

Review

Genomic Considerations for the Modification of *Saccharomyces cerevisiae* for Biofuel and Metabolite Biosynthesis

James T. Arnone

Department of Biology, William Paterson University, Wayne, NJ 07470, USA; arnonej@wpunj.edu;
Tel.: +973-720-3457

Received: 17 January 2020; Accepted: 24 February 2020; Published: 26 February 2020



Abstract: The growing global population and developing world has put a strain on non-renewable natural resources, such as fuels. The shift to renewable sources will, thus, help meet demands, often through the modification of existing biosynthetic pathways or the introduction of novel pathways into non-native species. There are several useful biosynthetic pathways endogenous to organisms that are not conducive for the scale-up necessary for industrial use. The use of genetic and synthetic biological approaches to engineer these pathways in non-native organisms can help ameliorate these challenges. The budding yeast *Saccharomyces cerevisiae* offers several advantages for genetic engineering for this purpose due to its widespread use as a model system studied by many researchers. The focus of this review is to present a primer on understanding genomic considerations prior to genetic modification and manipulation of *S. cerevisiae*. The choice of a site for genetic manipulation can have broad implications on transcription throughout a region and this review will present the current understanding of position effects on transcription.

Keywords: *Saccharomyces cerevisiae*; gene expression; genetic engineering; neighboring gene effects; adjacent gene co-regulation; spatial positioning; position effects

1. Background and Introduction

As the global population grows and countries continue to develop, the demand for fuel continues to rise. Traditional fuel sources are finite and contribute to climate change, stressing the importance of renewable sources of energy. This shift offers many benefits, including reductions in greenhouse gasses and improved energy security. Many countries use field crops such as corn, wheat, barley, sugar cane, and cassava to produce ethanol; however, their use as a biomass source competes with their use as food crops, increasing demands for land and water resources [1,2]. Although this strain is of a lesser concern on a global level, it can have substantial regional impacts, pushing for the exploration of other viable options [1].

The native metabolic pathways in diverse microbes, including wild yeasts and bacteria, are particularly attractive for use in the production of biofuels via fermentation, which can generate bioethanol from a variety of non-food crop substrates, such as glycerol [3,4]. The metabolic flexibility of fungi allows for the generation of ethanol from materials (including olive mill wastewater) and the production of microbial enzymes and lipids from milling and confectioners' wastes [3,5,6]. Screening of diverse fungal isolates has proven effective in expanding the available repertoire of biofuels and metabolites produced from myriad starting materials [7]. The ability to reconstitute these pathways into an easier-to-culture, non-native host can increase the yield and efficiency of metabolite production. Advances in genetic engineering enhance the production of biofuels (and other metabolites), allowing synthesis in non-native species and manipulations to maximize the desired output.

A number of microbes, including *Escherichia coli*, *Clostridium acetobutylicum*, *Streptomyces venezuelae*, and *Saccharomyces cerevisiae* have proven successful in the introduction of non-native pathways in each organism [8–10]. Each organism has its own advantages (and disadvantages); however, the budding yeast, *S. cerevisiae*, is particularly amenable to genetic manipulation in the production of many metabolites. The use of *S. cerevisiae* as a model system is widespread and has the metabolic flexibility to ensure myriad biofuel outputs, including the production of fatty acid ethyl esters, butanol isomers, fatty alkanes, and fatty and higher alcohols [11,12]. Additionally, *S. cerevisiae* has a genome that can easily be manipulated, which is especially important, as optimization for the production of a particular biofuel output can require multiple genetic changes to eliminate competing pathways and maximize synthetic and native pathways to maximize output [13–15]. As a case in point, the production of isobutanol is optimized by the concurrent deletion of alcohol dehydrogenase and pyruvate dehydrogenase genes and the introduction of transhydrogenase-like shunts [16].

As *S. cerevisiae* has been extensively utilized as a model organism, myriad genetic tools have been developed and optimized for genetic manipulation; they include homologous recombination mediated in vivo mutagenesis, DNA assembly cloning to construct biochemical pathways, sequence and ligation-independent cloning, Gibson assembly cloning, and CRISPR-Cas9 genetic editing [17–23]. This yeast is highly versatile, with stable haploid and diploid life cycles, and can tolerate multiple simultaneous genetic changes [11]. Recent advances in CRISPR technology has allowed for the simultaneous introduction of up to six genetic modifications in a single transformation step [24]. *S. cerevisiae* can also be modified for the utilization of non-traditional nutrient sources, including cyanamide and phosphite for nitrogen and phosphorus, respectively, which can prevent the growth of contaminating microbes [25]. This combination of available cloning techniques, combined with the tolerance for significant genetic manipulation, makes this microorganism ideal for the introduction and optimization of non-native pathways for use in biofuel and metabolite production.

Complete harnessing of the potential of *S. cerevisiae* requires a thorough understanding of the many mechanisms that are available to optimize the desired metabolite output. The use of plasmids for genetic engineering purposes, while easier, is much less stable than chromosomal manipulations, as plasmids require consistent selection for maintenance [26,27]. Chromosomal manipulations are much more stable, though there can be unintended second order effects that alter local transcription and can lead to unexpected complications. There are excellent reviews and resources that allow for a thorough understanding of gene expression and regulation in *S. cerevisiae* via *cis* regulatory DNA sequences and transcription factor interactions, global nucleosome positions, mRNA transcript lengths and signals, mRNA stability, decay, and their consequences on cell cycle, protein stability, and turnover [28–35]. The focus of this review is to provide an understanding of position effects and spatial positioning, the physical arrangement of genes along the chromosome in *S. cerevisiae*, and potential complications that can arise during chromosomal manipulations introduced by genetic engineering.

2. Heterochromatic Genomic Regions Silence Transcriptional Activity

Heterochromatin forms in transcriptionally repressed, typically gene poor regions of the genome. *S. cerevisiae* maintains heterochromatic regions at the telomere and sub-telomeric regions, the silent mating loci (*HML* and *HMR*) on chromosome III, and within the rDNA repeats found on chromosome XII [36–38].

