesyra. För att kontrollera om sådana mutationer har uppstått, skulle hela genomet för dessa kolonier kunna sekvenseras.

För att öka chansen att det muterande systemet skapar mutationer på *WS2* kan flera bindningssäten för det muterande systemet skapas. I vårt projekt skapades endast ett bindningssäte för systemet, genom den variabla delen på gRNA. En hypotes till att mutationer inte uppstod på *WS2* var att systemets mutationsfrekvens var för låg. För att erhålla en ökad mutationsfrekvens skulle fler bindningsställen kunna introduceras. I en studie av Finney Manchester och Maheshri [17] studeras ett fusionsprotein innehållande Mag1 för att åstadkomma riktade mutationer *in vivo* i jäst. I studien används flera bindningssäten för fusionsproteinet vilket resulterar i mutationer på målskvensen. Detta tyder på att flera bindningssäten skulle vara av relevans att testa i framtiden även i detta projekt.

En annan förändring av det muterande systemet som hade varit av relevans att studera är att undersöka om andra muterande enzym kan ge förbättrade resultat. Detta skulle kunna innebära att ersätta det muterande enzymet Mag1 med ett analogt enzym från en annan organism. Alternativt skulle Mag1 kunna ersättas av ett enzym med andra muterande egenskaper, till exempel ett DNA-metylas.

## 6 Slutsatser

Syftet med studien var att skapa ett muterande system genom att utnyttja CRISPR/Cas-systemet, detta för att skapa slumpvisa men till en specifik sekvens riktade mutationer *in vivo*.

En lyckad konstruktion av det muterade systemet skapades med hjälp av vektorer som uttrycker *WS2* och ett dCas9-Mag1-fusionsprotein. Dessa vektorer transformerades till jäststammen *S. cerevisiae* JV04, som sedan hos några kolonier visade högre resistens mot oljesyra efter att det muterande systemet hade inducerats med galaktos. Det muterande systemet har dock inte kunnat visas mutera *WS2* efter sekvensering av genen. Det går inte att

helt säkerställa att det inte fungerar utan ytterligare experimentella försök. Ytterligare experimentella försök skulle bland annat vara att sekvensera fler kolonier, att sekvensera hela p416ins-plasmiden för att se om mutationer skett på annan plats, att ha längre induceringstid, att ha ett ökat selektionstryck samt studera huruvida gRNA binder in till *WS2*.

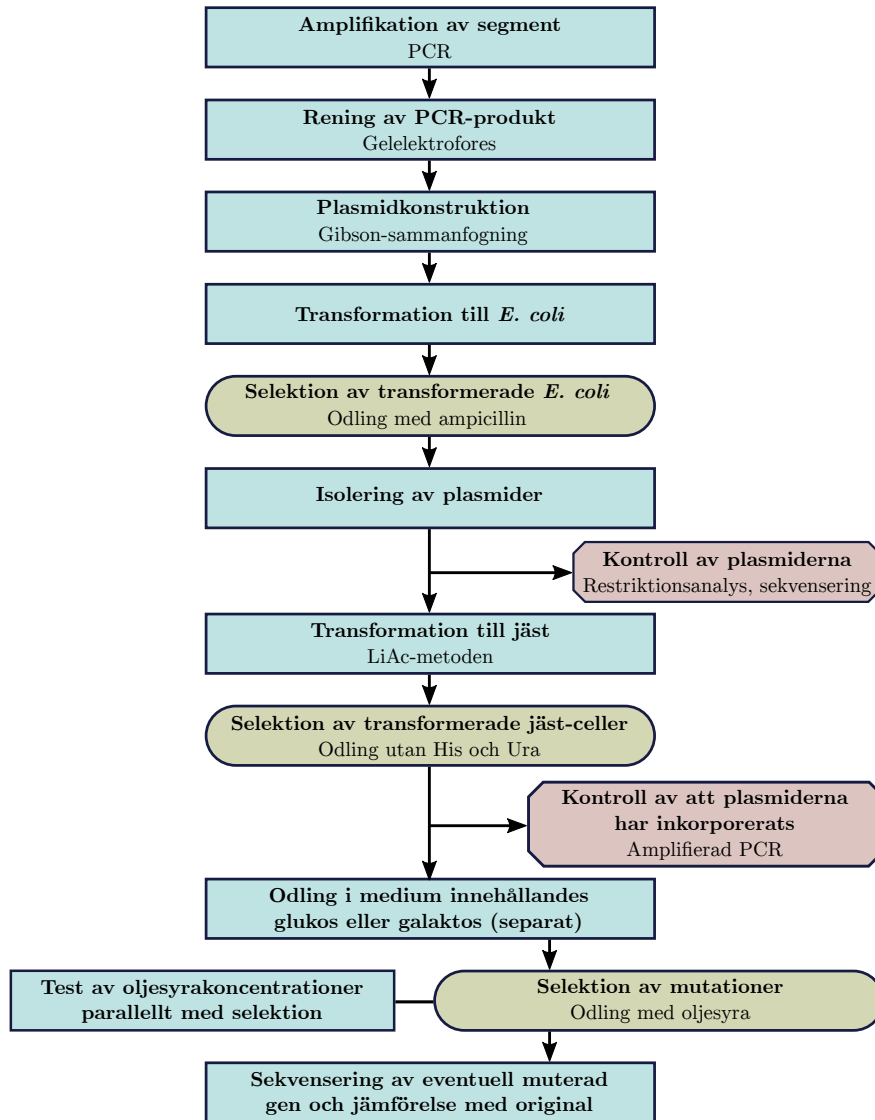
Andra experimentella försök som skulle vara av relevans är att sekvensera alla fragment i inserten för att påvisa deras närvaro, att studera att de ingående fragmenten korrekt uttryckts till proteiner, att skapa fler bindningssäten för systemet samt att byta systemets muterande enzym mot ett annat.

## Referenser

- [1] Mukhopadhyay A, Redding AM, Rutherford BJ, Keasling JD. Importance of systems biology in engineering microbes for biofuel production. *Curr Opin Biotechnol.* 2008;19(3):228 – 234.
- [2] Zhang F, Rodriguez S, Keasling JD. Metabolic engineering of microbial pathways for advanced biofuels production. *Curr Opin Biotechnol.* 2011 dec;22(6):775–783.
- [3] Lee JW, Na D, Park JM, Lee J, Choi S, Lee SY. Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat Chem Biol.* 2012 jun;8(6):536–546.
- [4] Nannemann DP, Birmingham WR, Scism RA, Bachmann BO. Assessing directed evolution methods for the generation of biosynthetic enzymes with potential in drug biosynthesis. *Future Med Chem.* 2011 maj;3(7):809–819.
- [5] Xu H, Petersen EI, Petersen SB, El-Gewely MR. Random mutagenesis libraries: optimization and simplification by PCR. *BioTechniques.* 1999 dec;27(6):1102–1104.
- [6] Cirino PC, Mayer KM, Umeno D. Generating Mutant Libraries Using Error-Prone PCR. I: Arnold FH, Georgiou G, redaktörer. *Directed Evolution Library Creation.* vol. 231 av *Methods in Molecular Biology™*. Humana Press; 2003. s. 3–9.
- [7] Fabret C, Poncet S, Danielsen S, Borchert TV, Ehrlich SD, Janniere L. Efficient gene targeted random mutagenesis in genetically stable *Escherichia coli* strains. *Nucleic Acids Res.* 2000 nov;28(21):E95.
- [8] Farzadfard F, Perli SD, Lu TK. Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. *ACS Synth Biol.* 2013 okt;2(10):604–613.
- [9] Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell.* 2013 feb;152(5):1173–1183.
- [10] Shi S, Valle-Rodríguez JO, Khoomrung S, Siewers V, Nielsen J. Functional expression and characterization of five wax ester synthases in *Saccharomyces cerevisiae* and their utility for biodiesel production. *Biotechnol Biofuels.* 2012;5:7.
- [11] Valle-Rodríguez JO, Shi S, Siewers V, Nielsen J. Metabolic engineering of *Saccharomyces cerevisiae* for production of fatty acid ethyl esters, an advanced biofuel, by eliminating non-essential fatty acid utilization pathways. *Appl Energy.* 2014;115:226–232.
- [12] Garbarino J, Padamsee M, Wilcox L, Oelkers PM, D’Ambrosio D, Ruggles KV, et al. Sterol and diacylglycerol acyltransferase deficiency triggers fatty acid-mediated cell death. *J Biol Chem.* 2009 nov;284(45):30994–31005.

