

---

Article

Not peer-reviewed version

---

# Genomic Analysis and Mechanisms Exploration of a Stress Tolerance and High-Yield Pullulan Producing Strain

---

Jing Yang , Wenru Wang , Ruihua Zhang , [Siqi Sun](#) , Biqi Li , Yue Shi , Junfeng Zeng , [Shulei Jia](#) \*

Posted Date: 5 July 2024

doi: 10.20944/preprints202407.0513.v1

Keywords: Pullulan and biomedicine; Secondary metabolites; Adaptive evolution; Comparative genomics



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

## Article

# Genomic Analysis and Mechanisms Exploration of a Stress Tolerance and High-Yield Pullulan Producing Strain

Jing Yang <sup>1</sup>, Wenru Wang <sup>1</sup>, Ruihua Zhang <sup>1</sup>, Siqi Sun <sup>2</sup>, Biqi Li <sup>1</sup>, Yue Shi <sup>1</sup>, Junfeng Zeng <sup>1</sup> and Shulei Jia <sup>3,\*</sup>

<sup>1</sup> School of Basic Medicine, Shanxi Medical University, Shanxi, 030606, China

<sup>2</sup> Department of Cardiology, Peking Union Medical College Hospital, Peking Union Medical College & Chinese Academy of Medical Sciences, Beijing, 100730, China

<sup>3</sup> School of Basic Medical Sciences, Tianjin Medical University, Tianjin, 300203, China.

\* Correspondence: jiashu320lei@126.com

**Abstract:** Pullulan is a kind of natural polymer, which is widely used in medicine and food because of its solubility, plasticity, edible, non-toxicity and good biocompatibility. It is of great significance to improve the yield of pullulan by genetic modification of microorganisms. It was previously reported that *Aureobasidium melanogenum* TN3-1 isolated from honey-comb could produce high-yield of pullulan, but the molecular mechanisms of its production of pullulan had not been completely solved. In this study, the reported strains of *Aureobasidium* spp. were further compared and analyzed at genome level. It was found that genome duplication and genome genetic variations might be the crucial factors for the high yield of pullulan and stress resistance. This particular phenotype may be the result of adaptive evolution, which can adapt to its environment through genetic variation and adaptive selection. In addition, the TN3-1 strain has a large genome, and the special regulatory sequences of its specific genes and promoters may ensure a unique characteristics. This study is a supplement of the previous studies, and provides basic data for the research of microbial genome modification in food and healthcare applications.

**Keywords:** Pullulan and biomedicine; secondary metabolites; adaptive evolution; comparative genomics

## 1. Introduction

Pullulan obtained by microbial fermentation is a kind of natural macromolecular polysaccharide, which has many biological activities and health care effects. It is a kind of  $\alpha$ -glucan with the  $\alpha$ -(1,6) and  $\alpha$ -(1,4) glucoside linkage alternating. It has been widely used in medicine, health food and other fields, and has many effects such as immune enhancement, anti-tumor, anti-aging, and blood lipid lowering [1]. For example, it can be used in the treatment of arthritis, inflammatory bowel disease and cancer. It can also be used as an auxiliary carrier of drugs to enhance the stability and bio-availability of drugs for the treatment of central nervous system diseases. Studies have shown that lulan nanoparticles can significantly reduce the expression level of oxidative stress markers in the hippocampus of mice models of neurodegenerative diseases, which can significantly improve the learning and memory ability of mice models of neurodegenerative diseases, and improve motor dysfunction [2]. Microbial polysaccharide capsules, mainly pullulan, have also gradually emerged in the field of biomedicine. Traditional capsules are made of gelatin, glycerin and water, but gelatin is not suitable for many types of drugs due to its shortcomings such as water loss hardening, water softening and cross-linking curing. The oxygen permeability of pullulan is 1/300 of hydroxypropyl methyl cellulose capsules and 1/8 of gelatin capsules, which can well protect the capsule contents from oxidation [3]. In addition, pullulan also has a good regulatory effect on the human immune system, which can enhance the immune function of the body, and improve the body's resistance.

*Aureobasidium* spp. (Taxonomy ID: 5579) is a genus of ascomycete fungi in the family Saccotheciacae. They are well known for their great biotechnology potentials [4–8]. The pullulan found so far was mainly synthesized by the *A. pullulans* and *Aureobasidium melanogenum* lineages [7]. As a promising microbial polysaccharide, how to increase the yield of pullulan by screening or improving microorganisms is an urgent problem to be solved. Studies showed that *Aureobasidium melanogenum* TN3-1 isolated from honey-comb can produce high levels of pullulan and can withstand high osmotic environment [5,6]. Compared with other strains of *Aureobasidium* spp., this strain has the stress tolerance and potential research values [5,9]. Isolates of *A. melanogenum* TN3-1 from hyperosmolar honey was able to produce high levels of pullulans from high concentrations of glucose due to hyperosmolar tolerance, while other strains such as *A. melanogenum* P16 and *A. melanogenum* CBS105.22 were not able to produce high levels of pullulans from high concentrations of glucose due to hyperosmolar intolerance. Especially, strain CBS105.22, isolating from granuloma, could not be able to produce any pullulan [5–9]. These different phenotypic differences may be closely related to their genomic variations during evolution. Although the differences and molecular mechanisms of high pullulan producing strains from *Aureobasidium* spp. have been analyzed at genome level [10,11], the mechanisms of high pullulan producing strains is still not fully resolved. Therefore, on basis of the available studies, we conducted a genomic comparison of *Aureobasidium* strains from different sources, and further analyzed the gene regulation and related signaling pathways from a new perspective, so as to provide supplementary for improving the pullulan production, and increase its applications in capsules, medicine or disease therapy.

## 2. Materials and Methods

### 2.1. Genome Download and Collation

Based on the existing literature [5–11], we downloaded the reported genomes of *Aureobasidium* spp. from the NCBI (<https://www.ncbi.nlm.nih.gov/assembly/>), and organized the phenotypic information of them. In addition, we downloaded other fungi genomes of Ascomycetes for comparison, including *Aspergillus niger*, *Saccharomyces cerevisiae*, *Yarrowia lipolytica*, *Schizosaccharomyces pombe* and so on. All the data were summarized and collated for further genomic analysis.

### 2.2. Gene Annotation and Functional Analysis

Sequence alignment was conducted between predicted genes and functional databases (BLASTP, E-value  $\leq 1e-5$ ), and the results with matching similarity  $\geq 40\%$  and coverage  $\geq 40\%$  were selected for annotation. The general functional databases annotated mainly included Swiss-Prot/TrEMBL, CAZy, GO, TCDB, KEGG and the NCBI non-redundant (NR) database. All known protein domains were searched by Pfam, ProDom, SMART database and InterProScan v7.0. Secondary metabolites biosynthesis gene clusters were annotated through antiSMASH v7.0 (<https://antismash.secondarymetabolites.org/>).

### 2.3. Gene Alignment and Genomic Collinearity Analysis

Protein structure and function analysis was conducted through the NCBI-BLAST online website (<https://www.ebi.ac.uk/Tools/ss/sncblast/>), and the TMHMM Server v.2.0 online website (<http://www.cbs.dtu.dk/services/TMHMM/>) was used to conduct membrane structure domain analysis. The collinearity analysis of *Aureobasidium* spp. was conducted through MCScanX toolkit (<https://github.com/wyp1125/MCScanX>) [12].