2.1. Telomeric Heterochromatin

The very distal tips of eukaryotic chromosomes are telomeres, specialized genomic regions maintained in unique structures to prevent inappropriate homologous recombination or triggering of the DNA damage response pathways [39]. Telomeres consist of simple, non-coding repeats (the sequence is $C_{1-3}A/TG_{1-3}$) that stretch for approximately 300 base-pairs \pm 75 base-pairs per chromosome [40]. Adjacent to these telomeric repeats are telomere associated sequences (TAS), which fall into two classes: X and Y'. X TAS sequences are found on all telomeres and Y' TAS sequences

are found on approximately half of the telomeres, which can vary by strain. There is significant diversity seen within each TAS class—in terms of size, composition, and insertions and deletions that can vary significantly from telomere to telomere [40].

Initial characterization has focused on individual telomeres, with an emerging consensus—telomeres exclude nucleosomes and are maintained in a transcriptionally silent, heterochromatic state. The integration of a transgene into telomeric regions in *S. cerevisiae* results in transcriptional repression of the transgene and is commonly known as the ‘telomere position effect’ (TPE) [38]. Evidence for this silencing focuses on the repression of *URA3* integrated into telomere *TEL07L*, which is mediated by the spread of heterochromatin nucleated at the telomere across the transgene by the Sirtuin family proteins and Rap1p [41,42]. The ubiquitous nature of the TPE has recently come into question, as global analysis utilizing the more sensitive RNA-sequencing methods identified active transcription at many native telomeres, albeit at levels that are lower than integration sites further from the telomere. Furthermore, this study shows that while the Sirtuin proteins are localized throughout the telomeres, their activity is not the primary mechanism of transcriptional repression, as only 6% of subtelomeric genes are silenced by Sir protein activity [43]. The current understanding of transcription within these genomic regions is more nuanced—under certain growth conditions (e.g., the induction of the environmental stress response, or ESR), there is transcriptional activation of telomeric genes [44]. This last point is germane, as the culturing conditions required for biofuel and metabolite production can trigger the budding yeast’s ESR, ultimately activating transcription within these regions. Indeed, our observations found that members of the stress-induced toxin-response gene family cluster in telomeric and sub-telomeric regions, playing a role in coordinating the transcription of this regulon [45,46].

2.2. Ribosomal DNA and the Silent Mating Loci

In addition to the heterochromatin that forms on telomeric regions, there are two additional sites where constitutive heterochromatin forms, at the silent mating loci, called *HML* and *HMR*, found on chromosome III, and at the rDNA repeats on chromosome XII. While there are similarities between the heterochromatin found in each region, there are also important differences compared to telomeric chromatin. Many of the chromatin modification proteins required for heterochromatin formation and maintenance are shared at all three heterochromatin loci, and each exhibit a lack of post-translational modifications to the tails [47,48]. One major difference is that transcriptional silencing at both *HML* and *HMR* is stronger but only extends over a shorter distance than that at the telomeres. Silencing at *HML* and *HMR* is nucleated by *cis* regulatory elements that flank each locus (*HML-E* and *HML-I* and *HMR-E* and *HMR-I*, respectively); however, the presence of insulators limit the genomic distance of silencing [49]. The rDNA repeats sequester to the nucleolus, and the number of active versus repressed regions varies, matched to the cell cycle and the demand for ribosomes. The transcriptional activation during stress does not affect the *HML* mating locus or the rDNA repeats [44].

The formation of heterochromatin facilitates higher order chromosome structures and the maintenance of subnuclear organization. Telomeres terminate into loops that fold back upon themselves, clustering together on the nuclear periphery into distinct puncta, as seen in the Rab1 nucleus [40,48,50]. The silencers flanking *HMR* form a local loop that facilitates transcriptional repression, while *HMR* and *HML* form an extended loop whereby Chromosome III folds across itself forming a long range interaction—all together, these interactions facilitate the characteristic subnuclear arrangement necessary for proper gene expression and genomic integrity [48,51].

3. Global Position Effects Result in Large Differences in Reporter and Transgene Expression

Saccharomyces cerevisiae was the first eukaryotic organism with a fully sequenced genome [52]. While estimates vary, there is consensus that there are just under 6000 protein coding genes, encoded by approximately 12 megabases of DNA divided into 16 chromosomes [52,53]. The completion of the genome sequence allowed researchers to construct the ‘Yeast Deletion Library’ (also known as

the ‘Yeast Knock Out Collection’, or YKO collection), a series of isogenic yeast strains, each with a single non-essential gene deleted [54,55]. The library consists of 5,916 isogenic yeast strains with the kanamycin resistance (KAN^R) gene replacing an individual open reading frame [55]. This resource allowed for rapid advances in the understanding of gene function on a level as never before, and led to further development of additional library collections, including the green-fluorescent protein (GFP) and tandem-affinity purification (TAP) tagged resources [56,57].

3.1. The Neighboring Gene Effect

The YKO collection expanded the understanding of gene function, significantly increasing the number of annotated gene functions. The design of the YKO collection integrated a KAN^R under the regulation of a relatively high strength promoter. Systematic characterization of genetic screens that used the deletion library led to a surprising result: the KAN^R integration site frequently disrupted the expression of a neighboring, adjacent gene [56–59]. The disruption caused by KAN^R to the expression of the adjacent gene was coined the ‘neighboring gene effect’ (NGE), which is estimated to affect transcription at 7%–15% of the targeted loci [58]. Follow-up analysis revealed that this effect results in miss-annotation of genetic interactions in up to 10% of screen results, rather than the identification of a bona fide genetic link (Table 1) [59].