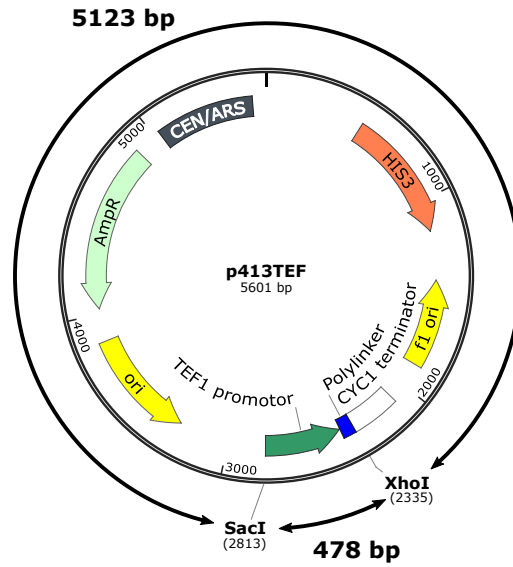
- [13] Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012 aug;337(6096):816–821.
- [14] Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol*. 2013 mar;31(3):233–239.
- [15] Villa JA, Cabezas M, de la Cruz F, Moncalian G. Use of limited proteolysis and mutagenesis to identify folding domains and sequence motifs critical for wax ester synthase/acyl coenzyme A:diacylglycerol acyltransferase activity. *Appl Environ Microbiol*. 2014 feb;80(3):1132–1141.
- [16] Barney BM, Mann RL, Ohlert JM. Identification of a residue affecting fatty alcohol selectivity in wax ester synthase. *Appl Environ Microbiol*. 2013 jan;79(1):396–399.
- [17] Finney-Manchester SP, Maheshri N. Harnessing mutagenic homologous recombination for targeted mutagenesis in vivo by TaGTEAM. *Nucleic Acids Res*. 2013 maj;41(9):e99.
- [18] Nacken V, Achstetter T, Degryse E. Probing the limits of expression levels by varying promoter strength and plasmid copy number in *Saccharomyces cerevisiae*. *Gene*. 1996 okt;175(1-2):253–260.
- [19] DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res*. 2013 apr;41(7):4336–4343.
- [20] Mumberg D, Muller R, Funk M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene*. 1995 apr;156(1):119–122.
- [21] Gibson DG. Chapter fifteen - Enzymatic Assembly of Overlapping DNA Fragments. I: Voigt C, redaktör. *Synthetic Biology, Part B Computer Aided Design and DNA Assembly*. vol. 498 av *Methods in Enzymology*. Academic Press; 2011. s. 349–361.
- [22] Gibson Assembly® Master Mix | NEB;. (Besökt 05/15/2015). <https://www.neb.com/products/e2611-gibson-assembly-master-mix>.
- [23] Ito H, Fukuda Y, Murata K, Kimura A. Transformation of intact yeast cells treated with alkali cations. *J Bacteriol*. 1983 jan;153(1):163–168.
- [24] Gietz RD, Woods RA. Genetic transformation of yeast. *BioTechniques*. 2001 apr;30(4):816–820.
- [25] Heigwer F, Kerr G, Boutros M. E-CRISP: fast CRISPR target site identification. *Nat Methods*. 2014 feb;11(2):122–123.
- [26] Glick BR, Pasternak JJ, Patten CL. *Molecular Biotechnology: Principles and Applications of Recombinant DNA*. 4. uppl. 1752 N St. NW, Washington, DC 20036-2904, USA: ASM Press; 2010.
- [27] Liu J, Martin-Yken H, Bigey F, Dequin S, Francois JM, Capp JP. Natural yeast promoter variants reveal epistasis in the generation of transcriptional-mediated noise and its potential benefit in stressful conditions. *Genome Biol Evol*. 2015;7(4):969–984.
- [28] Hanahan D. Studies on Transformation of *Escherichia coli* with Plasmids. *J Mol Biol*. 1983;166:969–984.
- [29] Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. *Molecular biology of the cell*. 5. uppl. 711 Third Avenue, 8th floor, New York, NY 10017, USA & 3 Park Square, Milton Park, Abingdon, OX14 4RN, UK.: Garland Science, Taylor & Francis Group; 2008.
- [30] GSL Biotech. SnapGene® software;. <http://www.snapgene.com/>.

## A Processchema

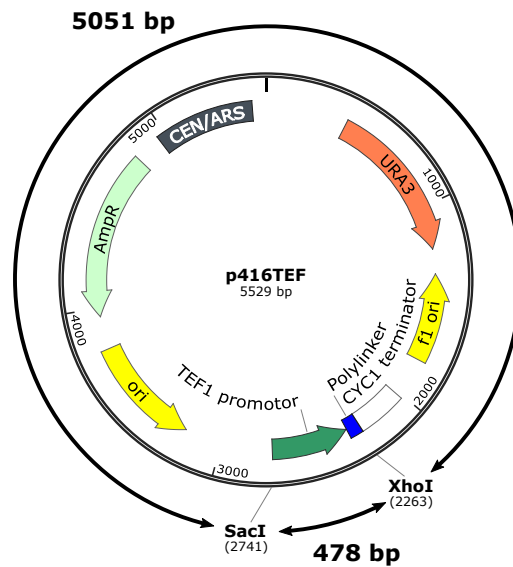


**Figur 17:** Processchema för det laborativa arbetet. Varje aktivitet skrivs i fetstil med relevanta metoder under i normal text.

## B Originalplasmider med restriktionsfragment för *SacI* och *XhoI*

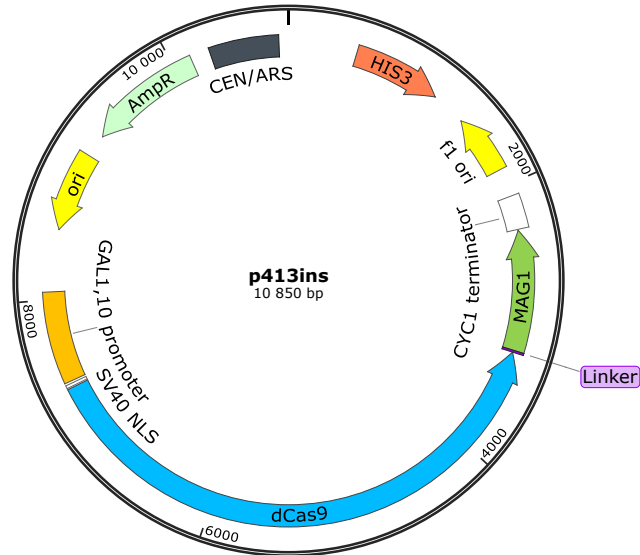


**Figur 18:** Genkarta över p413TEF med restriktionsfragment för *SacI* och *XhoI* markerade. Bilden genererades med SnapGene<sup>®</sup> software [30] och bearbetades.

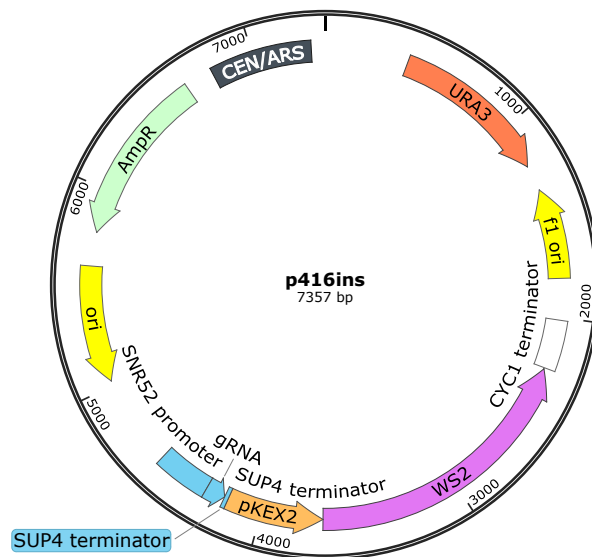


**Figur 19:** Genkarta över p416TEF med restriktionsfragment för *SacI* och *XhoI* markerade. Bilden genererades med SnapGene<sup>®</sup> software [30] och bearbetades.

## C Plasmidkonstruktioner



**Figur 20:** Genkarta över p413ins med *GAL1*-promotor, *SV40 NLS*, *dCas9* och *MAG1* infogat. *ampR* ger ampicillinresistens och *HIS3* möjliggör produktion av histidin. Bilden genererades med SnapGene<sup>®</sup> software [30] och bearbetades.



**Figur 21:** Genkarta över p416ins med *WS2*, *pKEX2* och integrationskassett för gRNA infogat. *ampR* ger ampicillinresistens och *URA3* möjliggör produktion av uracil. Bilden genererades med SnapGene<sup>®</sup> software [30] och bearbetades.

## D Förberedelse av och transformation till kompetenta jästceller

### Förberedelse av kompetenta jästceller

1. Späd en jästkultur över natten i 50 ml YPD till en  $OD_{600}$  på ungefär 0,25.
2. Inkubera kulturen vid 30 °C på en skakinkubator tills den når en  $OD_{600}$  på cirka 1,0 (ungefär 4-5 timmar)
3. Centrifugera cellerna i 5 minuter på 1000 g och kasta supernatanten.
4. Lös upp cellerna i 10 ml LTE-buffer.
5. Repetera steg 3.
6. Lös upp cellerna i 0,5 ml LTE-buffer.
7. Fördela 50  $\mu$ l av de upplösta cellerna till varsitt Eppendorfrör (ett för varje transformation).