### 2.4. 3D Structure Prediction of the $\alpha$ -1,3-Glucan Synthase

We submitted the  $\alpha$ -1,3 glucan synthase from P16, TN3-1, CBS105.22, and the type strain CBS10374 to AlphaFold Server 3.0 (<https://alphafoldserver.com/>) [13]. The protein three-dimensional structures of crucial genes involved in the synthesis of pullulan were predicted and compared. The

range standard for amino acid residue confidence pLDDT is as follows: pLDDT > 90 means high confidence and highly reliable structure; PLDDT 70-90 means moderate confidence and relatively reliable structure; PLDDT of 50-70 means low confidence and less reliable structure; PLDDT < 50 means very low confidence, and the accuracy of the structure may be poor. PTM is a score that measures the accuracy of the entire protein folding, and the higher the score, the better the predicted folding quality. The template modeling score for interface prediction, ipTM score of 0.8 or higher, indicates high prediction quality, with gray areas ranging from 0.6 to 0.8. However, cases below 0.6 are questionable.

### 2.5. Pan-Genome and Genome Evolution

GET\_HOMOLOGUES version 07112023 [14] was used to search homologous genes of different *Aureobasidium* strains, including four reference species of *A. pullulans* with whole genome sequencing [11]. The single copy homologous genes of different strains were connected and combined respectively, then the protein sequence was compared with the MAFFT software (default parameters), the sequences were clipped with TrimAl, and the phylogenetic tree was constructed with MEGA v7.0.14 through the Neighbor-Joining (NJ) method. The Bootstrap value was set to 1000.

### 2.6. Transcriptional Factor Analysis for Crucial Genes

The upstream promoters of the alpha-1,3-glucan synthase encoding gene (AGS2) and the associated genes were searched through BioEdit v7.0.9.0 and DNAMAN, and then predicted in the YEASTRACT+ website (<http://www.yeastRACT.com/>) [15].

## 3. Results

### 3.1. Genome Analysis of *A. melanogenum* TN3-1

The downloaded genomes of *Aureobasidium* spp. are shown in Table 1, and the phenotypes of most strains have been validated [5–11]. According to studies, the strain TN3-1 was collected from honey-comb [9]. It can tolerate high osmotic pressure environment and produce pullulan, which is quite valuable for the microbial polyploidization research [5]. The results showed that the TN3-1 strain had a genome size of 51.6 Mb, GC content of 52.33%, and a total of 17,915 coding genes, most of which were 900-1,000 bp long with the average length of genes being 1,562 bp (Figure 1A) (Table 2). Meanwhile, TN3-1 strain has the highest number of specific genes (n=3,627), which may ensure its uniqueness (Figure 1B) (Table 2). Notably, both *A. melanogenum* TN3-1 and *A. melanogenum* CBS105.22 had occurred genome duplication, with almost every gene in the genome being duplicated [10]. In addition, the homologous sequence comparison of all the protein sequences of TN3-1 strain in the NCBI-NR (Non-Redundant) database showed that, most of the genes are homologous genes with *A. melanogenum* (9,646) and *A. pullulans* (4,388), respectively. Quite a few genes also have high homology relationships with *Aureobasidium* subglaciale (2,487) (Figure 1C), indicating that this strain may be a completely different strain from *Aureobasidium* spp.

**Table 1.** Different fungal genomes download based on the available researches.

Strains	Accessions
<i>Aureobasidium pullulans</i> P25	GCA_003574545.1
<i>Aureobasidium melanogenum</i> P16	GCA_019915885.1
<i>Aureobasidium melanogenum</i> TN3-1	GCA_017949655.1
<i>Aureobasidium melanogenum</i> CBS105.22	GCA_018290055.1
<i>Aureobasidium melanogenum</i> P5	GCA_025604605.1
<i>Aureobasidium pullulans</i> var. <i>namibiae</i> CBS147.97	GCF_000721765.1
<i>Aureobasidium pullulans</i> var. <i>subglaciale</i> EXF-2481	GCF_000721755.1
<i>Aureobasidium pullulans</i> var. <i>pullulans</i> EXF-150	GCF_000721785.1
<i>Aureobasidium melanogenum</i> CBS 110374	GCF_000721775.1

Aureobasidium sp. P6	GCA_003992365.1
Aureobasidium melanogenum HN6.2	GCA_002156615.1
Aureobasidium sp. SLJ-2021a	GCA_019677095.1
Aureobasidium zeae	GCA_017580825.2
Saccharomyces cerevisiae S288C	GCF_000146045.2
Schizosaccharomyces pombe 972h-	GCF_000002945.1
Yarrowia lipolytica CLIB122	GCF_000002525.2
Aspergillus niger	GCF_000002855.4
Blastomyces parvus UAMH130	GCA_002572885.1
Lepidopterella palustris CBS459.81	GCA_001692735.1



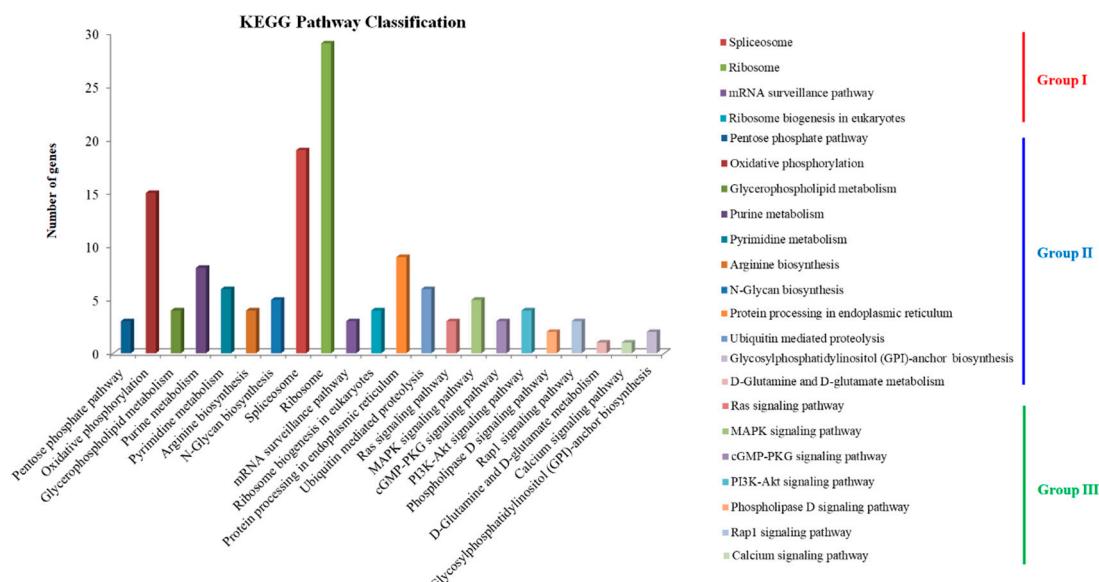
**Figure 1.** General genomic characteristics of the TN3-1 strain. A: Gene distribution characteristics of TN3-1; B: The common and unique genes of three reported *Aureobasidium* strains; C: Distribution of homologous genes in the TN3-1 strain.