3.2. Global Position Effect Variance

The completion of the sequencing of the yeast genome allowed for further characterization of transcriptional differences that arise when other, identical constructs are integrated into different genomic regions. These studies characterized the differences in expression of an insertion of a construct directly, as opposed to the disruption of the transcription of neighboring genes. Two studies investigated the integration of a green fluorescent protein (GFP) reporter and a red fluorescent reporter (RFP) throughout the genome and monitored position effects on their expression. When a GFP reporter was integrated at 482 genomic sites, the noise in the transcription of this reporter varied by 20-fold, resulting in a 15-fold difference at the level of protein production [60]. A global survey characterized the position effects observed when the red fluorescent protein (RFP) gene was integrated at 1044 different genomic positions. Based on location, there was a 13-fold difference in fluorescence seen between difference loci [61]. Additional analysis of KAN^R expression across the yeast deletion library revealed significant differences in the level of KAN^R expression from locus to locus, with up to 35% of expressional differences attributed to position effects (Table 1) [62]. The authors note that insertion of the KAN^R cassette does not disrupt the local chromatin environment [62].

Table 1. Summary of Position Effects on Gene Expression Characterized in *S. cerevisiae*.

Construct/Analysis	Position Effect Observed	Ref.
<i>pTEF1-LacZ</i> and <i>pACT1-LacZ</i>	8-fold difference in expression across 20 integration sites.	[63]
<i>pPRS3-LacZ</i>	14-fold difference in expression across 18 integration sites.	[64]
<i>pTEF-KAN^R</i>	4-fold variation in expression across chromosome I.	[62]
<i>pTEF-KAN^R</i> in yeast knock out collect	Characterized the ‘Neighboring Gene Effect’. Transcription disruption by the KAN^R cassette resulted in missannotation of gene function in genetic screens (parenthesis represents frequency of missannotation): telomere length mutant screen (24.1%), response to rapamycin screen (21.5%), topoisomerase mutant sensitivity (15.7%), and 5-fluorouracil sensitivity (7.2%).	[58]
Double deletion mutant screen	‘Neighboring Gene Effect’ resulted in 18% global missannotation of a systematic synthetic double deletion screen.	[59]
<i>pRPL5-GFP</i>	Characterized effects from 63 loci on chromosome I and 482 total sites globally. 22.6% of integration sites exhibited significant expressional differences, 36.5% demonstrated significant transcriptional noise, and there was 15-fold deviation in overall expression levels.	[60]
<i>pTEF1-RFP</i> and <i>pURA3-RFP</i>	13-fold difference in expression across 1044 integration sites. Different promoters resulted in different overall levels of expression; however, position affected both constructs to the same degree.	[61]

4. Adjacent Gene Co-Regulation and Functional Clustering

4.1. Adjacent Gene Co-Regulation

The production of mature, translationally competent ribosomes requires the coordinated gene expression of two distinct families of protein coding genes: the 129 ribosomal proteins (RPs) and the more than 200 rRNA and ribosome biosynthesis genes (*RRB* or *Ribi*), consuming significant amounts of intracellular energy stores and under tight regulation, predominantly at the level of transcription [65,66]. Each gene family is distinct—the RPs assemble and are incorporated into the ribosome during synthesis and remain there throughout the ribosomes' life, while the RRBs facilitate the synthesis of the ribosome but do not remain associated after the ribosome is mature and translationally competent. The different roles result in different levels of expression for each gene family. Each gene family is tightly co-regulated, enriched for distinct promoter sequences, and bind to separate, specific transcription factors [67–70]. Characterization of the genomic distribution of both families revealed that each family is found in a non-random distribution, predominantly clustered in pairs with other members of the same family [67,71].

The best-characterized locus is the RRB gene pair, *MPP10-MRX12*, found on Chromosome X. *MPP10*, a component of the small subunit processome and 90S preribosome complex, and *MRX12*, a protein that associates with mitochondrial ribosome, are clustered together under the transcriptional co-regulation of shared promoter elements only found upstream of *MPP10* (Figure 1A). These *cis*-regulatory promoter elements are called the ribosomal RNA processing element (RRPE) and the polymerase A and C (PAC) element, both of which are enriched within the RRB family as a whole [67]. The promoter region of *MRX12* contains no readily identifiable transcription factor-binding site; rather, the genome region immediately upstream is a nucleosome devoid autonomously replicating sequence (Figure 1) [72].

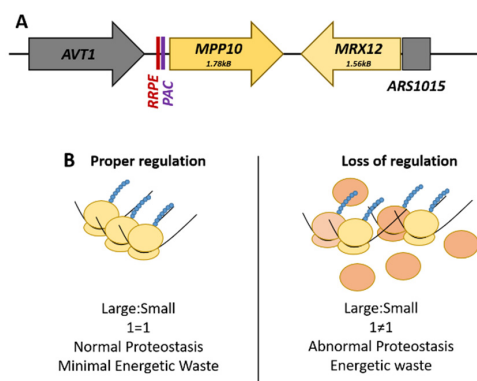


Figure 1. The *MPP10-MRX12* locus on Chromosome X. (A) The spatial arrangement of the ribosome biogenesis gene pair, *MPP10-MRX12*, highlighting the promoter regulatory elements. (B) Comparison of the potential effects that can arise from a loss of the proper transcriptional regulation of this gene pair. The loss of stoichiometric levels of these genes necessary for ribosome production is wasteful and could potentially alter cellular proteostasis.

Functional dissection of the RRPE and PAC promoter elements in *MPP10* by mutational analysis and gene expression profiling found that these sites are necessary for the transcriptional co-regulation of this RRB gene pair with the rest of the regulon during cellular growth and the induction of the stress response [71]. There is a physical requirement of adjacency for *MRX12* to *MPP10* for their co-regulation, as insertion of an actively transcribed *LEU2* gene effectively uncouples the pair from each other; this phenomenon is termed ‘adjacent gene co-regulation’ [73]. This surprising relationship demonstrates the distance that endogenous yeast promoters can potentially act with—as transcription of *MRX12* initiates across a chromosomal distance of almost four kilobases. The advantage of this arrangement is that it can potentially help buffer the expression of components needed at roughly

stoichiometric levels. The inability to do this is wasteful, potentially resulting in alterations to pools of ribosomal subunits, and ultimately to abnormal proteostasis (Figure 1B).