### Transformation

1. Tillsätt det DNA som ska transformeras, ungefär 1  $\mu$ g av varje plasmid.
2. Tillsätt 50  $\mu$ l ssDNA med en koncentration av 2 mg/ml.
3. Tillsätt 300  $\mu$ l av transformationsblandningen. Blanda väl genom att pipettera upp och sedan ner uppreparande gånger.
4. Inkubera i 30 minuter vid 30 °C.
5. Inkubera i 15 min vid 42 °C.
6. Tillsätt innehållet i varje tub till en lämplig platta. Sprid ut lösningen försiktigt på plattan och låt den torka tills all vätska absorberats.

#### LTE-buffer

0,1 M LiAc  
10 mM Tris-HCl (pH 7,5)  
1 mM EDTA

#### Transformationsblandning

40 % PEG 3350  
0,1 M LiAc  
10 mM Tris-HCL (pH 7,5)  
1 mM EDTA



## E PCR

**Tabell 2:** Volym av respektive komponent i en 50  $\mu$ l PCR-reaktion.

Komponent	50 $\mu$ l reaktion
H <sub>2</sub> O	späd till 50 $\mu$ l
5x Phusion HF Buffer	10 $\mu$ l
10 mM dNTPS	1 $\mu$ l
10 $\mu$ M Forward primer	2,5 $\mu$ l
10 $\mu$ M Reverse primer	2,5 $\mu$ l
DNA-templat	0,5 $\mu$ l
Phusion-polymeras	0,5 $\mu$ l

**Tabell 3:** PCR-schema för *MAG1*, *pGAL1*, gRNA, *pKEX2* samt *WS2*.

Steg i cykeln	Temperatur (°C)	Tid	Antal cykler
Initial denaturering	98,0	30 sek.	1
Denaturering	98,0	10 sek.	} 35
Hybridisering	61,3	30 sek.	
Polymerisering	72,0	30 sek.	
Slutlig polymerisering	72,0	10 min	1
	15,0	$\infty$	-

**Tabell 4:** PCR-schema för *dCas9*.

Steg i cykeln	Temperatur (°C)	Tid	Antal cykler
Initial denaturering	98,0	30 sek.	1
Denaturering	98,0	10 sek.	} 35
Hybridisering	63,0	30 sek.	
Polymerisering	72,0	65 sek.	
Slutlig polymerisering	72,0	10 min	1
	15,0	$\infty$	-

## F Laborationsinstruktioner för koloni-PCR

### Förberedelse av koloni-PCR

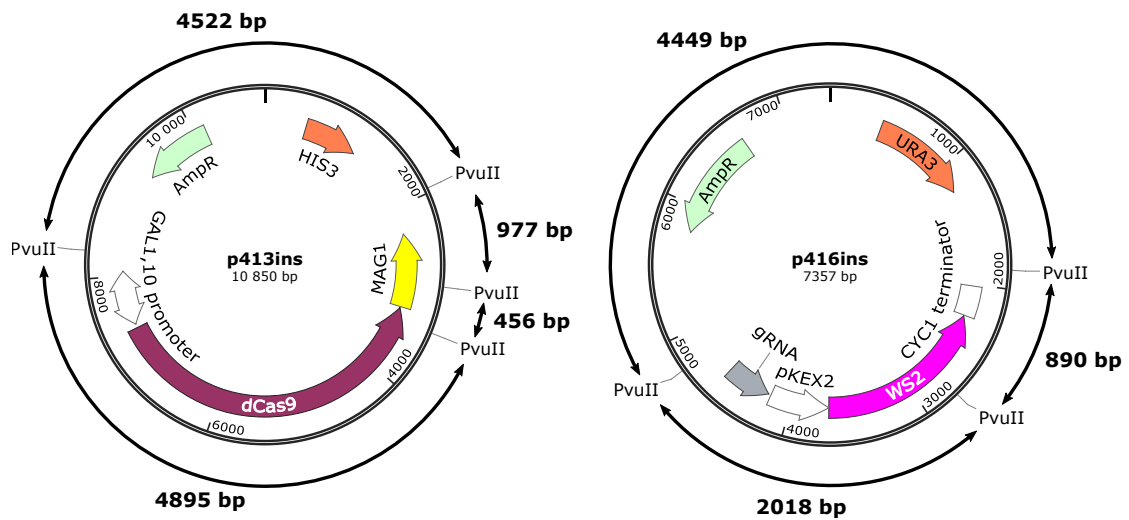
1. Förbered Eppendorfrören med 25  $\mu\text{l}$  av en 20 mM NaOH-lösning.
2. Tillsätt några få glaspärlor till varje tub.
3. Överför en liten bit av kulturen till tuben med en steril pipettspets och resuspendera jästcellerna i lösningen.
4. Koka provet vid 97,5 °C i 15 minuter.
5. Blanda lösningen med en vortex mixer i 15 sekunder.
6. Centrifugera ner lösningen under maximal hastighet i 1 minut.
7. Överför 0,5  $\mu\text{l}$  av lösningen till 24,5  $\mu\text{l}$  PCR-reaktionslösning.
8. Använd åtminstone 35 cykler för att få bra band.

### Förberedelse av PCR-reaktionslösningen

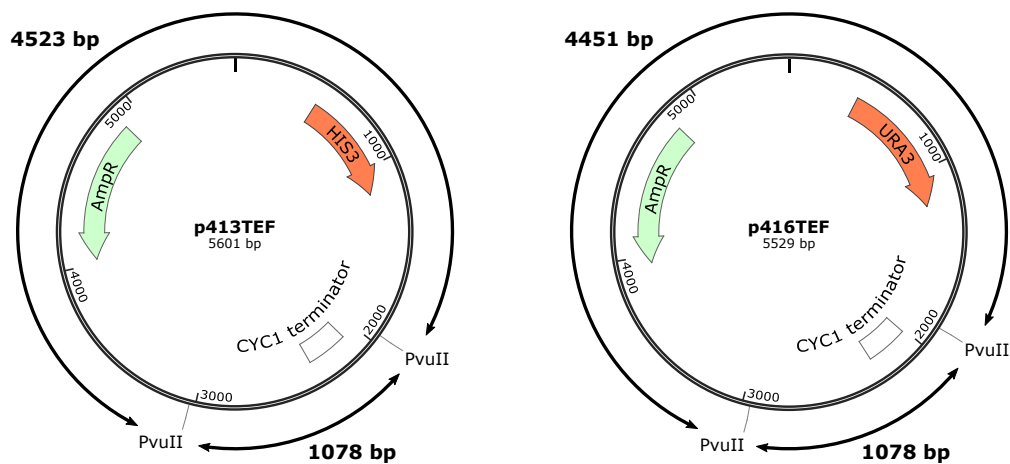
**Tabell 5:** Volym för respektive komponent i en 25  $\mu\text{l}$  PCR-reaktionslösning för koloni-PCR.

Komponent	Volym ( $\mu\text{l}$ )
Phusion-buffer	5
dNTPs	0,5
Forward primer	1,25
Reverse primer	1,25
Phusion-polymeras	0,25
DNA-templat	0,5
Vatten	Späd till 25 $\mu\text{l}$

## G Klyvningsställen *PvuII*, p413 och p416



**Figur 22:** Storlek av restriktionsfragment med *PvuII* för p413ins (vänster) och p416ins (höger). Bilden genererades med SnapGene<sup>®</sup> software [30] och bearbetades.



**Figur 23:** Storlek av restriktionsfragment med *PvuII* för p413TEF (vänster) och p416TEF (höger). Bilden genererades med SnapGene<sup>®</sup> software [30] och bearbetades.

## H DNA-sekvenser för primrar

### H.1 Primrar för förlängning av segment inför Gibsonsammanfogning

Tabell 6: Tabell med primrar för Gibsonsammanfogning av p413ins.

<i>MAG1</i> -segmentet	
<i>dCas9</i> → <u>Linker</u> → <i>MAG1</i>	GTGGT <u>GATGGAGGTGGAGGTGGAATGAAACTAAA</u> AAGGGAGTATGATG
Plasmid → <u><i>MAG1</i></u>	ATGTAAGCGTGACATAACTAATTACATGACTCGA GTTAGGATTTACGAAATTTTCTTC
<i>dCas9</i> -segmentet	
<i>MAG1</i> → <u>Linker</u> → <i>dCas9</i> -NLS	GTTTCATTCCACCTCCACCTCCATCACCACCTAA TTGAGACAAG
p <i>GAL1</i> → <u><i>dCas9</i>-NLS</u>	GGAGAAAAAACTATAATGTCTAGAGCCCCAAAGA AG
p <i>GAL1</i> -segmentet	
<i>dCas9</i> -NLS → <u>p<i>GAL1</i></u>	TGGGGCTCTAGACATTATAGTTTTTCTCCTTGA CGTTAAAG
Plasmid → <u>p<i>GAL1</i></u>	CCTCACTAAAGGGAACAAAAGCTGGAGCTCTTTC AAAAATTCTTACTTTTTTTTTTGG

**Tabell 7:** Tabell med primrar för Gibsonsammanfogning av p416ins.

<b>gRNA-kassetten</b>	
Plasmid → <u>gRNA</u>	CCTCACTAAAGGGAACAAAAGCTGGAGCTCTCTTTGAAAAGATAA <u>TGTATGATTATGC</u>
<u>pKEX2</u> → <u>gRNA</u>	CATCAGAGCCGCTGAAGACATAAAAAACAAAAAAGCAC
<b>pKEX2-segmentet</b>	
<u>gRNA</u> → <u>pKEX2</u>	TTGTTTTTTATGTCCTCAGCGGCTCTGATGTAGAT
<u>WS2</u> → <u>pKEX2</u>	ACCTAATCTCTTCATCTGATAATGGGTTTGTAGTTTATAATTATG <u>TG</u>
<b>WS2-segmentet</b>	
<u>pKEX2</u> → <u>WS2</u>	CAAACCCATTATCAGATGAAGAGATTAGGTACTCTAGACG
Plasmid → <u>WS2</u>	ATGTAAGCGTGACATAACTAATTACATGACTCGAGTTACTTTCTA <u>GTACGGGCACG</u>

## H.2 Primrar för sekvensering

**Tabell 8:** Primrar för sekvensering av insertet i p413ins. Samtliga primrar går i samma riktning och är avskiljda med ca 800 bp, då detta är längden inom vilket sekvenseringsresultatet är som säkrast.