**Table 2.** The general genomic characteristics of different *Aureobasidium* strains (BioProject: PRJNA661658).

Strains	Max length (bp)	N50 length (bp)	GC (%)	Genome size (bp)	Gene numbers	Gene length (bp)
P16	3,518,807	2,256,789	49.99	26,098,626	9,308	14,341,028
TN3-1	4,274,497	2,187,685	50.04	51,646,945	17,915	20,405,225
CBS105.22	4,273,107	2,075,361	49.99	48,644,395	17,817	27,348,966

Strains	Gene percentage (%)	Gene average length (bp)	% of Genome Specific (internal)	Specific genes	NRPS	T1PKS
P16	54.95	1,524	45.05	1,138	1	2
TN3-1	39.51	1,562	60.49	3,627	3	4
CBS105.22	56.22	1,547	43.78	2,499	1	5

The secondary metabolites annotation showed that TN3-1 strain contained 3 NRPS metabolic gene clusters (MGCs) and 4 type I PKS gene clusters (Table 2). Further analysis showed that, the PKS MGCs may be involved in the synthesis of secondary metabolites such as elsinochrome A/B/C, melanin, and naphthopyrone. Analysis of the specific genes showed that the specific genes of the high pullulan producing strain TN3-1 were 3,627, accounting for 28% of the whole genome (Figure 1B). KEGG annotation of these specific genes showed that they were involved in ~ 130 metabolic pathways, which could be divided into three groups, including the processing of genetic information such as DNA repair and recombination, RNA transport and ribosome generation; the metabolism related processes such as pentose phosphate pathway, respiratory chain oxidative phosphorylation, glycosylphosphatidylinositol (GPI) biosynthesis, D-glutamine metabolism; the metabolic pathway processes such as calcium ion signaling pathway, MAPK and RAS signaling pathway (Figure 2). The characteristics of specific genes in P16 strain was similar with that of TN3-1, but that in CBS105.22 strain did not show the Ras signaling pathway and the metabolic pathway related to glycosylphosphatidylinositol (GPI) biosynthesis. The MAPK signaling pathway within CBS105.22 strain was mainly associated with cell wall pressure response (K11244). However, in the pullulan producing strain P16 and TN3-1, the MAPK signaling pathway was mainly related to osmotic pressure response.



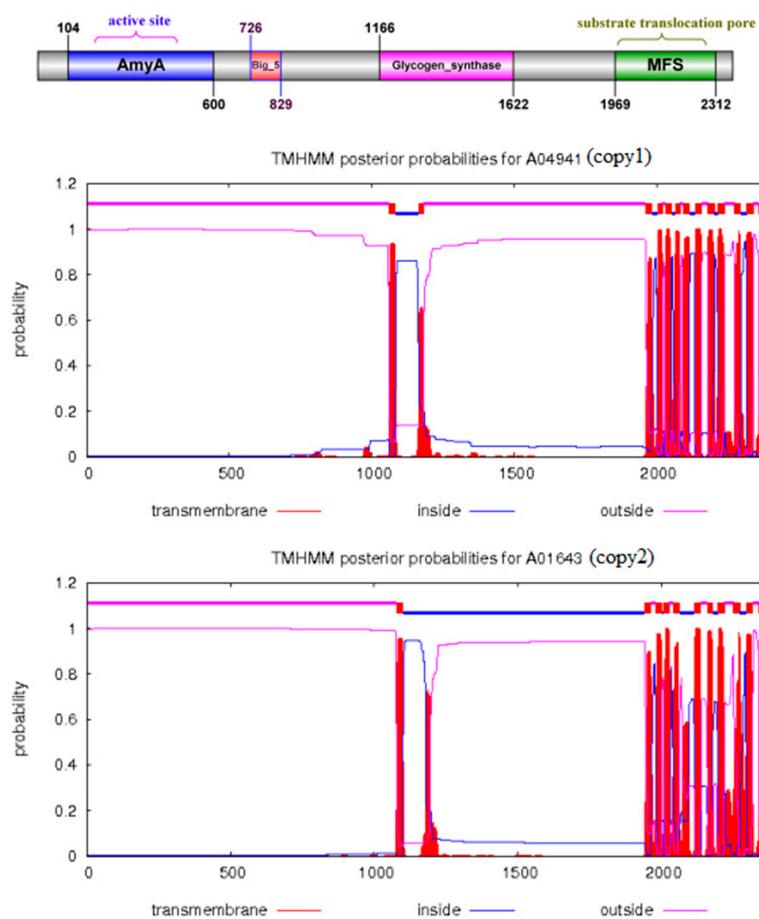
**Figure 2.** Enrichment analysis of the unique genes in the TN3-1 strain.

Interestingly, we found some crucial genes from the specific genes, such as type II polyketo synthase-non-ribosomal peptide synthase, DNA repair protein, small GTP binding protein, 1,3- $\beta$ -glucan transferase GEL2, 1,3- $\beta$ -glucan transferase GEL3, the Ras family protein, Rho/Rac family GTPase activating protein-like protein, vacuole separation protein 55, mitogen-activated protein kinase, reverse transcriptase,  $\beta$ -glucosidase and cell wall mannoprotein hsp150, and they played important roles in secondary metabolite synthesis, DNA repair, spore formation and cell wall synthesis against osmotic pressure. In contrast, the specific genes in P16 strain included extracellular signal-regulated kinase, guanine exchange factor GEF, mitogen-activated protein kinase, WD repeat protein, and glucosyltransferase family proteins, which played an important role in signal transduction,

transport, chromosome modification, spore formation, activation of Rho-GTPase activity and regulation of cell wall synthesis. Besides, we also found that there were many unknown functional domains encoding non-characterized gene family proteins, such as DUF1479 domain protein, DUF1793 domain protein, DUF218 domain protein and the DUF89 domain protein. Evolutionary analysis showed that members of the DUFs protein family might have many important functions. DUF family member DUF1742 domain protein was also found in CBS105.22 strain, and the other specific genes included  $\alpha$ -1,2-mannosyltransferase mnn23, mitogen-activated protein kinase, stress response A/B domain protein, canavalin A-like lectin/glucanase, and copper Zinc superoxide dismutase like protein. These specific genes may have different characteristics in osmotic stress response, spore formation and free radical scavenging.