4.2. Functional Clustering

A systematic survey in budding yeast revealed that there are many functionally related gene families found in a non-random genomic distribution—this phenomenon is not limited to genes involved in the synthesis of the ribosome. Approximately, one-quarter of functionally related gene families in *S. cerevisiae* exhibit clustering throughout the genome. This arrangement results in tighter transcriptional co-regulation throughout the cell cycle, compared to their singleton (unpaired) counterparts [74]. It is likely that there are many different drivers that lead to functional clustering in such widely diverse gene families; however, the functionally clustered genes aggregate in areas of the genome where neighboring genes influence transcription of their neighbors [45].

Proper gene regulation via genomic clustering can also involve transcriptional interference, which results in an anti-correlated expression of neighboring genes due to mutually exclusive expression. The canonical example of this type of transcriptional regulation is seen at the *SRG1-SER3* genomic locus found on chromosome V. *SRG1* is a non-coding RNA that prevents the transcription of *SER3*—which can only be actively transcribed when the *SRG1* transcript is repressed—as the transcript extends over the transcription start site for *SER3*, altering the underlying chromatin [75,76].

The co-regulation of clustered genes can help to prevent the accumulation of molecules that may be potentially toxic [74,77]. The *GAL* genes cluster on chromosome II—there are three genes necessary for galactose catabolism (Figure 2A). This locus consists of *GAL1*, a galactokinase that catalyzes the conversion of alpha-D-galactose to galactose-1-phosphate; *GAL10*, a UDP-glucose-4-epimerase; and *GAL7*, a galactose-1-phosphate uridyl transferase. The genomic arrangement of these genes results in buffering of the cell from transcriptional noise, maintaining comparable levels of expression, which would not be possible if the genes were unlinked. The balance in the expression of *GAL1* to that of *GAL7* avoids the accumulation of the cytotoxic galactose-1-phosphate metabolite (Figure 2B) [78].

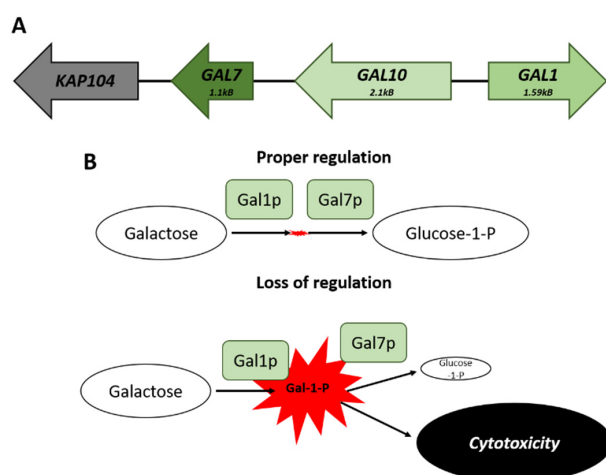


Figure 2. The *GAL* locus on Chromosome II. (A) Spatial arrangement of the genes that comprise the *GAL* locus. (B) Comparison of the potential effects that can arise from a loss of proper transcriptional regulation of the cluster. The failure to maintain appropriate levels of *GAL7* matched to *GAL1* leads to a buildup of Galactose-1-Phosphate and subsequent cytotoxic effects.

5. Lessons and Conclusions

Saccharomyces cerevisiae offers significant advantages for the introduction of non-native biological pathways, compared to some of its less-complex, prokaryotic peers. It also presents additional levels of complexity and challenges that require consideration and planning to ameliorate. The genomic site for modification must be chosen with care, as there is the ability to introduce unintended consequences—the

second order effects. The targeted integration site can profoundly affect the levels of expression of any gene—and vice versa—potentially altering delicate cellular processes and leading to a lower metabolic output.

Global analyses of position effects reveal that centromeric and telomeric regions, a distance defined as ± 20 kb from the repeat sequence, result in reduced gene expression compared to integration sites positioned further away [61]. These regions, however, are not completely silent in all cases. The researcher looking for a site to target for lower levels of gene expression may very well consider these regions. Genomic regions further away represent target sites that are more conducive for higher levels of gene expression, although position effects only result in high levels of expression in 25% of the genome. Hot spots that support high levels of expression include the *YDR448W*, *YGR240C*, *YHR142W*, *YML059C*, *YPL014W*, and *YPR028W* loci, and represent excellent starting points for researchers who want to synthesize highly expressed pathways [61]. One caveat is the effect of any non-native metabolite on the budding yeast. *S. cerevisiae* has a robust stress response, which triggers easily [46,79]. Such a response could result in the increase in transcription of regions that are typically low expressed and compound matters.

Our research group previously performed an analysis of the position effects on endogenous gene expression that arose due to the budding yeast's ESR [45]. Analysis focused on the induction of the environmental stress response, including a heat-shock response, DNA damage response, oxidative stress, nitrogen depletion, and the switch in carbon source from glucose to glycerol. As a resource for those concerned about induction of the ESR, the data from 30 representative loci have been parsed from that dataset and is presented in Table 2.

Table 2. The Spearman's correlation coefficient for 30 genomic regions during stress response and nutrient stressors.

Chromosome	Arm	Coordinates	SCC (10kB Window) *	SCC (10 Gene Window) **
II	R	415,500	0.0340	0.0130
IV	R	1301,500	-0.0500	-0.0280
IV	R	1451,200	-0.0160	-0.0112
IV	L	130,100	-0.0170	0.0030
IV	L	117,200	-0.0140	-0.0220
IV	L	217,200	0.0100	0.0330
IV	L	308,200	0.0240	0.0350
V	L	53,400	-0.0404	-0.0250
VII	L	254,000	-0.1820	-0.0750
VII	L	366,200	-0.0720	-0.0200
VII	R	649,200	-0.0390	-0.0220
VII	L	310,400	-0.0320	-0.0330
VII	R	270,600	0.0330	-0.0083
VIII	L	36,200	0.0690	0.0480
IX	L	316,400	-0.3000	0.0850
IX	L	99,600	0.0600	-0.0043
X	L	75,600	0.0644	0.1030
XI	L	431,500	0.0011	-0.0080
XII	R	932,000	-0.0500	0.0120
XII	R	805,500	-0.0190	0.0100
XII	R	282,600	0.0380	0.0620
XII	R	522,600	0.0720	-0.0280
XII	R	1028,500	0.3100	0.2470
XIII	R	754,500	0.0900	0.0300
XIV	L	330,800	-0.0470	-0.0680
XIV	L	65,200	-0.0190	-0.0300
XV	R	444,200	-0.0038	0.0211
XV	L	80,000	0.0053	-0.0610
XVI	L	173,000	0.0180	0.0290
XVI	L	135,600	0.1040	0.1170

* The average SCC for every pairwise combination of genes throughout a 10 kB window centered on the specified target region. ** The average SCC for every pairwise combination of genes throughout a 10-gene window centered on the specified target region (25kB–30kB region).

