

Supplementary Information



Alberto Scarampi

Contents

1 Methyl viologen treatment leads to enhanced light-dependent intracellular ROS accumulation	1
2 Adaptation to methyl viologen is not due to its abiotic degradation over time	2
3 Spot plates of Strains Grown in BG11 without or with MV (6 μ M)	4
4 Chronoamperometries strain by strain	6
5 Addition of methyl viologen at various growth phases	6
6 The gene <i>prqA</i> is not essential for spontaneous evolution of MV resistance	7
7 Whole genome sequencing and bioinformatic analyses	8
7.1 Genome sequencing QC report	9
7.2 Sequenced strains compared to various substrains	13
7.3 SNPs Polimorphism	15

1 Methyl viologen treatment leads to enhanced light-dependent intracellular ROS accumulation

Methyl viologen (MV) mediates its cytotoxic effects through the production of reactive oxygen species (ROS). To assess the impact of MV on ROS generation, we employed a ROS quantification assay using cultures of *Synechocystis* sp. PCC 6803 incubated with the ROS-sensitive fluorophore 2',7'-dichlorofluorescein diacetate (DCFH-DA). Given that MV can act as an electron acceptor from Photosystem I (PSI) and its mechanism of action is contingent upon photosynthetic electron flow, the assay was executed under both light and dark conditions. In accordance with our hypothesis, cultures treated with 10 μ M MV exhibited a significant intracellular accumulation of ROS. Notably, this ROS accumulation was markedly enhanced under light conditions, thereby corroborating that MV toxicity is mediated via a photosynthetic electron flow-dependent mechanism.

It is important to note that DCFH-DA is a non-specific ROS indicator, limiting our capacity to differentiate among various ROS types. However, prior studies have demonstrated that MV is capable

of generating superoxide anions, hydrogen peroxide, and hydroxyl radicals upon interaction with PSI, both *in vitro* and in cyanobacterial cells *in vivo*.