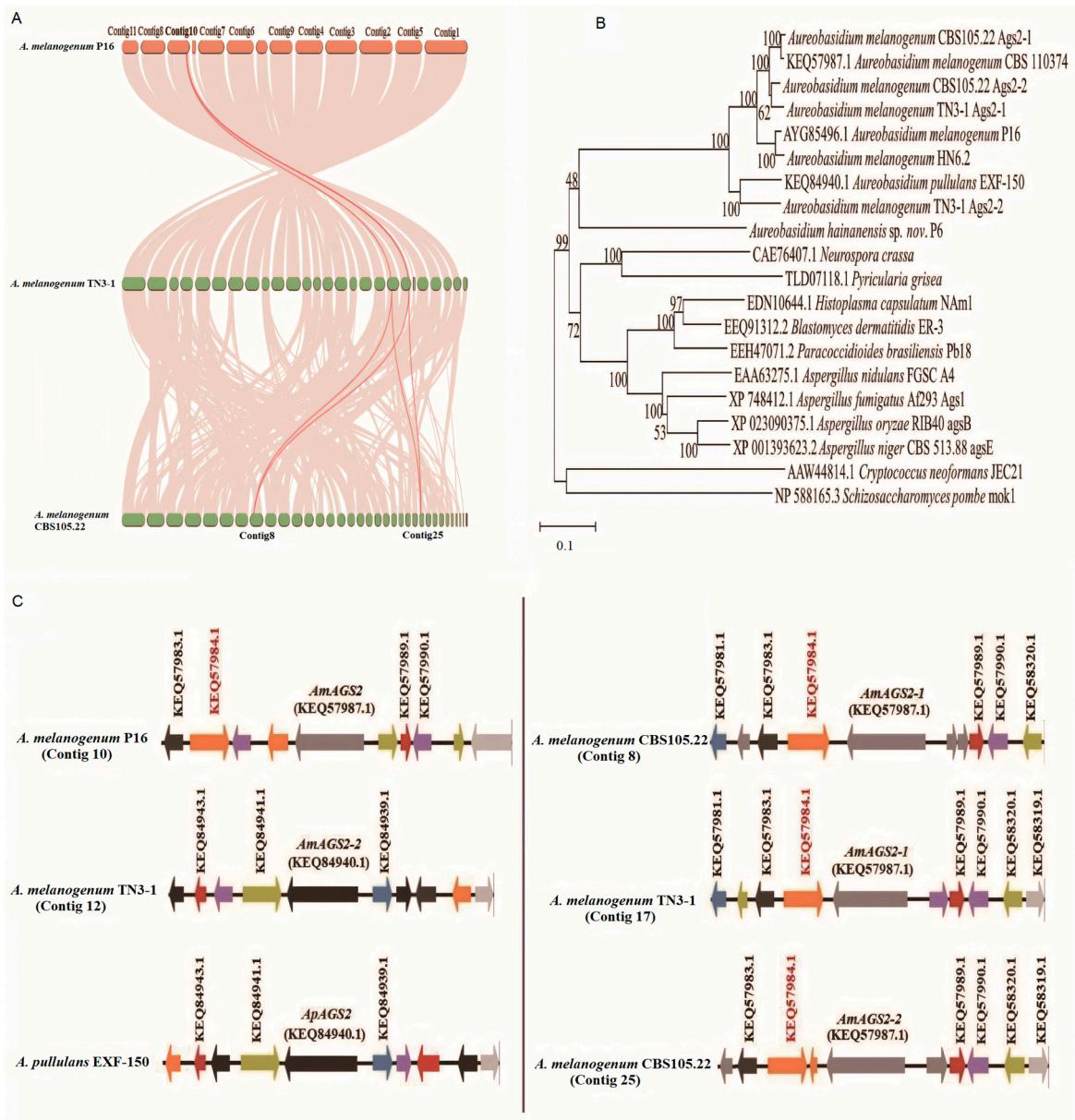
### 3.2. TN3-1 Strain Had a Unique Pullulan Synthesis Gene

The  $\alpha$ -1,3-glucan synthase (Ags2)(EC 2.4.1.183) has been reported to play a crucial role in catalyzing the synthesis of pullulan, and this gene is prevalent in the strains of *Aureobasidium* [6,7]. However, the  $\alpha$ -1,3-glucan synthase coding gene was not found in the genomes of *Y. lipolytica* and *S. cerevisiae*, which might explain the reason why they could not produce pullulans. Further analysis showed that  $\alpha$ -1,3-glucan synthase had four conserved protein domains, which might play a synergistic role in pululan synthesis. The core catalytic domain belonged to the alpha-amylase family, and alpha-1,3-glucan was also one of the components of the cell wall. Studies showed that only one of the two homologous  $\alpha$ -1,3-glucan synthase coding genes played a crucial role in TN3-1 strain, while the other AGS2 gene could not catalyze the synthesis of pullulan. Interestingly, the two genes were not directly the "superimposed" relationships [16]. Thus, we analyzed the transmembrane structure of Ags2 proteins, and found that the two homologous Ags2 proteins of TN3-1 strain might be located on the cell membrane and were the membrane proteins, among which one Ags2 belonged to extracellular protein, while the other Ags2 might be a transmembrane protein located in both intracellular and extracellular cells (Figure 3).



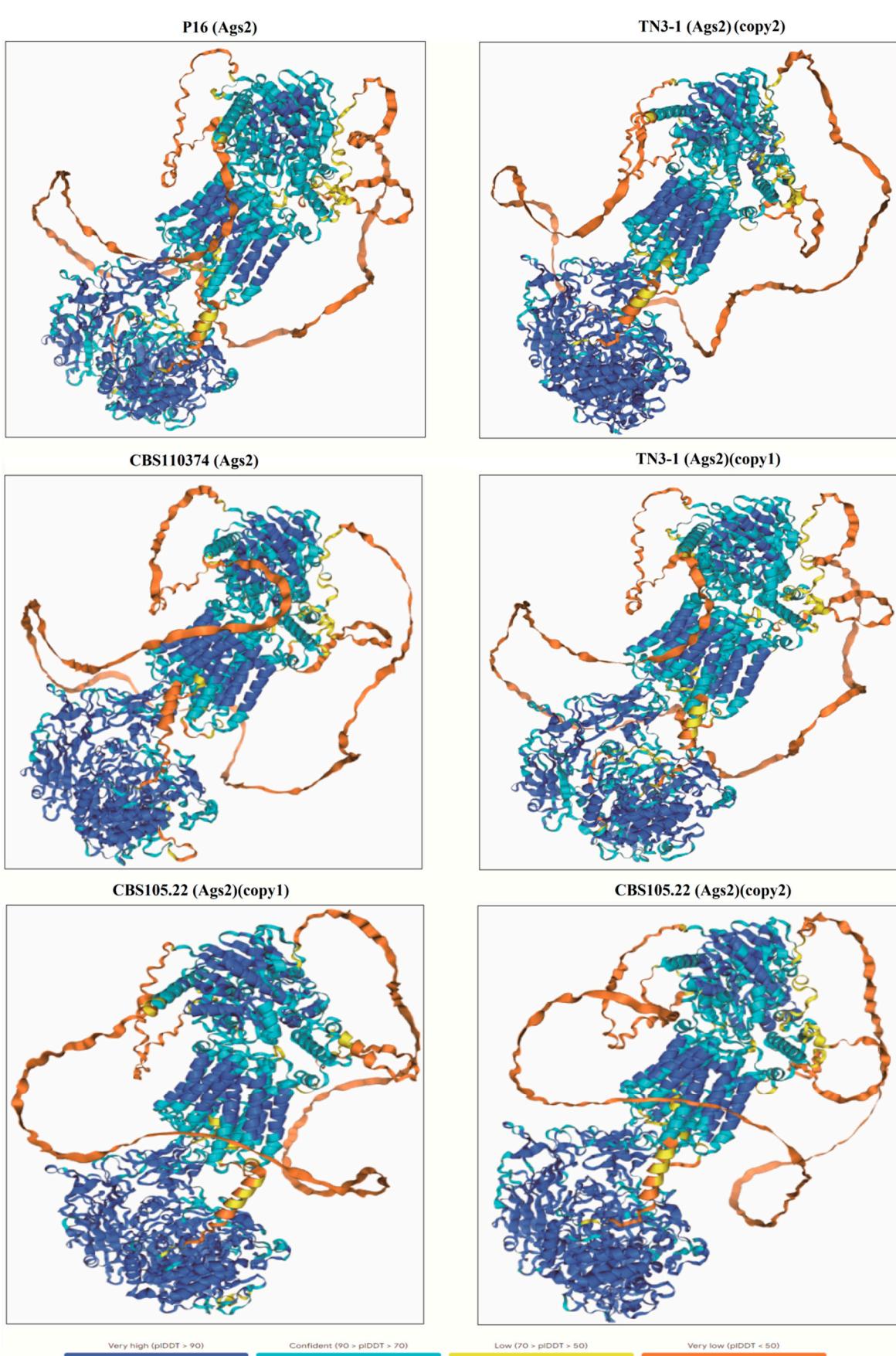
**Figure 3.** The conserved domain and transmembrane analysis of  $\alpha$ -1,3-glucan synthase in TN3-1. A: Domains of Ags2 in the TN3-1 strain; B: Transmembranes of two Ags2 proteins in TN3-1.