Due to its versatility and potential, *S. cerevisiae* will continue to be an attractive choice for genetic engineering for a variety of purposes, both inside and outside the biofuel sector. Yeasts are important sources of metabolites within the pharmaceutical and biotechnology industries, with *S. cerevisiae* widely used for these purposes, including for the production of metabolites such as dihydroartemisinic acid

(an important precursor molecule in the production of the anti-malarial agent artemisinin), and the production of beta-carotene [80–83]. Years of research and discovery on the microorganism has opened up incredible avenues and opportunities to harness this biological system as a mini factory. Furthermore, these observations are not simply limited to *S. cerevisiae*; they likely extend throughout the fungal kingdom and to related eukaryotes [74,84].

Funding: This research received no external funding.

Acknowledgments: The author would like to acknowledge the William Paterson University's Office of the Provost for Assigned Release Time for research and the Office of the Dean of the College of Science and Health and the Center for Research for their support of this work.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Fraiture, C.d.; Giordano, M.; Liao, Y. Biofuels and implications for agricultural water use: Blue impacts of green energy. *Water Policy* **2008**, *10*, 67–81. [[CrossRef](#)]
2. Sanchez, O.J.; Cardona, C.A. Trends in biotechnological production of fuel ethanol from different feedstocks. *Bioresour. Technol.* **2008**, *99*, 5270–5295. [[CrossRef](#)] [[PubMed](#)]
3. Sarris, D.; Papanikolaou, S. Biotechnological production of ethanol: Biochemistry, processes and technologies. *Eng. Life Sci.* **2016**, *16*, 307–329. [[CrossRef](#)]
4. Filippousi, R.; Antoniou, D.; Tryfinopoulou, P.; Nisiotou, A.A.; Nychas, G.J.; Koutinas, A.A.; Papanikolaou, S. Isolation, identification and screening of yeasts towards their ability to assimilate biodiesel-derived crude glycerol: Microbial production of polyols, endopolysaccharides and lipid. *J. Appl. Microbiol.* **2019**, *127*, 1080–1100. [[CrossRef](#)] [[PubMed](#)]
5. Dourou, M.; Kancelista, A.; Juszczyk, P.; Sarris, D.; Bellou, S.; Triantaphyllidou, I.-E.; Rywinska, A.; Papanikolaou, S.; Aggelis, G. Bioconversion of olive mill wastewater into high-added value products. *J. Clean. Prod.* **2016**, *139*, 957–969. [[CrossRef](#)]
6. Tsakona, S.; Skiadaresis, A.G.; Kopsahelis, N.; Chatzifragkou, A.; Papanikolaou, S.; Kookos, I.K.; Koutinas, A.A. Valorisation of side streams from wheat milling and confectionery industries for consolidated production and extraction of microbial lipids. *Food Chem.* **2016**, *198*, 85–92. [[CrossRef](#)] [[PubMed](#)]
7. Carsanba, E.; Papanikolaou, S.; Fickers, P.; Erten, H. Screening Various *Yarrowia lipolytica* Strains for Citric Acid Production. *Yeast* **2019**, *36*, 319–327. [[CrossRef](#)]
8. Atsumi, S.; Liao, J.C. Metabolic engineering for advanced biofuels production from *Escherichia coli*. *Curr. Opin. Biotechnol.* **2008**, *19*, 414–419. [[CrossRef](#)]
9. Lee, S.K.; Chou, H.; Ham, T.S.; Lee, T.S.; Keasling, J.D. Metabolic engineering of microorganisms for biofuels production: From bugs to synthetic biology to fuels. *Curr. Opin. Biotechnol.* **2008**, *19*, 556–563. [[CrossRef](#)]
10. Liu, Y.; Nielsen, J. Recent trends in metabolic engineering of microbial chemical factories. *Curr. Opin. Biotechnol.* **2019**, *60*, 188–197. [[CrossRef](#)]
11. Sherman, F. Getting started with yeast. *Methods Enzym.* **2002**, *350*, 3–41.
12. Duina, A.A.; Miller, M.E.; Keeney, J.B. Budding yeast for budding geneticists: A primer on the *Saccharomyces cerevisiae* model system. *Genetics* **2014**, *197*, 33–48. [[CrossRef](#)] [[PubMed](#)]
13. Blombach, B.; Eikmanns, B.J. Current knowledge on isobutanol production with *Escherichia coli*, *Bacillus subtilis* and *Corynebacterium glutamicum*. *Bioeng. Bugs* **2011**, *2*, 346–350. [[CrossRef](#)] [[PubMed](#)]
14. Atsumi, S.; Hanai, T.; Liao, J.C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **2008**, *451*, 86. [[CrossRef](#)]
15. Yamamoto, S.; Suda, M.; Niimi, S.; Inui, M.; Yukawa, H. Strain optimization for efficient isobutanol production using *Corynebacterium glutamicum* under oxygen deprivation. *Biotechnol. Bioeng.* **2013**, *110*, 2938–2948. [[CrossRef](#)]
16. Matsuda, F.; Ishii, J.; Kondo, T.; Ida, K.; Tezuka, H.; Kondo, A. Increased isobutanol production in *Saccharomyces cerevisiae* by eliminating competing pathways and resolving cofactor imbalance. *Microb. Cell Factories* **2013**, *12*, 119. [[CrossRef](#)]
17. Storici, F.; Lewis, L.K.; Resnick, M.A. In vivo site-directed mutagenesis using oligonucleotides. *Nat. Biotechnol.* **2001**, *19*, 773. [[CrossRef](#)]