Primer #	Sekvens
1	GGGAGGGCGTGAATGTAAG
2	CATGTCTAGCAATATATTTCTCAGG
3	AGTTGGAGAATCAAACCCAC
4	ACCTCTATTCTTATCGGATCTAG
5	AGGCATTGAATCTATCTTCGAC
6	GAGCCTTAGTAATTTGCGGTGTTG
7	AAACCTTGAACCTTCTTGGATG

**Tabell 9:** Primrar för sekvensering av insertet i p416ins. Samtliga primrar går i samma riktning och är avskiljda med ca 800 bp, då detta är längden inom vilket sekvenseringsresultatet är som säkrast.

Primer #	Sekvens
1	TTTTCGGTTAGAGCGGATGTG
2	CAGGGTGTCTGACAGAATG
3	GGCTTATGACAGAATTATGCATATG

**Tabell 10:** Primrar för sekvensering av *WS2*. Primrarna sitter på överlappen och är därför motsatt riktade.

Primerriktning	Sekvens
Forward	CGTGACATAACTAATTACATGACTC
Reverse	CCATTATCAGATGAAGAGATTAGG

### H.3 Primrar ytterligare kontroll av insert

**Tabell 11:** Primrar för kontroll av inserten p413ins och p416ins.

Primer	Sekvens
p413 forward 1	GGAGGGCGTGAATGTAAGC
p413 forward 2	GATGTTCCCTCCACCAAAGG
p416 forward	GGAAGCATATTTGAGAAGATGC
Reverse	CCTCACTAAAGGGAACAAAAGC

# I DNA-sekvenser för konstruerade plasmider

## I.1 p413ins

tCYC1	bp 2086-2333
MAG1	bp 2340-3230
Linker	bp 3231-3245
dCas9	bp 3246-7346
SV40 NLS	bp 7356-7376
pGAL1	bp 7393-8057

```
1 TCGCGGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA
61 CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGCGTG
121 TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA GCAGATTGTA CTGAGAGTGC
181 ACCATAAATT CCCGTTTTAA GAGCTTGGTG AGCGCTAGGA GTCACTGCCA GGTATCGTTT
241 GAACACGGCA TTAGTCAGGG AAGTCATAAC ACAGTCCTTT CCCGCAATTT TCTTTTTCTA
301 TTACTCTTGG CCTCCTCTAG TACACTCTAT ATTTTTTAT GCCTCGGTAA TGATTTTCAT
361 TTTTTTTTTT CCCCTAGCGG ATGACTCTTT TTTTTCTTA GCGATTGGCA TTATCACATA
421 ATGAATTATA CATTATATAA AGTAATGTGA TTCTTCGAA GAATATACTA AAAAATGAGC
481 AGGCAAGATA AACGAAGGCA AAGATGACAG AGCAGAAAAG CCTAGTAAAG CGTATTACAA
541 ATGAAACCAA GATTCAGATT GCGATCTCTT TAAAGGGTGG TCCCCTAGCG ATAGAGCACT
601 CGATCTTCCC AGAAAAAGAG GCAGAAGCAG TAGCAGAACA GGCCACACAA TCGCAAGTGA
661 TTAACGTCCA CACAGGTATA GGGTTTCTGG ACCATATGAT ACATGCTCTG GCCAAGCATT
721 CCGGCTGGTC GCTAATCGTT GAGTGCATG GTGACTTACA CATAGACGAC CATCACACCA
781 CTGAAGACTG CGGGATTGCT CTCGGTCAAG CTTTTAAAGA GGCCCTACTG GCGCGTGGAG
841 TAAAAAGGTT TGGATCAGGA TTTGCGCCTT TGGATGAGGC ACTTCCAGA GCGGTGGTAG
901 ATCTTTCGAA CAGGCCGTAC GCAGTTGTCG AACTTGGTTC GCAAAGGGAG AAAGTAGGAG
961 ATCTCTCTTG CGAGATGATC CCGCATTTC TGTAAAGCTT TGCAGAGGCT AGCAGAATTA
1021 CCCTCCAGCT TGATTGTCTG CGAGGCAAGA ATGATCATCA CCGTAGTGAG AGTGCCTTCA
1081 AGGCTCTTGC GGTGCCATA AGAGAAGCCA CCTCGCCCAA TGGTACCAAC GATGTTCCCT
1141 CCACCAAAGG TGTTCTTATG TAGTGACACC GATTATTTAA AGCTGCAGCA TACGATATAT
1201 ATACATGTGT ATATATGTAT ACCTATGAAT GTCAGTAAGT ATGTATACGA ACAGTATGAT
1261 ACTGAAGATG ACAAGGTAAT GCATCATTCT ATACGTGTCA TTCTGAACGA GCGCGCTTT
1321 CTTTTTTTCT TTTTGCTTTT TCTTTTTTTT TCTCTTGAAC TCGACGGATC TATGCGGTGT
1381 GAAATACCGC ACAGATGCGT AAGGAGAAAA TACCGCATCA GGAAATTGTA AACGTTAATA
1441 TTTTGTAAA ATTTCGCTTA AATTTTTGTT AAATCAGCTC ATTTTTTAAC CAATAGGCCG
1501 AAATCGGCAA AATCCCTTAT AAATCAAAAAG AATAGACCGA GATAGGGTTG AGTGTGTGTC
1561 CAGTTTGAA CAAGAGTCCA CTATTAAGA ACGTGGACTC CAACGTCAA GGGCGAAAAA
1621 CCGTCTATCA GGGCGATGGC CCACTACGTG AACCATCACC CTAATCAAGT TTTTGGGGT
1681 CGAGGTGCCG TAAAGCACTA AATCGGAACC CTAAGGGAG CCCCGATT TTTAGCTTGAC
1741 GGGGAAAGCC GGCGAACGTG GCGAGAAAAG AAGGGAAGAA AGCGAAAGGA GCGGGCGCTA
1801 GGGCGCTGGC AAGTGTAGCG GTCACGCTGC GCGTAACCAC CACACCCGCC GCGCTTAATG
1861 CGCCGTACA GGGCGCTCG CGCCATTCCG CATTGAGGCT GCGCAACTGT TGGGAAGGGC
1921 GATCGGTGCG GGCCTCTTCG CTATTACGCC AGCTGGCGAA AGGGGATGT GCTGCAAGGC
1981 GATTAAGTTG GGTAAAGCCA GGGTTTTCCC AGTCACGACG TTGTAAAACG ACGGCCAGTG
2041 AGCGCGGTA ATACGACTCA CTATAGGGCG AATTGGGTAC CGGCCCAAAA TTTAAAGCCTT
2101 CGAGCGTCCC AAAACCTTCT CAAGCAAGGT TTTCAGTATA ATGTTACATG CGTACACCGG
2161 TCTGTACAGA AAAAAAGAA AAATTTGAAA TATAAATAAC GTTCTTAATA CTAACATAAC
2221 TATAAAAAAA TAAATAGGGA CTAAGACTTC AGGTTGTCTA ACTCCTTCCT TTTGCGTTAG
2281 AGCGGATGTG GGGGAGGGG GTGAATGTAA GCGTGACATA ACTAATTACA TACTCGAGT
```