In genetics,  $Ka/Ks$  or  $d_N/ds$  represents the ratio between non-synonymous substitution (Ka) and synonymous substitution (Ks), and this ratio means whether there is selection pressure acting on this protein-coding gene. A nucleotide variation that does not result in an amino acid change is called a synonym mutation, and the reverse is called a non-synonym mutation. It is generally believed that synonymous mutations are not subject to natural selection, while non-synonymous mutations are subject to natural selection. If  $d_N/ds > 1$ , it is considered to have positive selection effect, if  $d_N/ds = 1$ , it is considered to have neutral selection, if  $d_N/ds < 1$ , it is considered to have purification selection [17]. The selection pressure analysis of AGS2 of different strains was performed by PAML 4.9 software. The results showed that omega ( $d_N/ds$ ) value was only 0.04,  $d_N/ds < 0.25$ , and the synonymous mutation rate was greater than the non-synonymous mutation rate. The results indicated that the AGS2 gene of TN3-1 strain was purified and selected in the process of evolution, eliminating harmful mutations to maintain its important biological functions, and gradually retained for the offspring during evolution. Whole-genome collinearity analysis showed that the two AGS2 genes of TN3-1 strain had collinearity with that of the P16 strain, while the two AGS2 genes of CBS105.22 strain had collinearity with each AGS2 gene of the TN3-1 strain (Figure 4A). Therefore, it can be inferred that the three strains may have a common ancestor, the genome duplication events may occur before species differentiation, and most homologous genes have the same or similar functions, similar regulatory pathways, and conserved functions. Furthermore, it was found that the two AGS2 genes of strain CBS105.22 had high similarity values, and all of them were close to one AGS2 gene of strain TN3-1. The duplicated AGS2 genes of strain CBS105.22 may be paralogous genes produced by genome duplication. Due to the lack of original natural selection force, the duplicated AGS2 gene can be freely mutated and may acquire new functions according to the theory of genome polyploidization. The collinearity analysis of the AGS2 gene showed that the AGS2 gene was located in the collinearity block of the TN3-1, P16 and CBS105.22 strains, and had obvious paralogous characteristics in CBS105.22 strain. Further analysis showed that one copy of TN3-1 had an obvious collinearity relationship with the arrangement of the two AGS2 genes in strain CBS105.22, and the similarity value of this gene was high (96%). In addition, the gene arrangement of the other AGS2 gene in TN3-1 strain was more consistent with that of *A. pullulans* EXF-150 than the other two strains (Figure 4B), and the similarity between this gene was about 84.51%, higher than that of the P16 strain (80.58%). That meant that this AGS2 gene might have associations with *A. pullulans* strain, and underwent the process of natural selection, gradually being inherited by eliminating harmful mutations.



**Figure 4.** The  $\alpha$ -1,3-glucan synthase encoding genes of different *Aureobasidium* strains. **A:** The collinearity of *AGS2* genes; **B:** The phylogenetic tree of *Ags* proteins; **C:** Arrangement of the *AGS2* gene and its neighboring genes.

Furthermore, a 3D model of the  $\alpha$ -1,3 glucan synthase was also constructed to further predict its 3D structure. The results showed that the 3D structure of the crucial pullulan synthesis enzyme (copy2) in TN3-1 strain was different from that in any other strain of *A. melanogenum* spp., and had local similarity with that of the P16 strain. In contrast, the other  $\alpha$ -1,3-glucan synthase (copy1) in TN3-1 strain shared a similar folding pattern with that of CBS105.22 and the type strain CBS110374 (Figure 5). These  $\alpha$ -1,3-glucan synthetases from different strains of *A. melanogenum* seem to display different folding patterns, and the special 3D structure folding of the crucial *Ags2* protein (copy2) in TN3-1 strain may affect its gene expression efficiency.



**Figure 5.** The 3D structure comparison of  $\alpha$ -1,3-glucan synthase from different strains of *Aureobasidium melanogenum* spp.

### 3.3. TN3-1 Strain Activated Distinct Pathways in Response to Environmental Changes

Under low nitrogen and osmotic stress environments, yeast can respond to osmotic stress through two signaling pathways, namely the MAPK cascade cell wall integrity signaling pathway and Ras/cAMP signaling pathway, which helps cells search for nutrients [18]. Multiple binding sites of transcription factor (TF) Tec1 and Ste12 were found in the  $\beta$ -1,3-glucan (FKS) gene promoter of the TN3-1, P16 and CBS105.22 strains. The two transcription factors can synergically regulate mycelia growth in different metabolic pathways. The transcription factor Swi4/6 was a zip-specific transcriptional activator in regulating the cell wall integrity signaling pathway (CWI), however, there was no TF binding sites of Swi4/6 in TN3-1 and P16 strains, but the TF binding site (CACGAAA) was only found in the promoter of the  $\beta$ -1, 3-glucan synthase encoding gene of CBS105.22 strain (Table 3). In the cell wall integrity pathway, Bck2 gene first activates the expression of M/G1 gene Swi4 by combining with transcription factor Mcm1, and then further regulates the expression of downstream genes by transcription factor Swi4 [19]. In strain CBS105.22, we found multiple binding sites (CCYWWNNRG) of Mcm1 in the SWI4 promoter region, indicating a complex regulation network. In addition, in TN3-1 and P16 strains, the binding sites (ACCAGC) of transcription factor Ace2 were observed in the upstream region of the AGS2 gene (Table 3). The C-terminal structure of ACCAGC was identical to that of transcription factor Msn2, which played an important role in regulating mycelium morphology, adhesion, membrane formation and cell separation [20]. Thus, due to the differences in gene regulation, the expression of AGS2 gene may be relatively weak in the non-pullulan producing strain CBS105.22, while the CWI pathway may play a dominant role in CBS105.22 strain, so as to cope with the osmotic stress response.