18. Storici, F.; Resnick, M.A. The delitto perfetto approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. *Methods Enzymol.* **2006**, *409*, 329–345.
19. Shao, Z.; Zhao, H.; Zhao, H. DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res.* **2008**, *37*, e16. [[CrossRef](#)]
20. Shao, Z.; Zhao, H. Construction and engineering of large biochemical pathways via DNA assembler. *Synth. Biol.* **2013**, *1073*, 85–106.
21. Li, M.Z.; Elledge, S.J. SLIC: A method for sequence-and ligation-independent cloning. *Gene Synth.* **2012**, *852*, 51–59.
22. Gibson, D.G.; Young, L.; Chuang, R.-Y.; Venter, J.C.; Hutchison III, C.A.; Smith, H.O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **2009**, *6*, 343. [[CrossRef](#)] [[PubMed](#)]
23. DiCarlo, J.E.; Norville, J.E.; Mali, P.; Rios, X.; Aach, J.; Church, G.M. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* **2013**, *41*, 4336–4343. [[CrossRef](#)] [[PubMed](#)]
24. Mans, R.; Van Rossum, H.M.; Wijsman, M.; Backx, A.; Kuijpers, N.G.; Van den Broek, M.; Daran-Lapujade, P.; Pronk, J.T.; Van Maris, A.J.; Daran, J.-M.G. CRISPR/Cas9: A molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *Fems Yeast Res.* **2015**, *15*, fov004. [[CrossRef](#)] [[PubMed](#)]
25. Shaw, A.J.; Lam, F.H.; Hamilton, M.; Consiglio, A.; MacEwen, K.; Brevnova, E.E.; Greenhagen, E.; LaTouf, W.G.; South, C.R.; Van Dijken, H. Metabolic engineering of microbial competitive advantage for industrial fermentation processes. *Science* **2016**, *353*, 583–586. [[CrossRef](#)]
26. Zhang, Z.; Moo-Young, M.; Chisti, Y. Plasmid stability in recombinant *Saccharomyces cerevisiae*. *Biotechnol. Adv.* **1996**, *14*, 401–435. [[CrossRef](#)]
27. Da Silva, N.A.; Srikrishnan, S. Introduction and expression of genes for metabolic engineering applications in *Saccharomyces cerevisiae*. *Fems Yeast Res.* **2012**, *12*, 197–214. [[CrossRef](#)]
28. Schmid, M.; Jensen, T.H. Controlling nuclear RNA levels. *Nat. Rev. Genet.* **2018**, *19*, 518. [[CrossRef](#)]
29. Christiano, R.; Nagaraj, N.; Fröhlich, F.; Walther, T.C. Global proteome turnover analyses of the Yeasts *S. cerevisiae* and *S. pombe*. *Cell Rep.* **2014**, *9*, 1959–1965. [[CrossRef](#)]
30. Zhu, C.; Byers, K.J.; McCord, R.P.; Shi, Z.; Berger, M.F.; Newburger, D.E.; Saulrieta, K.; Smith, Z.; Shah, M.V.; Radhakrishnan, M. High-resolution DNA-binding specificity analysis of yeast transcription factors. *Genome Res.* **2009**, *19*, 556–566. [[CrossRef](#)]
31. De Boer, C.G.; Hughes, T.R. YeTFaSCo: A database of evaluated yeast transcription factor sequence specificities. *Nucleic Acids Res.* **2011**, *40*, D169–D179. [[CrossRef](#)] [[PubMed](#)]
32. De Boer, C.G.; Van Bakel, H.; Tsui, K.; Li, J.; Morris, Q.D.; Nislow, C.; Greenblatt, J.F.; Hughes, T.R. A unified model for yeast transcript definition. *Genome Res.* **2014**, *24*, 154–166. [[CrossRef](#)] [[PubMed](#)]
33. Jiang, C.; Pugh, B.F. A compiled and systematic reference map of nucleosome positions across the *Saccharomyces cerevisiae* genomes. *Genome Biol.* **2009**, *10*, R109. [[CrossRef](#)] [[PubMed](#)]
34. Hurowitz, E.H.; Brown, P.O. Genome-wide analysis of mRNA lengths in *Saccharomyces cerevisiae*. *Genome Biol.* **2003**, *5*, R2. [[CrossRef](#)]
35. Nagalakshmi, U.; Wang, Z.; Waern, K.; Shou, C.; Raha, D.; Gerstein, M.; Snyder, M. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **2008**, *320*, 1344–1349. [[CrossRef](#)]
36. Rine, J.; Herskowitz, I. Four genes responsible for a position effect on expression from HML and HMR in *Saccharomyces cerevisiae*. *Genetics* **1987**, *116*, 9–22.
37. Smith, J.S.; Boeke, J.D. An unusual form of transcriptional silencing in yeast ribosomal DNA. *Genes Dev.* **1997**, *11*, 241–254. [[CrossRef](#)]
38. Gottschling, D.E.; Aparicio, O.M.; Billington, B.L.; Zakian, V.A. Position effect at *S. cerevisiae* telomeres: Reversible repression of Pol II transcription. *Cell* **1990**, *63*, 751–762. [[CrossRef](#)]
39. Zakian, V.A. Structure, function, and replication of *Saccharomyces cerevisiae* telomeres. *Annu. Rev. Genet.* **1996**, *30*, 141–172. [[CrossRef](#)]
40. Wellinger, R.J.; Zakian, V.A. Everything you ever wanted to know about *Saccharomyces cerevisiae* telomeres: Beginning to end. *Genetics* **2012**, *191*, 1073–1105. [[CrossRef](#)]
41. Grunstein, M. Molecular model for telomeric heterochromatin in yeast. *Curr. Opin. Cell Biol.* **1997**, *9*, 383–387. [[CrossRef](#)]
42. Wright, J.H.; Gottschling, D.E.; Zakian, V.A. *Saccharomyces cerevisiae* telomeres assume a non-nucleosomal chromatin structure. *Genes Dev.* **1992**, *6*, 197–210. [[CrossRef](#)] [[PubMed](#)]