2341 TAGGATTTC A CGAAATTTTC TTCTGCCTTC ATCATGGCAT CTGTATTTGT GCTCGCGAGC  
2401 CTCCAAAGTA TGAACATAAA CACAGACCTA TACGGAGAAA ATGTTTCAGA GCATTTTTCC  
2461 ATTATGTGCT CGTCATATAT TTTCCAGTTG TATTTCTTAT GCTTAATCTT ACTCTTTTTA  
2521 ACTACTTTTC TTTACGCAT TAATTCCTTT TCCAATCTG GCTTATCTGA AAGGTATTTT  
2581 GAAAAACCCC TAGCAATACC TAGATCTTCA GGAGCAAATA CATCCATTCT TTTCAATCCG  
2641 GAGATCAAGA ACATTTTGGC ACTCCATGGG CCTATACCTT TTACATTCGT AACTAAACTT  
2701 TCAATCACTT CCTCATCATI ATCTTTTTGA CCGAAGAGCT TTTGATATC CTTATATTTT  
2761 TCAGTAAAAAT AGACAGCAAG AGACTCTAGA TATATCATTI TCCTTTTACT CAATCCACAT  
2821 TTTGCGATTT CTGCACATTT TGCTGGGTCT TTGAAGTCTT CGAAAAGGAT TTTGTAATCA  
2881 GGAATGCAC CGCCATAAAG ACTGACAACC CTGCTTGA TGCTTTCAGC TGCTGGCCA  
2941 CTGATCTGTT GAGACAAAAT TGTGCTTGCA AGCCTAATAA AATAATCTTC GAGTGTATTA  
3001 GGGACTTGAG TCTCCTTCAA GTACAACGTA AATTCATTAT TCTTAAGTAT GGGAAAAAGT  
3061 GATGGATCTT TCTCTAAAAT GTGTTGCAA GCCATATTGA ACTTTTCTTC ATGTCTAGCA  
3121 ATATATTTCT CAGGAAGAGC AACCTCTAGA GGTCGAGACC CTAATCTTTT TGCTATTTCC  
3181 TTAACAGCGT CTGCTTTTAT TAACTCATCA TACTCCCTTT TTAGTTTCAT TCCACCTCCA  
3241 CCTCCATCAC CACCTAATTG AGACAAGTCG ATTCTAGTTT CGTACAAACC AGTAATAGAT  
3301 TGGTGATCA AAGTAGCGTC CAAAACCTCT TTGGTAGAGG TGTATCTCTT TCTGTGATA  
3361 GTGGTATCGA AATACTTAAA AGCAGCTGGA GCACCCAAGT TAGTCAAAGT GAACAAGTGA  
3421 ATGATGTTTT CAGCTTGTTC TCTAATTGGC TTATCTCTGT GCTTGTGTA AGCTGACAAA  
3481 ACTTTGTCCA AATTGGCGTC AGCCAAAATG ACTCTTTTGG AAAATTCGGA AATTTGTTCA  
3541 ATGATTTGCT CCAAGTAGTG CTTGTGTTGT TCGACAAATA ATTGCTTTTG TTCGTTATCT  
3601 TCAGGAGAAC CTTTCAATTT TTGCTAGTGG GAGGCCAAGT ATAAGAAAT AACGTACTTG  
3661 GATGGCAAAG CCAATTCATT ACCCTTTTGT AATTCACCGG CAGAAGCCAA CATCTCTTT  
3721 CTACCATTTT CCAATTCGAA CAAACTGTAC TTTGGCAACT TAATGATCAA GTCCTTCTTG  
3781 ACTTCTTTGT AACCTTGGC TTCCAAAAAG TCGATAGGAT TCTTTTCGAA TGAGGATCTT  
3841 TCCATGATAG TGATACCTAA TAATCTTTG ACGGACTTCA ACTTCTTACT CTTACCCTTT  
3901 TCGAGATAG CAACAACCAA AACAGAGTAA GCAACAGTTG GAGAATCAA ACCACCCTAC  
3961 TTTTTGGGT TTTTCTAGCG ATCAACTTAT CCGAGTTTCT CTTAGGCAAA  
4021 ATGGATTCTT TGGAGAAACC ACCAGTTTGA ACTTCGGTCT TTTTGACGAT GTTTACTTGT  
4081 GGCATGGACA ATACCTTTCT AACAGTGGCA AAATCTCTAC CCTTATCCCA AACGATTCA  
4141 CCTGTTTAC CATTAGTTTC AATCAATGGT CTTTTTCTAA TTTACCGTT GGCCAAAGTG  
4201 ATTTCCGTCT TAAAGAAAT CATAATATG GAGTAGAAGA AGTATTGGC AGTAGCTTTA  
4261 CCAATTTCTT GTTCGGACTT GGGATCATC TTTCTAACGT CGTAAACCTT GTAATCACCG  
4321 TAGACAAAT CCGATTCCAA CTTTGGATAC TCTTAAATCA AGGCAGTACC AACAAACAGCA  
4381 TTCAAGTAAG CATCATGTGC ATGATGGTAG TTGTTGATTT CTCTGACCTT GTAGAATTGG  
4441 AAGTCCTTC TGAATCGGA AACCAACTTG GACTTCAAGG TAATGACCTT AACTTCTCTG  
4501 ATCAATTTGT CGTTTTGCTG TACTTAGTGT TTCATTCTAC TATCCAAAAT TTGAGCAACG  
4561 TGCTTGGTGA TTTGTCTGGT TTCTACTAAT TGTCTTTTGA TGAACCTGC CTTATCCAAT  
4621 TCAGACAAAC CACCTCTTTC AGCCTTGGTC AAATTGTCAA ACTTTCTTTG GGTGATCAAC  
4681 TTGGCCTTCA ATAATTGTCT CCAATAGTTC TTCATTTTCT TAACGACTTC TTCAGATGGA  
4741 ACGTTATCGG ACTTACCTCT ATTCTTATCG GATCTAGTCA AGACCTTGTI GTCAATGGAG  
4801 TCATCCTTCA AAAAAAGATT TGGAACGATA GCGTCAACAT CGTAATCAGA CAATCTGTTG  
4861 ATGTCCAATT CTTGGTCAAC GTACATATCT CTACCGTTTT GCAAGTAGTA CAAGTACAAC  
4921 TTTTCGTTTT GTAATTGGGT GTTTTCAACT GGGTGTCTT TCAAAATTTG GGAACCCAAT  
4981 TCTTTGATAC CTTCTTCGAT TCTTTCATT CTTTCTCTGG AGTTCTTTTG ACCCTTTTGG  
5041 GTGGTTTGAT TTTCTCTAGC CATTTCGATA ACGATTTTT CTGGCTTATG TCTACCCATA  
5101 ACTTTGACCA ATTCGTCAAC GACCTTAAAG GTTTGTAAAA TACCCTTCTT AATAGCTGGA  
5161 GAACCAGCCA AGTTAGCGAT ATGTTTATGT AAAGAATCAC CTTGACCAGA GACTTGGGCC  
5221 TTTTGTATAT CTTCTTTGAA AGTCAAGGAA TCGTCTGGA TTAATGTCAT GAAGTTCTG  
5281 TTAGCGAAAC CATCAGACTT CAAAAAGTCC AAAATGGTCT TACCGGATTG CTTGTCTCTG  
5341 ATACCGTTAA TCAATTTTCT GGACAATCTA CCCCAACCAG TGTATCTTCT TCTCTTAAAT  
5401 TGCTTCATAA CCTTATCATC GAACAAGTGG GCGTAGGTTT TCAATCTTTC TTCAATCATI  
5461 TCTCTATCTT CAAACAAGGT CAATGTCAAG ACAATATCTT CCAAAATATC TTCGTTTTCT  
5521 TCGTTGTCCA AGAAATCCTT GTCCTTGATG ATTTTCAACA AGTCATGGTA AGTACCTAAA