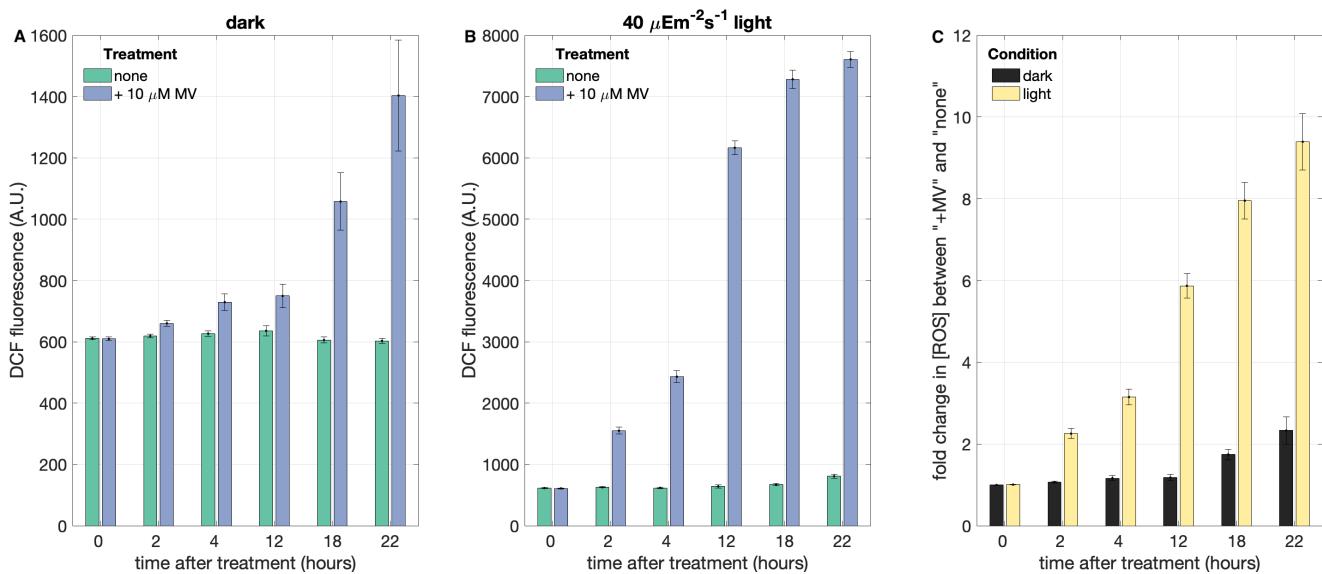


Figure 1: ROS quantification assay based on the substrate DCFH-DA, which becomes fluorescent upon reaction with intracellular ROS. A) DCF fluorescence in cultures of *Synechocystis* at various time points after treatment with 10 μM MV (in the dark). B) DCF fluorescence at various time points after treatment with 10 μM MV (under 40 $\mu\text{Em}^{-2}\text{s}^{-1}$ of white light illumination). C) Fold changes in DCF fluorescence between cultures treated with MV and no treatment control in the darkness and light at various time points. Error bars represent standard error of the mean from three biological replicates (individual colonies treated independently).

2 Adaptation to methyl viologen is not due to its abiotic degradation over time

Synechocystis cells were able to resume growing after growth-inhibiting treatments with MV. However such growth resumption could have been a result of the abiotic chemical degradation of methyl viologen inside the flasks, especially given that the growth curves were performed for a considerably long time at 30 degrees and in the presence of light. To discard this hypothesis, three cultures from adapted flasks were re-diluted, alongside unadapted controls and their susceptibility against addition of methyl viologen from a freshly prepared stock was quantified in a new growth curve. As demonstrated in Figure 2, previously adapted cultures were insensitive to addition of fresh methyl viologen, which instead expectedly inhibited the growth of wild-type, "unadapted" cultures. This confirmed that spontaneous degradation of methyl viologen was not responsible for the observed adaptation, which may be then due to the evolution of biotic resistance mechanisms.

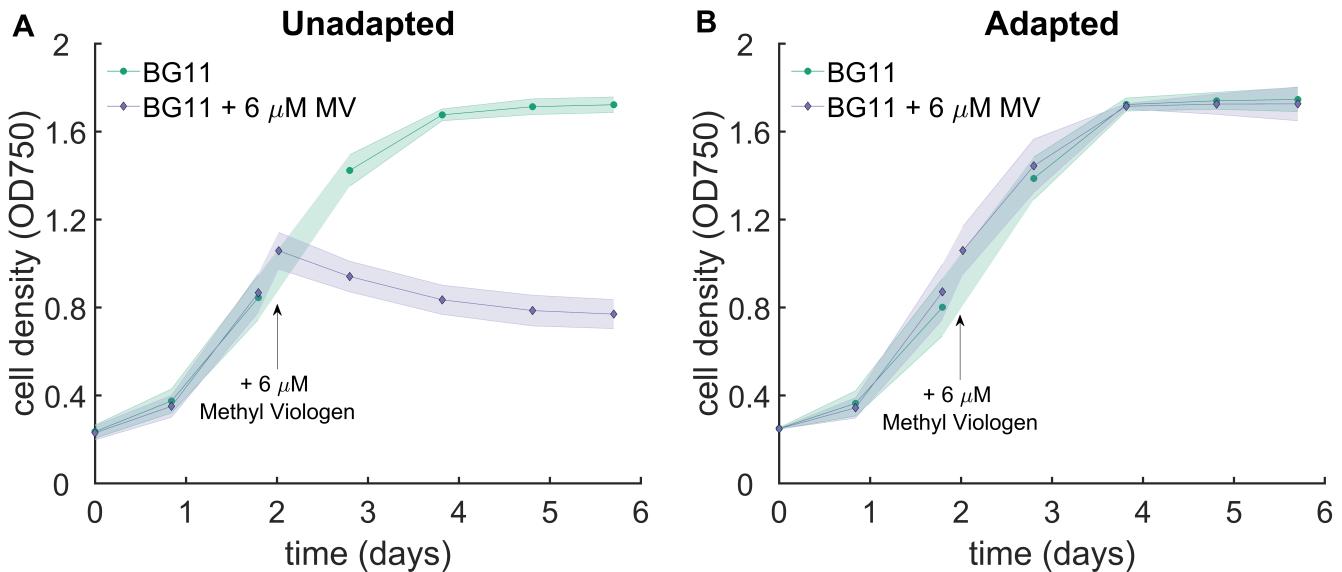


Figure 2: Growth curves of three independent unadapted (A) and MV-adapted (B) cultures of *Synechocystis* sp. PCC 6803 growing at 30 °C under constant illumination (intensity = $100 \mu E m^{-2} s^{-1}$) in the absence (green trace) and upon addition of $6 \mu M$ of methyl viologen (from a fresh stock) during exponential growth (purple trace). Shaded regions represent standard deviations from three biological replicates.