43. Ellahi, A.; Thurtle, D.M.; Rine, J. The chromatin and transcriptional landscape of native *Saccharomyces cerevisiae* telomeres and subtelomeric domains. *Genetics* **2015**, *200*, 505–521. [[CrossRef](#)] [[PubMed](#)]
44. Ai, W.; Bertram, P.G.; Tsang, C.K.; Chan, T.F.; Zheng, X.F. Regulation of subtelomeric silencing during stress response. *Mol Cell* **2002**, *10*, 1295–1305. [[CrossRef](#)]
45. Cera, A.; Holganza, M.K.; Hardan, A.A.; Gamarra, I.; Eldabagh, R.S.; Deschaine, M.; Elkamhawy, S.; Sisso, E.M.; Foley, J.J.; Arnone, J.T. Functionally Related Genes Cluster into Genomic Regions That Coordinate Transcription at a Distance in *Saccharomyces cerevisiae*. *mSphere* **2019**, *4*, e00063-19. [[CrossRef](#)]
46. Gasch, A.P.; Spellman, P.T.; Kao, C.M.; Carmel-Harel, O.; Eisen, M.B.; Storz, G.; Botstein, D.; Brown, P.O. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell.* **2000**, *11*, 4241–4257. [[CrossRef](#)]
47. Aparicio, O.M.; Billington, B.L.; Gottschling, D.E. Modifiers of position effect are shared between telomeric and silent mating-type loci in *S. cerevisiae*. *Cell* **1991**, *66*, 1279–1287. [[CrossRef](#)]
48. Gartenberg, M.R.; Smith, J.S. The nuts and bolts of transcriptionally silent chromatin in *Saccharomyces cerevisiae*. *Genetics* **2016**, *203*, 1563–1599. [[CrossRef](#)]
49. Haber, J.E. Mating-type genes and MAT switching in *Saccharomyces cerevisiae*. *Genetics* **2012**, *191*, 33–64. [[CrossRef](#)]
50. Schober, H.; Kalck, V.; Vega-Palas, M.A.; Van Houwe, G.; Sage, D.; Unser, M.; Gartenberg, M.R.; Gasser, S.M. Controlled exchange of chromosomal arms reveals principles driving telomere interactions in yeast. *Genome Res.* **2008**, *18*, 261–271. [[CrossRef](#)]
51. Valenzuela, L.; Dhillon, N.; Dubey, R.N.; Gartenberg, M.R.; Kamakaka, R.T. Long-range communication between the silencers of HMR. *Mol. Cell. Biol.* **2008**, *28*, 1924–1935. [[CrossRef](#)]
52. Goffeau, A.; Barrell, B.G.; Bussey, H.; Davis, R.; Dujon, B.; Feldmann, H.; Galibert, F.; Hoheisel, J.; Jacq, C.; Johnston, M. Life with 6000 genes. *Science* **1996**, *274*, 546–567. [[CrossRef](#)] [[PubMed](#)]
53. Engel, S.R.; Dietrich, F.S.; Fisk, D.G.; Binkley, G.; Balakrishnan, R.; Costanzo, M.C.; Dwight, S.S.; Hitz, B.C.; Karra, K.; Nash, R.S. The reference genome sequence of *Saccharomyces cerevisiae*: Then and now. *G3 GenesGenomesGenet.* **2014**, *4*, 389–398. [[CrossRef](#)] [[PubMed](#)]
54. Winzeler, E.A.; Shoemaker, D.D.; Astromoff, A.; Liang, H.; Anderson, K.; Andre, B.; Bangham, R.; Benito, R.; Boeke, J.D.; Bussey, H. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **1999**, *285*, 901–906. [[CrossRef](#)] [[PubMed](#)]
55. Giaever, G.; Chu, A.M.; Ni, L.; Connelly, C.; Riles, L.; Veronneau, S.; Dow, S.; Lucau-Danila, A.; Anderson, K.; Andre, B. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **2002**, *418*, 387. [[CrossRef](#)] [[PubMed](#)]
56. Krogan, N.J.; Cagney, G.; Yu, H.; Zhong, G.; Guo, X.; Ignatchenko, A.; Li, J.; Pu, S.; Datta, N.; Tikuisis, A.P. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* **2006**, *440*, 637. [[CrossRef](#)]
57. Ghaemmaghami, S.; Huh, W.-K.; Bower, K.; Howson, R.W.; Belle, A.; Dephoure, N.; O’Shea, E.K.; Weissman, J.S. Global analysis of protein expression in yeast. *Nature* **2003**, *425*, 737. [[CrossRef](#)]
58. Ben-Shitrit, T.; Yosef, N.; Shemesh, K.; Sharan, R.; Ruppim, E.; Kupiec, M. Systematic identification of gene annotation errors in the widely used yeast mutation collections. *Nat. Methods* **2012**, *9*, 373. [[CrossRef](#)]
59. Atias, N.; Kupiec, M.; Sharan, R. Systematic identification and correction of annotation errors in the genetic interaction map of *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **2015**, *44*, e50. [[CrossRef](#)]
60. Chen, X.; Zhang, J. The genomic landscape of position effects on protein expression level and noise in yeast. *Cell Syst.* **2016**, *2*, 347–354. [[CrossRef](#)]
61. Wu, X.-L.; Li, B.-Z.; Zhang, W.-Z.; Song, K.; Qi, H.; Dai, J.-b.; Yuan, Y.-J. Genome-wide landscape of position effects on heterogeneous gene expression in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **2017**, *10*, 189. [[CrossRef](#)] [[PubMed](#)]
62. Chen, M.; Licon, K.; Otsuka, R.; Pillus, L.; Ideker, T. Decoupling epigenetic and genetic effects through systematic analysis of gene position. *Cell Rep.* **2013**, *3*, 128–137. [[CrossRef](#)]
63. Flagfeldt, D.B.; Siewers, V.; Huang, L.; Nielsen, J. Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*. *Yeast* **2009**, *26*, 545–551. [[CrossRef](#)]
64. Thompson, A.; Gasson, M.J. Location effects of a reporter gene on expression levels and on native protein synthesis in *Lactococcus lactis* and *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **2001**, *67*, 3434–3439. [[CrossRef](#)] [[PubMed](#)]