5581 GAGGCATTGA ATCTATCTTC GACACCAGAA ATTTCAACGG AGTCGAAGCA TTCGATTTTC  
5641 TTGAAGTAAT CTTCTTTTAA TTGCTTGACA GTAACCTTTC TGTTGGTCTT GAACAACAAA  
5701 TCGACGATAG CCTTCTTTTG TTCACCGGAT AAAAAAGCTG GTTTTCTCAT ACCTTCGGTG  
5761 ACGTACTTAA CCTTAGTCAA TTCGTTGTAG ACGGTGAAGT ATTCGTACAA CAAAGAGTGC  
5821 TTGGGCAAAA CCTTTCGTT TGGCAAGTTC TTGTCGAAGT TGGTCATTCT TTCAAATAAG  
5881 GATTGGGCAG AAGCACCTT ATCTACAAC TCTTCGAAGT TCCAAGGAGT AATAGTTTCT  
5941 TCGGACTTTC TAGTCATCCA GGCAATCTT GAATTACCTC TAGCCAATGG ACCAACGTAG  
6001 TAAGGGATTC TGAAGGTCAG AATCTTTTCG ATCTTTTCTC TGTTGTCTT CAAGAATGGA  
6061 TAAAAATCTT CTTGTCTTCT CAAAATTGGC TGCAATTCAC CCAAGTGAAT TTGATGTGGA  
6121 ATAGAACCGT TGTCGAAGGT TCTTTGCTTT CTCAACAAAT CTTCTCTGTT CAACTTGACC  
6181 AATAATCTT CAGTACCGTC CATCTTTTCC AAGATTGGCT TGATGAACCT GTAAAAATCT  
6241 TCTTGAGAAG CACCACCATC AATATAACCG GCGTAACCGT TCTTGATTG ATCGAAGAAA  
6301 ATTTCTTTGT ACTTTTCAGG TAATTGTGT CTGACCAAAG CCTTCAACAA AGTCAAGTCT  
6361 TGGTGGTGT CATCGTATCT TTTGATCATA GAGGCAGACA ATGGAGCCTT AGTAATTTGC  
6421 GTGTTGACTC TCAAAATATC GGACAACAAA ATGGCATCGG ACAAGTTCTT AGCAGCCAAA  
6481 AACAAATCAG CGTATTGATC ACCGATTTGA GCCAACAAAT TATCCAAATC GTCATCGTAG  
6541 GTATCCTTAG ATAATTGCAA CTTGGCATCT TCAGCCAAAT CGAAATGGA CTTGAAGTTT  
6601 GGAGTCAAAC CCAAAGACAA GGCAATCAAA TTACCGAACA AACCATCTT CTTTTCACCT  
6661 GGTAATTGGG CGATCAAAT TTCTAATCTT CTGGACTTGG ACAATCTGGC AGACAAAATA  
6721 GCTTTAGCAT CAACACCAGA GCGTGTGATT GGGTTTTCTT CGAATAATTG ATTTAGGTT  
6781 TGGACTAAT GGATGAACAA CTTATCAACA TCGGAGTTAT CTGGGTTCAA ATCACCTTCG  
6841 ATCAAGAAAT GACCTCTGAA CTTAATCATG TGGGCCAAAG CCAAATAGAT CAATCTCAAA  
6901 TCAGCCTTAT CGGTAGAGTC AACCAACTTC TTTCTCAAAAT GGTAGATAGT TGGGTACTTT  
6961 TCGTGGTAAG CAACTTCATC AACGATGTTA CCGAAGATTG GATGTCTTTC GTGCTTCTTA  
7021 TCTTCTCGA CCAAGAAGGA TTCTTCCAAT CTGTGGAAGA ATGAATCATC AACCTTGGCC  
7081 ATTTCTGTTG AGAAGATTTC TTGCAAGTAG CAGATTCTGT TCTTTCTTCT GGTGTATCTT  
7141 CTTCTAGCGG TTCTTTTCAA TCTAGTAGCT TCAGCAGTTT CACCAGAATC GAACAACAAG  
7201 GCACCAATCA AGTTCTTTTT GATGGAGTGT CTATCAGTGT TACCCAAAAC CTTGAAATTC  
7261 TTGGATGGAA CCTTGTATTC ATCAGTAATA ACAGCCCAAC CAACAGAGT TGTACCGATA  
7321 GCCAAACCAA TAGAGTATTT CTTATCCCG GGTCTAACTT TTCTCTTCTT CTTTGGGGCT  
7381 CTAGACATTA TAGTTTTTTC TCCTTGACGT TAAAGTATAG AGGTATATTA ACAATTTTTT  
7441 GTTGATACTT TTATTACAT TGAATAAGAA GTAATACAAA CCGAAAATGT TGAAAGTAT  
7501 AGTTAAAGTC GTTATGCAGT TTTTGCATTT ATATATCTGT TAATAGATCA AAAATCATCG  
7561 CTTCCGCTGAT TAATTACCCC AGAATAAAG CTAAAAAACT AATCGCATT TCACTCCTATG  
7621 GTTGTTAATT TGATTCGTTT ATTTGAAGGT TTGTGGGGCC AGGTACTGCT CAATTTTTCC  
7681 TCTTCATAAC CATAAAAGCT AGTATTGTAG AATCTTTATT GTTCGGAGCA GTGCGGCGCG  
7741 AGGCACATCT GCGTTTCAGG AACGCGACCG GTGAAGACGA GGACGCACGG AGGAGAGTCT  
7801 TCCTTCGGAG GCGTGTCCAC CGCTCGGCGG CTTCTAATCC GTACTTCAA ATAGCAATGA  
7861 GCAGTTAAGC GTATTACTGA AAGTTCCAAA GAGAAGGTTT TTTTAGGCTA AGATAATGGG  
7921 GCTCTTTACA TTTCCACAAC ATATAAGTAA GATTAGATAT GGATATGTAT ATGGATATGT  
7981 ATATGGTGGT AATGCCATGT AATATGATTA TTAACCTTCT TTGCGTCCAT CCAAAAAAAA  
8041 AGTAAGAATT TTTGAAAAG CTCCAGCTTT TGTCCCTTT AGTGAGGGTT AATGCGCGC  
8101 TTGGCGTAAT CATGGTCATA GCTGTTTCTT GTGTGAAATT GTTATCCGCT CACAATTTCA  
8161 CACAACATAG GAGCCGGAAG CATAAAGTGT AAAGCCTGGG GTGCCTAATG AGTGAGGTAA  
8221 CTCACATTA TCGCTTGGC CTCACTGCCG GCTTTCAGT CCGGAAACCT GTCGTGCCAG  
8281 CTGCATTAAT GAATCGGCCA ACGCGCGGGG AGAGGCGGTT TGCGTATTGG GCGCTTCTCC  
8341 GCTTCCTCGC TCACTGACTC GCTGCGCTCG GTCGTTCCGG TGCGGCGAGC GGTATCAGCT  
8401 CACTCAAAGG CGGTAATACG GTTATCCACA GAATCAGGGG ATAACGCAGG AAAGAACATG  
8461 TGAGCAAAAAG GCCAGCAAAA GGCCAGGAAC CGTAAAAAGG CCGCGTTGCT GCGTTTTTTC  
8521 CATAGGCTCC GCCCCCTGA CGAGCATCAC AAAAATCGAC GCTCAAGTCA GAGGTGGCGA  
8581 AACCCGACAG GACTATAAAG ATACCAGCG TTTCCCTCGT GAAGCTCCCT CGTCCGCTCT  
8641 CCTGTTCCGA CCCTGCCGCT TACCGGATAC CTGTCCGCTT TTCTCCCTTC GGAAGCGGTG  
8701 GCGCTTTCTC ATAGCTCAGC CTGTAGGTAT CTCAGTTCCG TGTAGGTCGT TCGTCCAAG  
8761 CTGGGCTGTG TGCACGAACC CCCCCTTCCG CCCGACCGCT GCGCCTTATC CCGTAACTAT

```

8821 CGTCTTGAGT CCAACCCGGT AAGACACGAC TTATCGCCAC TGGCAGCAGC CACTGGTAAC
8881 AGGATTAGCA GAGCGAGGTA TGTAGGCGGT GCTACAGAGT TCTTGAAGTG GTGGCCTAAC
8941 TACGGCTACA CTAGAAGGAC AGTATTTGGT ATCTGCGCTC TGCTGAAGCC AGTTACCTTC
9001 GGAAAAAGAG TTGGTAGCTC TTGATCCGGC AAACAAACCA CCGCTGGTAG CGGTGGTTTT
9061 TTTGTTTGCA AGCAGCAGAT TACGCGCAGA AAAAAAGGAT CTCAAGAAGA TCCTTTGATC
9121 TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACTCAC GTTAAGGGAT TTTGGTCATG
9181 AGATTATCAA AAAGGATCTT CACCTAGATC CTTTTAAATT AAAAATGAAG TTTTAAATCA
9241 ATCTAAAGTA TATATGAGTA AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA
9301 CCTATCTCAG CGATCTGTCT ATTTTCGTTCA TCCATAGTTG CCTGACTCCC CGTCGTGTAG
9361 ATAACACGA TACGGGAGGG CTTACCATCT GGCCCCAGTG CTGCAATGAT ACCGCGAGAC
9421 CCACGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACCAGC CAGCCGGAAG GGCCGAGCGC
9481 AGAAGTGGTC CTGCAACTTT ATCGCCTCC ATCCAGTCTA TTAATTGTTG CCGGAAGCT
9541 AGAGTAAGTA GTTCGCCAGT TAATAGTTTG CGCAACGTTG TTGCCATTGC TACAGGCATC
9601 GTGGTGTAC GCTCGTCGTT TGGTATGGCT TCATTCAGCT CCGGTTCCCA ACGATCAAGG
9661 CGAGTTACAT GATCCCCAT GTTGTGCAAA AAAGCGGTTA GCTCCTTCGG TCCTCCGATC
9721 GTTGTACAGAA GTAAGTTGGC CGCAGTGTTA TCACTCATGG TTATGGCAGC ACTGCATAAT
9781 TCTCTTACTG TCATGCCATC CGTAAGATGC TTTTCTGTGA CTGGTGAGTA CTCAACCAAG
9841 TCATTCTGAG AATAGTGTAT GCGGCGACCG AGTTGCTCTT GCCCGGCGTC AATACGGGAT
9901 AATACCGCGC CACATAGCAG AACTTTAAAA GTGCTCATCA TTGAAAAACG TTCTTCGGGG
9961 CGAAAACTCT CAAGGATCTT ACCGCTGTTG AGATCCAGTT CGATGTAACC CACTCGTGCA
10021 CCCAACTGAT CTTCAGCATC TTTTACTTTC ACCAGCGTTT CTGGGTGAGC AAAAACAGGA
10081 AGGC AAAATG CCGCAAAAAA GGAATAAAG GCGACACGGA AATGTTGAAT ACTCATACTC
10141 TTCTTTTTC AATATTATTG AAGCATTAT CAGGGTATT GTCTCATGAG CGGATACATA
10201 TTTGAATGTA TTTAGAAAAA TAAACAAATA GGGGTTCCGC GCACATTTCC CCGAAAAAGTG
10261 CCACCTGGGT CCTTTTCATC ACGTGCTATA AAAATAATTA TAATTTAAAT TTTTAAATAT
10321 AAATATATAA ATTA AAAATA GAAAGTAAAA AAAGAAATTA AAGAAAAAAT AGTTTTTGTG
10381 TTCCGAAGAT GTAAAAGACT CTAGGGGGAT CGCCAACAAA TACTACCTTT TATCTTGCTC
10441 TTCTGTCTCT CAGGTATTAA TGCCGAATTG TTTCATCTTG TCTGTGTAGA AGACCACACA
10501 CGAAAAATCCT GTGATTTTAC ATTTTACTTA TCGTTAATCG AATGTATATC TATTTAATCT
10561 GCTTTTCTTG TCTAATAAAT ATATATGTAA AGTACGCTTT TTGTTGAAAT TTTTAAACC
10621 TTTGTTTATT TTTTTTCTT CATTCCGTAA CTCTTCTACC TTCTTTATTT ACTTTCTAAA
10681 ATCCAAATAC AAAACATAAA AATAAATAAA CACAGAGTAA ATTCCCAAAT TATTCCATCA
10741 TTA AAAAGATA CGAGGCGCGT GTAAGTTACA GGCAAGCGAT CCGTCCTAAG AAACCATTAT
10801 TATCATGACA TTAACCTATA AAAATAGGCG TATCAGCAGG CCCTTTCGTC