**Table 3.** Comparison of transcription factor binding sites in different promoter regions.

Promoter	Strains	SWI4/6	Promoter	Strains	MSN2
FKS	P16	-	PFK	P16	-
	TN3-1	-		TN3-1	-
	CBS105.22	+		CBS105.22	+
Promoter	Strains	ACE2	Promoter	Strains	MSN2
AGS2	P16	+	PGM	P16	+
	TN3-1	+		TN3-1	+
	CBS105.22	-		CBS105.22	-

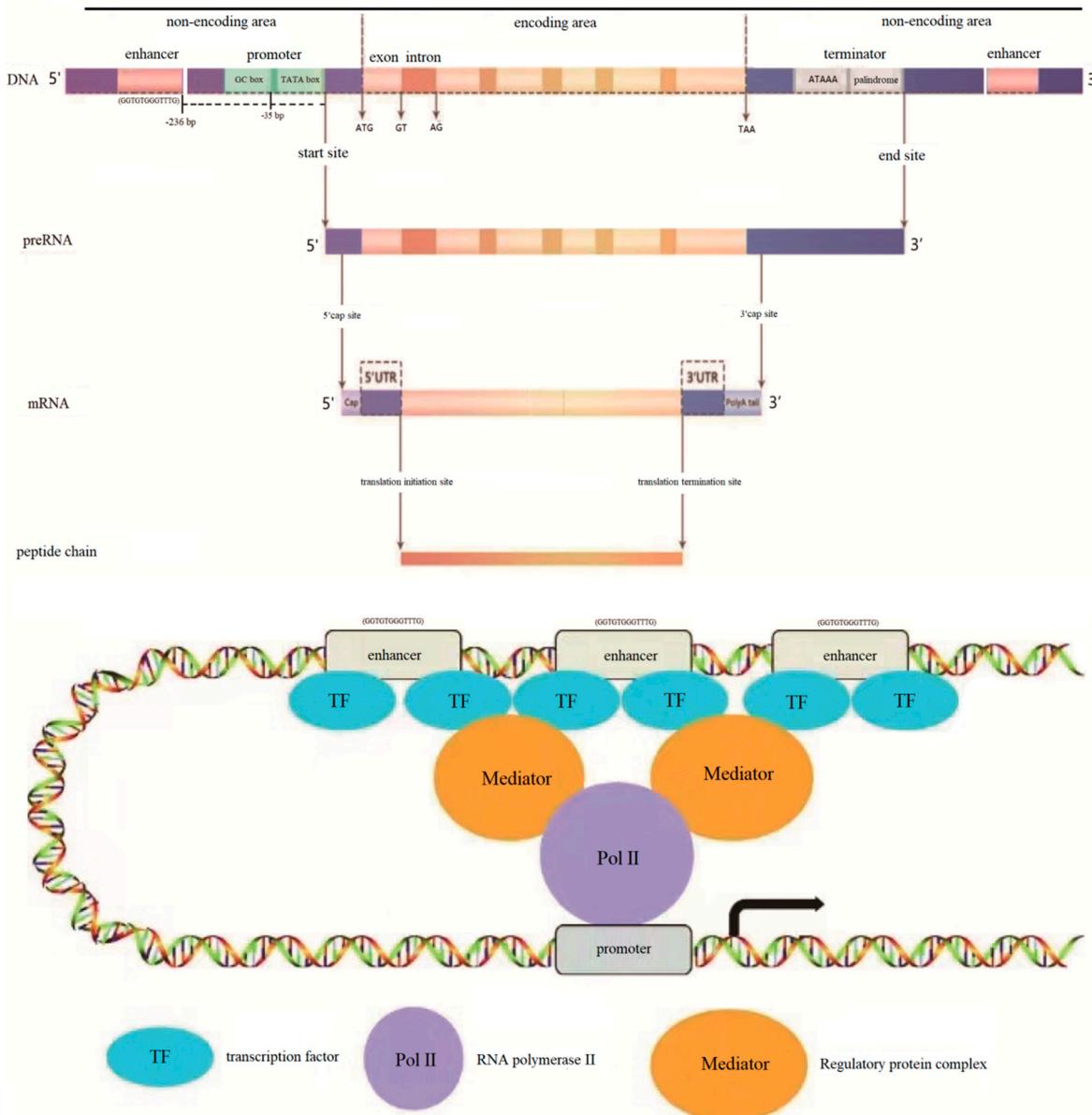
+: at least one transcription factor binding site was presented; -: no transcription factor binding sites were detected.

Studies have shown that cells produce more ATP for the metabolism of polymeric precursors to synthesize more pullulans under low nitrogen conditions, and the expression of glycolysis genes is mainly regulated by transcription factor Msn2 [21,22]. However, we found no Msn2 binding site (AGGGG) in the upstream region of phosphofructokinase gene (PFK) in TN3-1 and P16 strains, while in strain CBS105.22, the Msn2 binding site (AGGGG) existed in the upstream of the two PFK gene copies (Table 3). Therefore, in the pullulan producing strain P16 and TN3-1, carbon flow was possibly more concentrated in the glycogen synthesis pathway, while in strain CBS105.22, the carbon flow might go through the glycolysis pathway to generate acetyl CoA, which was conducive to growth [23]. In addition to the AGS2 gene, other important genes involved in the synthesis process of pullulan include the UDPG-pyrophosphorylase gene (UGP), the glucose-mutase phosphate gene (PGM), the UDPG-glucosyl transferase gene (UGT1), and the glucokinase gene (Gluk). Among them, PGM catalyzes the mutation of glucose 6-phosphate to form glucose 1-phosphate, UGP catalyzes glucose 1-phosphate and UTP to form UDPG, and UGT1 catalyzes the substrate UDPG to synthesize polysaccharide chain, which are considered to be important rate-controlling steps in the synthesis of pullulan [24]. Through the transcription regulation analysis, the results showed that in TN3-1 and P16 strains, there was an Msn2 transcription factor regulatory site (AGGGG) in the promoter region of the PGM gene (Table 3). Under the regulation of transcription factors, The high expression of this

gene can produce a large amount of glucose-1-phosphate, which leads to the process of glycogen synthesis. In addition, the binding site (CCCT) of Msn2 was also found in *Y. lipolytica* and *S. cerevisiae*, however, it was not observed in strain CBS105.22. The expression of this gene may be relatively weak during the synthesis of pullulan precursors, so the carbon flow possibly diverted to cell growth in this strain.

### 3.4. Efficient Energy Supply Provides Strong Assurance

Analysis of the crucial genes in TCA cycle and respiratory chain showed that in the TN3-1 strain, one copy of the succinyl-CoA synthetase  $\beta$  subunit encoding gene contained an enhancer at the distance of -236 bp from the transcription start site. The core sequence of this enhancer was "GGTGTGGGTTG", while no enhancer was found in the promoter region of the P16 and CBS105.22 strains. Meanwhile, a GC-box (GGGCGG) was also observed at -35 bp from the transcription start site, which could enhance the transcription start frequency (Figure 6). The above characteristics may ensure the efficient energy supply of TN3-1 strain in the process of pullulan synthesis.