65. Fromont-Racine, M.; Senger, B.; Saveanu, C.; Fasiolo, F. Ribosome assembly in eukaryotes. *Gene* **2003**, *313*, 17–42. [[CrossRef](#)]
66. Henras, A.K.; Plisson-Chastang, C.; O'Donohue, M.F.; Chakraborty, A.; Gleizes, P.E. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip. Rev. RNA* **2015**, *6*, 225–242. [[CrossRef](#)]
67. Wade, C.H.; Umbarger, M.A.; McAlear, M.A. The budding yeast rRNA and ribosome biosynthesis (RRB) regulon contains over 200 genes. *Yeast* **2006**, *23*, 293–306. [[CrossRef](#)]
68. Wade, C.; Shea, K.A.; Jensen, R.V.; McAlear, M.A. EBP2 is a member of the yeast RRB regulon, a transcriptionally coregulated set of genes that are required for ribosome and rRNA biosynthesis. *Mol. Cell. Biol.* **2001**, *21*, 8638–8650. [[CrossRef](#)]
69. Bosio, M.C.; Negri, R.; Dieci, G. Promoter architectures in the yeast ribosomal expression program. *Transcription* **2011**, *2*, 71–77. [[CrossRef](#)]
70. Bosio, M.C.; Fermi, B.; Dieci, G. Transcriptional control of yeast ribosome biogenesis: A multifaceted role for general regulatory factors. *Transcription* **2017**, *8*, 254–260. [[CrossRef](#)]
71. Arnone, J.T.; McAlear, M.A. Adjacent gene pairing plays a role in the coordinated expression of ribosome biogenesis genes MPP10 and YJR003C in *Saccharomyces cerevisiae*. *Eukaryot. Cell* **2011**, *10*, 43–53. [[CrossRef](#)] [[PubMed](#)]
72. Lee, W.; Tillo, D.; Bray, N.; Morse, R.H.; Davis, R.W.; Hughes, T.R.; Nislow, C. A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.* **2007**, *39*, 1235. [[CrossRef](#)] [[PubMed](#)]
73. Arnone, J.T.; Arace, J.R.; Soorneedi, A.R.; Citino, T.T.; Kamitaki, T.L.; McAlear, M.A. Dissecting the cis and trans elements that regulate adjacent-gene coregulation in *Saccharomyces cerevisiae*. *Eukaryot. Cell* **2014**, *13*, 738–748. [[CrossRef](#)] [[PubMed](#)]
74. Eldabagh, R.S.; Mejia, N.G.; Barrett, R.L.; Monzo, C.R.; So, M.K.; Foley, J.J.; Arnone, J.T. Systematic identification, characterization, and conservation of adjacent-gene coregulation in the budding yeast *Saccharomyces cerevisiae*. *mSphere* **2018**, *3*, e00220–18. [[CrossRef](#)]
75. Martens, J.A.; Laprade, L.; Winston, F. Intergenic transcription is required to repress the *Saccharomyces cerevisiae* SER3 gene. *Nature* **2004**, *429*, 571. [[CrossRef](#)]
76. Martens, J.A.; Wu, P.-Y.J.; Winston, F. Regulation of an intergenic transcript controls adjacent gene transcription in *Saccharomyces cerevisiae*. *Genes Dev.* **2005**, *19*, 2695–2704. [[CrossRef](#)]
77. McGary, K.L.; Slot, J.C.; Rokas, A. Physical linkage of metabolic genes in fungi is an adaptation against the accumulation of toxic intermediate compounds. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 11481–11486. [[CrossRef](#)]
78. Xu, H.; Liu, J.-J.; Liu, Z.; Li, Y.; Jin, Y.-S.; Zhang, J. Synchronization of stochastic expressions drives the clustering of functionally related genes. *Sci. Adv.* **2019**, *5*, eaax6525. [[CrossRef](#)]
79. Kurdistani, S.K.; Robyr, D.; Tavazoie, S.; Grunstein, M. Genome-wide binding map of the histone deacetylase Rpd3 in yeast. *Nat. Genet.* **2002**, *31*, 248–254. [[CrossRef](#)]
80. Li, Q.; Sun, Z.; Li, J.; Zhang, Y. Enhancing beta-carotene production in *Saccharomyces cerevisiae* by metabolic engineering. *Fems Microbiol. Lett.* **2013**, *345*, 94–101. [[CrossRef](#)]
81. Beekwilder, J.; Van Rossum, H.M.; Koopman, F.; Sonntag, F.; Buchhaupt, M.; Schrader, J.; Hall, R.D.; Bosch, D.; Pronk, J.T.; Van Maris, A.J. Polycistronic expression of a β -carotene biosynthetic pathway in *Saccharomyces cerevisiae* coupled to β -ionone production. *J. Biotechnol.* **2014**, *192*, 383–392. [[CrossRef](#)] [[PubMed](#)]
82. Guo, Y.; Dong, J.; Zhou, T.; Auxillos, J.; Li, T.; Zhang, W.; Wang, L.; Shen, Y.; Luo, Y.; Zheng, Y. YeastFab: The design and construction of standard biological parts for metabolic engineering in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **2015**, *43*, e88. [[CrossRef](#)] [[PubMed](#)]
83. Westfall, P.J.; Pitera, D.J.; Lenihan, J.R.; Eng, D.; Woolard, F.X.; Regentin, R.; Horning, T.; Tsuruta, H.; Melis, D.J.; Owens, A. Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E111–E118. [[CrossRef](#)]
84. Arnone, J.T.; Robbins-Pianka, A.; Arace, J.R.; Kass-Gergi, S.; McAlear, M.A. The adjacent positioning of co-regulated gene pairs is widely conserved across eukaryotes. *Bmc Genom.* **2012**, *13*, 546. [[CrossRef](#)] [[PubMed](#)]