```

## I.2 p416ins

t <i>CYC1</i>	bp 2014-2261
WS2	bp 2268-3689
p <i>KEX2</i>	bp 3690-4177
gRNA-kassett	bp 4178-4565

```

1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA
61 CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG
121 TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA GCAGATTGTA CTGAGAGTGC
181 ACCATAACCAC AGCTTTTCAA TTCAATTCAT CATTTTTTTT TTATTTCTTT TTTTGATTTT
241 GTTTTCTTTG AAATTTTTTT GATTCGGTAA TCTCCGAACA GAAGGAAGAA CGAAGGAAGG
301 AGCACAGACT TAGATTGGTA TATATACGCA TATGTAGTGT TGAAGAAACA TGA AATTGGC
361 CAGTATCTT AACCCAACGT CACAGAACAA AAACCTGCAG GAAACGAAGA TAAATCATGT
421 CGAAAGCTAC ATATAAGGAA CGTGCTGCTA CTCATCCTAG TCCTGTTGCT GCCAAGCTAT

```

481 TTAATATCAT GCACGAAAAG CAAACAACT TGTGTGCTTC ATTGGATGTT CGTACCACCA  
541 AGGAATTACT GGAGTTAGTT GAAGCATTAG GTCCEAAAAT TTGTTTACTA AAAACACATG  
601 TGGATATCTT GACTGATTTT TCCATGGAGG GCACAGTTAA GCCGCTAAAG GCATTATCCG  
661 CCAAGTACAA TTTTTFCTC TTCGAAGACA GAAAATTTGC TGACATTGGT AATACAGTCA  
721 AATTGCAGTA CTCTGCGGGT GTATACAGAA TAGCAGAATG GGCAGACATT ACGAATGCAC  
781 ACGGTGTGGT GGGCCCAGGT ATTGTTAGCG GTTGAAGCA GCGGCAGAA GAAGTAACAA  
841 AGGAACCTAG AGGCCTTTG ATGTTAGCAG AATTGTCATG CAAGGGCTCC CTATCTACTG  
901 GAGAATATAC TAAGGGTACT GTTGACATTG CGAAGAGCGA CAAAGATTTT GTTATCGGCT  
961 TTATTGCTCA AAGAGACATG GGTGGAAGAG ATGAAGTTA CGATTGGTTG ATTATGACAC  
1021 CCGGTGTGGG TTTAGATGAC AAGGGAGACG CATTGGGTCA ACAGTATAGA ACCGTGGATG  
1081 ATGTGTCTC TACAGGATCT GACATTATTA TTGTTGGAAG AGGACTATTT GCAAAGGGAA  
1141 GGGATGCTAA GGTAGAGGGT GAACGTTACA GAAAAGCAGG CTGGGAAGCA TATTGAGAA  
1201 GATGCGGCCA GCAAAACTAA AAAACTGTAT TATAAGTAAA TGCATGTATA CTAAACTCAC  
1261 AAATTAGAGC TTCAATTTAA TTATATCAGT TATTACCCTA TGCGGTGTGA AATACCGCAC  
1321 AGATGCGTAA GGAGAAAATA CCGCATCAGG AAATTGTAAA CGTTAATATT TTGTTAAAA  
1381 TCGCGTAAA TTTTGTAAA ATCAGCTCAT TTTTAAACA ATAGGCCGAA ATCGGCAAAA  
1441 TCCCTTATAA ATCAAAAGAA TAGACCGAGA TAGGGTTGAG TGTTGTCCA GTTTGGAACA  
1501 AGAGTCCACT ATTAAGAAGC GTGGACTCCA ACGTCAAAGG GCGAAAAACC GTCTATCAGG  
1561 GCGATGGCCC ACTACGTGAA CCATCACCTT AATCAAGTTT TTTGGGGTCG AGGTGCCGTA  
1621 AAGCACTAAA TCGGAACCCT AAAGGGAGCC CCCGATTTAG AGCTTGACGG GGAAAGCCGG  
1681 CGAACGTGGC GAGAAAGGAA GGAAGAAAAG CGAAAGGAGC GGGCGCTAGG GCGCTGGCAA  
1741 GTGTAGCGGT CACGCTGCGC GTAACCACCA CACCCGCGC GCTTAATGCG CCGTACAGG  
1801 GCGCGTCGCG CCATTCGCCA TTCAGGCTGC GCAACTGTTG GGAAGGGCGA TCGGTGCGGG  
1861 CCTCTCGCT ATTACGCCAG CTGGCGAAAAG GGGGATGTGC TGCAAGGCGA TTAAGTTGGG  
1921 TAACGCCAGG GTTTTCCCAG TCACGACGTT GTAAAACGAC GGCCAGTGAG CCGCGTAAAT  
1981 ACGACTCACT ATAGGGCGAA TTGGGTACCG GCCGCAAAAT AAAGCCTTCG ACGGTCCCAA  
2041 AACCTTCTCA AGCAAGGTTT TCAGTATAAT GTTACATGCC TACACGCGTC TGTACAGAAA  
2101 AAAAAAGAAA ATTTGAAAATA TAAATAACGT TCTTAATACT AACATAACTA TAAAAAATA  
2161 AATAGGGACC TAGACTTCAG GTTGTCTAAC TCCTTCCTTT TCGGTTAGAG CGGATGTGGG  
2221 GGGAGGGCGT GAATGTAAGC GTGACATAAC TAATTACATG ACTCGAGTTA CTTTCTAGTA  
2281 CGGGCAGCCT TCTTTGGTGG CAGAATTAGA GATTGAGTT CATCCAATGC TTCACCAGTA  
2341 TAGACTGCTA ACTTTTGCAT TGATGGTAAA GTATCACGAC AACCGGTAAA CCCAAAGTTA  
2401 AGGGAGCCAG CGTAAGACAG GCATGTGATG TTCAAGGCAC CACCGTGGG GATAAGGGAA  
2461 ACTGGGTACA TGGCTTCCAA TCTAGCTCCT TCGTAATACA ATGTCCCCTC AGGCCAGGG  
2521 ACATTGCTTA TAGTAACATT GAAAACCTGGC CTCATTCTAC CTCCAATCC GCTCATCAAT  
2581 TGTAAGATGT AAGGTGACAT CAGCAGCATT GTGTATTGTG TTAGTGCTGA CTTAGGCAAC  
2641 TTTTGGAGAT GCTCCTTCGC CCTTCGTGTT GAGGTTTTGA TCTGTGTAA TCTATTGAGA  
2701 GGATCAGCTT CGTCAGTAGC CAAACTAGCA ATCATGAATG AGATTGTGT TCCGGTGCCT  
2761 TCGTCATCAG CTGGTCTGAT ATTCACTGGA ATTCCTGCCG TTAAAGGTGT GTCAGGCAAG  
2821 TTATTCTGCT CAGCCAAAAA CCTTCTCAAG GCAGTACCAC ATAAGTAAAG CACTATATCA  
2881 TTCAGTGAAC CACCCTGGC ATGCGCTAAG TTTTCAATC TATCTAGTTG GTAGTGTGA  
2941 GTAGCGAATC TTCTCTGCGC TGTAACCTA TGGTTAAGGA CAGACACTGG ACCGGTGAAT  
3001 GGAGCTGTTA AGCCATCCTC AGGGTGTCTG ACAGAAATGTA CTAGACGATT GCCAGCTTGC  
3061 CATAGTCTAG GGGCCATATC CGCTTGGAGC TTCAAGGCGT CCATTGCTTG AGAAACCGCA  
3121 GCAGGCACAG AAGCCTCTTT ATCTGTTTTA GCTCCTCTTC TCTGGTGAGG TCTAACTGTC  
3181 CATGGAGGAG GCATGTTACA TCTCTCTGGG TCAGTGGTCA ATACTCTTTG CATCAATCTT  
3241 ACTCCAGATA TCCCATCAAT CATAGAGTGA TGCATCTTAG TATACAGTGC AAATCTGTTG  
3301 TTTTCAAGAC CTTCAATGAC ATGGCATTCC CATAGTGGTC TAGAGAAATC TAAAGGGTTG  
3361 GAGTGTAACT TAGAAAACAAG TATCCCTAGC TCTCTTTCAC CACCTGGTCT TGGCAATGCA  
3421 CTATGTCCGA CATGATAACT CAAATCGATA TCTTTGCTA CTTCCAGGC AGGAAATGTA  
3481 ACCCTGCCAA GGAAACCGGA CCTATGCGAGC TTGTAACCCC ATGGTGGGGC AACATCTCCT  
3541 GCCTCTTTC TCTTGTAAAC CATATCACGT AGAAATGTCT CTGGAGCACC TTCTGGTAAT  
3601 GAGAAGATTT GGAGAGTGCC CACGTGCATT GCGGTATCTT CGGATTGAC TGCAAGCCAA  
3661 CTAGCGTCTA GAGTACCTAA TCTCTTCATC TGATAATGGG TTTGTAGTTT ATAATTATGT