**Figure 6.** Schematic diagram of the upstream promoter elements and regulation of the crucial genes.

#### 4. Discussion

Consistent with existing research [10], we found that TN3-1 strain had a larger genome and more genes than any other strains of *Aureobasidium* spp.. Especially, it carried many more specific genes than the other *Aureobasidium* strains. The specific genes of TN3-1 strain involved the osmotic pressure related genes, as well as the genes of wall synthesis. According to the gene collinearity, the specific genes in TN3-1, P16 and CBS105.22 strains may originate from a common ancestor and have different functions with differentiation. Among the specific genes, several Ras and GTP protein family members as well as the genes related to regulation of cell wall synthesis were found in TN3-1 strain, which might have special advantages in coping with osmotic stress response. In addition, there are a certain number of homologous genes with different species of *Aureobasidium* spp., indicating that TN3-1 strain is possibly a heterozygous strain. Indeed, it was speculated to be a new species formed by the fusion of different *Aureobasidium* strains [10]. In this study, we proposed that the TN3-1 strain carried a unique AGS2 gene, mainly manifested in its unique transmembrane structure and three-dimensional folding pattern. However, it seemed that the unique AGS2 gene of TN3-1 did not undergo positive selection, but retained its original functions. The gene dose effect might not be applicable to *Aureobasidium* spp. with genome duplication, as studies found that the two AGS2 copies of TN3-1 strain were not superimposed [5,10]. This may be related to the ancient evolutionary status of the TN3-1 strain (late Oligocene) [10], which is more ancient than other strains of *Aureobasidium* spp.. After genome duplication, one copy of the AGS2 gene might lose its original function due to the accumulation of harmful mutations, and the subsequent evolution of the non-pullulan producing strain CBS105.22 inherited one copy of the AGS2 gene from TN3-1 strain, and further underwent gene duplication. After all, in the genomes of eubacteria and archaea, ~ 30%-50% of genes belonged to paralogues, which were higher in eukaryotic genome [25]. Previous studies speculated that TN3-1 strain was formed by the fusion of the ancestor of CBS105.22 and an ancient strain similar to P16 strain, thus possessing the characteristics of both strains [10]. In this case, strain TN3-1 can tolerate high temperature environments, which is similar to the CBS105.22 strain isolated from granuloma of child, and also has the ability of pullulan production similar to the P16 strain.

In this article, we found that the TN3-1 strain may respond to environmental changes by initiating specific signaling pathways. The response to osmotic stress is demonstrated through the cell wall integrity signaling pathway and the Ras/cAMP signaling pathway, which have been shown to be associated with the pullulans synthesis. For example, there are transcription factor binding sites in the upstream of some crucial genes, while these regulatory sites do not exist in P16 and CBS105.22 strains. Similarly, regulatory sites of transcription factors were also found in the upstream of the glycolytic genes. Under the regulation of transcription factors, high expression of these genes can synthesize a large amount of glucose-1-phosphate, leading to a shift in metabolic flow towards glycogen synthesis. However, in the CBS105.22 strain, the expression of these genes may be relatively weak, resulting in a greater shift in carbon flow towards cell growth. This could be supported by the previous studies [7,9]. Combined with the above analysis, it was speculated that under osmotic pressure conditions, different strains of *Aureobasidium* spp. could activate different signaling pathway response strategies in response to environmental changes, which was in line with the characteristics of adaptive evolution [26]. The pullulan producing strain TN3-1 and P16 may be mainly regulated by the Ras/cAMP signaling pathway, while strain CBS105.22 may resist osmotic pressure mainly through the cell wall integrity signaling pathway (CWI) due to the high sensitivity and permeability of biofilm to glucose concentration. In strain CBS105.22, the FKS gene was regulated by transcription factor Swi4 to synthesize  $\beta$ -1,3-glucan for cell wall synthesis, which showed filamentous characteristics [27]. As reported, previous studies showed that enlarged cells and thick-walled spores were the main causes of pululan formation, while conidium and mycelium had no effect on the synthesis of pullulan [28]. Thus, the Ras/cAMP signaling pathway may be a major pathway for pullulan synthesis under low nitrogen and osmotic stress conditions, and the downstream gene expression is mainly regulated by ACE2, MSN2 and other transcription factors [20,29,30]. In addition, the enhancer in the upstream of the TCA cycle associated gene may enhance the expression of this gene, ensuring efficient energy supply during metabolic process. Usually, the

transcriptional enhancer has a transcriptional enhancement effect, which can generally increase the transcription frequency by 10-200 times [31]. In the process of pullulan synthesis, GTP synthesis by succinyl-CoA synthetase in the TCA cycle is the only substrate level phosphorylation reaction, and the generated GTP can be used for energy supply or rapidly activate the Ras/cAMP signaling pathway, making it at an advantage in stress response.

Although we have further explored the mechanisms of stress tolerance and pullulan production, there still exist some shortcomings. First, we only supplemented the results based on previous studies and did not compare it with more polyploid strains, which limited the research scope. Second, our research was limited to the downloaded genomic data and lacked necessary experimental validations. Finally, we did not make analysis of the interactions between promoters and the crucial proteins through AlphaFold Server 3.0. Until now, NCBI has totally released 176 *Aureobasidium* genomes (3th July, 2024) (<https://www.ncbi.nlm.nih.gov/assembly/?term=Aureobasidium>). With the rapid development of next sequencing technology, more and more *Aureobasidium* genomes will be released. Full-scale genome analysis will be beneficial for discovering interesting strains and functional genes, revealing the potential molecular mechanisms of stress tolerance and pullulan synthesis.

## 5. Conclusion

On basis of previous studies, we further analyzed the genomic characteristics of TN3-1 strain and the possible molecular mechanisms of its high-yield ability of pullulan production. The genome duplication, gene function differentiation and promoter region changes of TN3-1 strain may be the important factors for its special phenotype. Meanwhile, it may have heterozygous advantages in the evolutionary process. It will be a type strain for the study of microorganisms synthesizing pullulans, and provide new insights for studying adaptive evolution of genome duplication in *Aureobasidium* spp.

