

**Table 4. Major paralogous gene families in BF, BT, and PG**

Top 20	BF strain YCH46		BT strain VPI-5482		PG strain W83	
	Function	No. of genes	Function	No. of genes	Function	No. of genes
1	Outer membrane protein, SusC homolog	54	Outer membrane protein, SusC homolog	79	ABC transporter, ATP-binding protein 1	14
2	RNA polymerase ECF-type sigma factor	38	Outer membrane protein, SusD homolog	49	DNA-binding protein, histone-like family	9
3	Antisigma factor	25	RNA polymerase ECF-type sigma factor	43	Hypothetical protein	6
4	ABC transporter, ATP-binding protein	25	Antisigma factor	25	RNA polymerase ECF-type sigma factor	4
5	Outer membrane protein, SusD homolog	23	Two-component system sensor histidine kinase/response regulator, hybrid	24	$\alpha$ -1,2-mannosidase	4
6	Two-component system sensor histidine kinase	11	Site-specific recombinase (tyrosine type)	23	Oxidoreductase	4
7	Glycosyltransferase (GT family 2)	11	$\alpha$ -1,2-mannosidase	23	Peptidyl-prolyl <i>cis-trans</i> isomerase, FKBP type	4
8	Hexosaminidase precursor	10	ABC transporter ATP-binding protein	23	Na <sup>+</sup> -transporting NADH:ubiquinone oxidoreductase	4

Top 20	BF strain YCH46		BT strain VPI-5482		PG strain W83	
	Function	No. of genes	Function	No. of genes	Function	No. of genes
9	Putative transcriptional regulator, UpxY homolog	10	DNA-binding protein, histone-like family	15	Hypothetical protein	4
10	Outer membrane protein (Omp121 family)	9	Outer membrane protein (Omp121 family)	15	Hypothetical protein	4
11	$\alpha$ -1,2-mannosidase	9	Glycosyltransferase (GT family 2)	15	Hypothetical protein	4
12	Arylsulfatase	9	Two-component system sensor histidine kinase	13	Two-component system response regulator	3
13	Site-specific recombinase (tyrosine type)	9	Endo-1,4- $\beta$ -xylanase	12	Site-specific recombinase (tyrosine type)	3
14	Hypothetical protein	9	$\beta$ -hexosaminidase precursor	12	ABC transporter, ATP-binding protein, 2	3
15	DNA-binding protein, histone-like family	9	Arylsulfatase	12	ABC transporter, ATP-binding protein, 3	3
16	Conserved hypothetical protein	9	Oxidoreductase	11	Transmembrane cation efflux protein	3
17	AraC-type transcription regulator 1	8	AraC-type transcription regulator	11	2-amino-3-ketobutyrate CoA ligase	3
18	AraC-type transcription regulator 2	8	<i>N</i> -acetylmuramoyl-L-alanine amidase	10	Nitroreductase family protein	3
19	Two-component system response	8	Hypothetical protein	10	Carboxyl-terminal protease	3

Top 20	BF strain YCH46		BT strain VPI-5482		PG strain W83	
	Function	No. of genes	Function	No. of genes	Function	No. of genes
	regulator					
20	Conserved hypothetical protein, UpxZ homolog	8	Conserved hypothetical protein	10	Glycerate dehydrogenase	3

The largest paralogous gene families (top 20) found in BF strain YCH46, BT strain VPI-5482, and PG strain W83 genomes are listed. Genes located on mobile genetic elements were excluded from this analysis. Paralogues were defined as the genes whose amino acid sequences showed >30% identity over 60% of the length. The numbers of gene included in each paralogous gene family are thus not necessarily identical to those based on the annotation. For example, there are 34 tyrosine type site-specific recombinases in the BF strain YCH46 genome (in addition, 10 are on mobile genetic elements). But these have been divided into several families by the criterion used in this analysis.