3721 GATGAGGCCA AAAACAACG GGATTAGTT ATCCTTAGTA TGGTAGCCGA ATGAATGAAA  
3781 TATGTGGAAT AGATTCTCT TTTGCAAAAC GGGATATTTA AGCCAGGAAA AAGAACAGGC  
3841 TTATGACAGA ATTATGCATA TGAACGAGA GAAAGAACAC GTCCAGCAA TATGGAGTAC  
3901 CTTGTTTGCC TTTAGAAAAT GTTTCGGTT CAATACTGTT CTCTTCAGCG ATGCTGTTGA  
3961 AGTATCAGCT CAGTTATTTA ATATGCTAAC CTTCCCTTG TTTGTTTTCT TTTTCTATCT  
4021 TTATTTTTTT TCCATTTACC GTTTTTGCG TTTCAGAAAA TGACCCTCAT TGATTATATG  
4081 AATAATAATA TACACGTTAT CATTTTTAT TTCTTTTATG TATATATAGT AAAAAATAAA  
4141 CATATCGAGA TCGGTGTATC TACATCAGAG CCGCTGAAGA CATAAAAAAC AAAAAAGCA  
4201 CCACCGACTC GGTGCCACTT TTTCAAGTTG ATAACGGACT AGCCTATTTT TAACTTGCTA  
4261 TTTCTAGCTC TAAAACGCTT TACGCTGGCT CCCTTAGATC ATTTATCTTT CACTGCGGAG  
4321 AAGTTTCGAA CGCCGAAACA TGGCACCAC CTTTCACCTC TACAGCGTTT GACCAAAATC  
4381 TTTTGAACAG AACATTGTAG GGTGTGAAA AATGCGCACC TTTACCGCTA GCCCAAGAGG  
4441 GCACCTACAA ATCTAGAGTT GTACTTCAA CGTACATGTA ATCACCCTGT ATATACTCGA  
4501 AAGAAAACAT CAAGTTTCTG TATAAATATG AGTGAAAGCA TAATCATACA TTATCTTTTC  
4561 AAAGAAGACT CCAGCTTTTG TTCCTTTAG TGAGGGTTAA TTGCGCGCTT GCGGTAATCA  
4621 TGGTCATAGC TGTTTCCTGT GTGAAATGT TATCCGCTCA CAATTCACA CAACATAGGA  
4681 GCCGGAAGCA TAAAGTGTA AGCCTGGGT GCCTAATGAG TGAGGTAACT CACATTAATT  
4741 GCGTTGCGCT CACTGCCCGC TTTCCAGTCG GAAACCTGT CGTGCCAGCT GCATTAATGA  
4801 ATCGGCCAAC GCGCGGGAG AGGCGGTTG CGTATTGGC GCTCTCCGC TTCCTCGCTC  
4861 ACTGACTCGC TCGCTCGGT CGTTCGGCTG CCGCGAGCG TATCAGCTCA CTCAAAGCG  
4921 GTAATACGGT TATCCACAGA ATCAGGGAT AACGCAGAA AGAACATGTG AGCAAAGGC  
4981 CAGCAAAGG CCAGGAACCG TAAAAGGCC GCGTTGCTGG CGTTTTCCA TAGGCTCCGC  
5041 CCCCCTGACG AGCATCAAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA  
5101 CTATAAAGAT ACCAGGCGTT TCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTCCGACC  
5161 CTGCCGCTTA CCGGATACCT GTCCGCCTT CTCCCTCGG GAAGCGTGGC GCTTTCTCAT  
5221 AGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG  
5281 CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATCG TCTTGAGTCC  
5341 AACCCGTAA GACACGACT ATGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA  
5401 GCGAGGTATG TAGGCGGTG TACAGAGTTC TTGAAGTGGT GGCCTAACTA CCGGTACACT  
5461 AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCG AAAAAAGGTT  
5521 GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTT TGTTTGCAAG  
5581 CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGG  
5641 TCTGACGCTC AGTGAACGA AACTCACGT TAAGGATTT TGGTCATGAG ATTATCAAAA  
5701 AGGATCTTCA CCTAGATCCT TTTAAATTA AAATGAAGTT TAAATCAAT CTAAGTATA  
5761 TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG  
5821 ATCTGTCTAT TTCGTTATC CATAGTTGCC TGACTCCCG TCGTGTAGAT AACTACGATA  
5881 CGGGAGGGCT TACCATCTGG CCCAGTGTG GCAATGATAC CGCGAGACC ACGCTCACCG  
5941 GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGAAGGG CCGAGCGCAG AAGTGGTCTT  
6001 GCAACCTTAT CCGCTCCAT CCAGTCTATT AATTGTTGCC GGAAGCTAG AGTAAGTAGT  
6061 TCGCCAGTTA ATAGTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTACGC  
6121 TCGTCTTTG GTATGGCTTC ATTCAGCTCC GGTCCCAAC GATCAAGCG AGTTACATGA  
6181 TCCCCATGT TGTGCAAAA AGCGTTAGC TCCTTCGGTC CTCCGATCGT TGTGAGAAGT  
6241 AAGTTGGCCG CAGTGTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTACTGTC  
6301 ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA  
6361 TAGTGTATGC GCGACCGAG TTGCTCTTGC CCGCGCTCAA TACGGGATAA TACCAGCCA  
6421 CATAGCAGAA CTTTAAAAGT GTCATCATT GAAAAACGTT CTTCGGGGCG AAAACTCTCA  
6481 AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT  
6541 TCAGCATCTT TTACTTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC  
6601 GCAAAAAGG GAATAAGGGC GACACGGAAA TGTGAATAC TCATACTCTT CCTTTTTCAA  
6661 TATTATTGAA GCATTTATCA GGGTATTGT CTCATGAGCG GATACATATT TGAATGTATT  
6721 TAGAAAAATA AACAAATAGG GGTCCGCGC ACATTTCCCC GAAAAGTGCC ACCTGGGTCC  
6781 TTTTCATCAC GTGCTATAAA AATAATTATA ATTTAAATTT TTTAATATAA ATATATAAAT  
6841 TAAAAATAGA AAGTAAAAA AGAAATTAAG GAAAAATAG TTTTGTGTTT CCGAAGATGT  
6901 AAAAGACTCT AGGGGATCG CCAACAAATA CTACCTTTTA TCTTGCTCTT CCTGCTCTCA

6961 GGTATTAATG CCGAATTGTT TCATCTTGTC TGTGTAGAAG ACCACACACG AAAATCCTGT  
7021 GATTTTACAT TTTACTTATC GTTAATCGAA TGTATATCTA TTTAATCTGC TTTTCTTGTC  
7081 TAATAAATAT ATATGTAAAG TACGCTTTT GTTGAATTT TTTAAACCTT TGTTTATTTT  
7141 TTTTCTTCA TTCCGTAACCT CTCTACCTT CTTTATTAC TTTCTAAAAT CCAAATACAA  
7201 AACATAAAAA TAAATAAACA CAGAGTAAAT TCCCAAATTA TTCCATCATT AAAAGATACG  
7261 AGGCGCGTGT AAGTTACAGG CAAGCGATCC GTCCTAAGAA ACCATTATAT CATGACATTA  
7321 ACCTATAAAA ATAGGCGTAT CACGAGGCC TTTTCGTC