**Author Contributions:** Conceptualization, J.Y., S.J.; Methodology, W.W., J.Z.; Investigation and data curation, R.Z.; software, R.Z., S.S., Y.S., B.L.; writing—original draft preparation, J.Y.; writing—review and editing, S.J. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work were supported by the Youth Project of National Natural Science Foundation of China, 82301759; Youth Project of Shanxi Basic Research Program, No. 20210302124561.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** The computational requirements for this work were supported in part by the Tianjin Medical University High Performance Computing (HPC) Center's resources and personnel.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Tiwari S, Patil R, Dubey SK, Bahadur P. Derivatization approaches and applications of pullulan. *Adv Colloid Interface Sci.* 2019 Jul;269:296-308.
2. Wang Y, Liang Z, Liang Z, Lv W, Chen M, Zhao Y. Advancements of Prussian blue-based nanoplatforms in biomedical fields: Progress and perspectives. *J Control Release.* 2022 Nov;351:752-778.
3. Singh RS, Kaur N, Singh D, Purewal SS, Kennedy JF. Pullulan in pharmaceutical and cosmeceutical formulations: A review. *Int J Biol Macromol.* 2023 Mar 15;231:123353.
4. Jia SL, Chi Z, Liu GL, Hu Z, Chi ZM. Fungi in mangrove ecosystems and their potential applications. *Crit Rev Biotechnol.* 2020 Sep; 40 (6) : 852-864.
5. Chen L, Wei X, Liu GL, Hu Z, Chi ZM, Chi Z. Glycerol, trehalose and vacuoles had relations to pullulan synthesis and osmotic tolerance by the whole genome duplicated strain *Aureobasidium melanogenum* TN3-1 isolated from natural honey. *Int J Biol Macromol.* 2020; 165(Pt A):131-140.

6. Chen TJ, Liu GL, Wei X, Wang K, Hu Z, Chi Z, Chi ZM. A multidomain alpha-Glucan synthetase 2 (AmAgs2) is the crucial enzyme for pullulan biosynthesis in *Aureobasidium melanogenum* P16. *Int J Biol Macromol.* 2020 May 1; 150:1037-1045.
7. Chen TJ, Chi Z, Jiang H, Liu GL, Hu Z, Chi ZM. Cell wall integrity is required for pullulan biosynthesis and glycogen accumulation in *Aureobasidium melanogenum* P16. *Biochim Biophys Acta Gen Subj.* 2018 Jun;1862(6):1516-1526.
8. Zhang T, Yang R, Yang S, Guan J, Zhang D, Ma Y, Liu H. Research progress of self-assembled nanogel and hybrid hydrogel systems based on pullulan derivatives. *Drug Deliv.* 2018 Nov;25(1):278-292.
9. Xue SJ, Chen L, Jiang H, Liu GL, Chi ZM, Hu Z, Chi Z. High pullulan biosynthesis from high concentration of glucose by a hyperosmotic resistant, yeast-like fungal strain isolated from a natural comb-honey. *Food Chem.* 2019 Jul 15;286:123-128.
10. Jia SL, Zhang M, Liu GL, Chi ZM, Chi Z. Novel chromosomes and genomes provide new insights into evolution and adaptation of the whole genome duplicated yeast-like fungus TN3-1 isolated from natural honey. *Funct Integr Genomics.* 2023 Jun 19; 23 (3) : 206.
11. Gostinčar C, Ohm RA, Kogej T, et al. Genome sequencing of four *Aureobasidium pullulans* varieties: biotechnological potential, stress tolerance, and description of new species. *BMC Genomics.* 2014 Jul 1;15:549.
12. Wang Y, Tang H, Wang X, Sun Y, Joseph PV, Paterson AH. Detection of colinear blocks and synteny and evolutionary analyses based on utilization of MCScanX. *Nat Protoc.* 2024.
13. Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature.* 2024 Jun;630(8016):493-500.
14. Contreras-Moreira B, Vinuesa P. GET\_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl Environ Microbiol.* 2013 Dec;79(24):7696-701.
15. Monteiro PT, Oliveira J, Pais P, et al. YEASTRACT+: a portal for cross-species comparative genomics of transcription regulation in yeasts. *Nucleic Acids Res.* 2020 Jan 8;48(D1):D642-D649.
16. Qi CY, Jia SL, Wei X, Yang G, Chi Z, Liu GL, Hu Z, Chi ZM. The differences between fungal alpha-glucan synthase determining pullulan synthesis and that controlling cell wall Alpha-1,3 glucan synthesis. *Int J Biol Macromol.* 2020 Nov 1; 162:436-444.
17. Álvarez-Carretero S, Kapli P, Yang Z. Beginner's Guide on the Use of PAML to Detect Positive Selection. *Mol Biol Evol.* 2023 Apr 4;40(4):msad041.
18. Vinod PK, Sengupta N, Bhat PJ, Venkatesh KV. Integration of global signaling pathways, cAMP-PKA, MAPK and TOR in the regulation of FLO11. *PLoS One.* 2008 Feb 27; 3(2):e1663.
19. Bastajian N, Friesen H, Andrews BJ. Bck2 acts through the MADS box protein Mcm1 to activate cell-cycle-regulated genes in budding yeast. *PLoS Genet.* 2013 May;9(5):e1003507.
20. Yang G, Liu GL, Wang SJ, Chi ZM, Chi Z. Pullulan biosynthesis in yeast-like fungal cells is regulated by the transcriptional activator Msn2 and cAMP-PKA signaling pathway. *Int J Biol Macromol.* 2020 Aug 15; 157:591-603.
21. Rødkær S, Færgeman N. Glucose- and nitrogen sensing and regulatory mechanisms in *Saccharomyces cerevisiae*[J]. *FEMS yeast research,* 2014, 14(5): 683-696.
22. Gorner W, Durchschlag E, Wolf J, Brown EL, Ammerer G, Ruis H, Schuller C. Acute glucose starvation activates the nuclear localization signal of a stress-specific yeast transcription factor. *EMBO J.* 2002 Jan 15; 21 (1-2) : 135-144.
23. Kuang Z, Ji H, Boeke JD. Stress response factors drive regrowth of quiescent cells. *Curr Genet.* 2018 Aug; 64 (4) : 807-810.
24. Li Y, Chi Z, Wang GY, Wang ZP, Liu GL, Lee CF, Ma ZC, Chi ZM. Taxonomy of *Aureobasidium* spp. and biosynthesis and regulation of their extracellular polymers. *Crit Rev Microbiol.* 2015 Jun; 41 (2) : 228-237.
25. Koonin EV. Orthologs, paralogs, and evolutionary genomics. *Annu Rev Genet.* 2005; 309-338
26. Causton HC, Ren B, Koh SS, Harbison CT, Kanin E, Jennings EG, Lee TI, True HL, Lander ES, Young RA. Remodeling of yeast genome expression in response to environmental changes. *Mol Biol Cell.* 2001 Feb; 12 (2) : 323-337.
27. Hochstenbach F, Klis FM, Van d EH, et al. Identification of a putative alpha-glucan synthase essential for cell wall construction and morphogenesis in fission yeast[J]. *Proceedings of the National Academy of Sciences,* 1998, 95(16):9161-9166.
28. Campbell BS, Siddique AB, McDougall BM, Seviour RJ. Which morphological forms of the fungus *Aureobasidium pullulans* are responsible for pullulan production? *FEMS Microbiol Lett.* 2004 Mar 19; 232 (2): 225-228.
29. Estruch F. Stress-controlled transcription factors, stress-induced genes and stress tolerance in budding yeast. *FEMS Microbiol Rev.* 2000 Oct; 24 (4): 469-486.

30. Kottom TJ, Limper AH. The *Pneumocystis* Ace2 transcription factor regulates cell wall-remodeling genes and organism virulence. *J Biol Chem.* 2013 Aug 16;288(33):23893-902.
31. Haberle V, Stark A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat Rev Mol Cell Biol.* 2018 Oct;19(10):621-637.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.