Figure 3 shows the absorbance spectra of unadapted (A) and adapted (B) strains after treatment with methyl viologen after 6 days of growth. Addition of methyl viologen to unadapted strains dramatically altered the pigment content of the strains, resulting in no detectable chlorophyll peaks. On the other hand, adapted strains were able to tolerate fresh addition of methyl viologen. Addition of methyl viologen to these strains results in non-lethal loss of chlorophyll concentration per cell compared to non-treated controls.

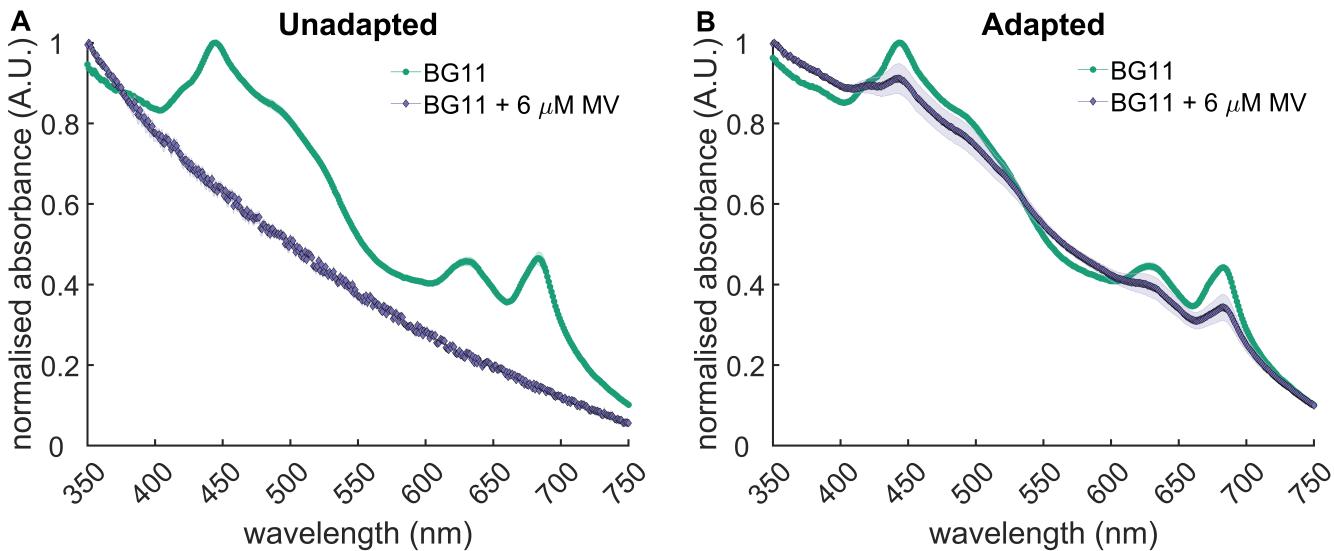


Figure 3: Absorbance spectra of unadapted (A) and adapted (B) cultures of *Synechocystis* sp. PCC 6803 after 6 days of growth at 30 ° under constant illumination (intensity = $100 \mu Em^{-2}s^{-1}$) in the absence (green trace) and upon addition of $6 \mu M$ of methyl viologen during exponential growth (purple trace). Shaded regions represent standard errors from three biological replicates.

3 Spot plates of Strains Grown in BG11 without or with MV (6 μM)

During the spontaneous evolution experiments, the cultures had been growing for over a month and had been exposed to multiple selective pressures. Therefore in the different flask, different sub-populations could have been present, complicating future DNA sequencing analyses. To overcome this limitation, adapted and unadapted cultures were spotted at different dilutions on BG11 plates containing no or $6 \mu M$ of methyl viologen. As shown in Figure 5, only previously adapted strains (labelled with "mvR") were able to survive on methyl viologen containing plates. This allowed to isolate individual colonies, which were then inoculated, grown and their gDNA sent for sequencing.

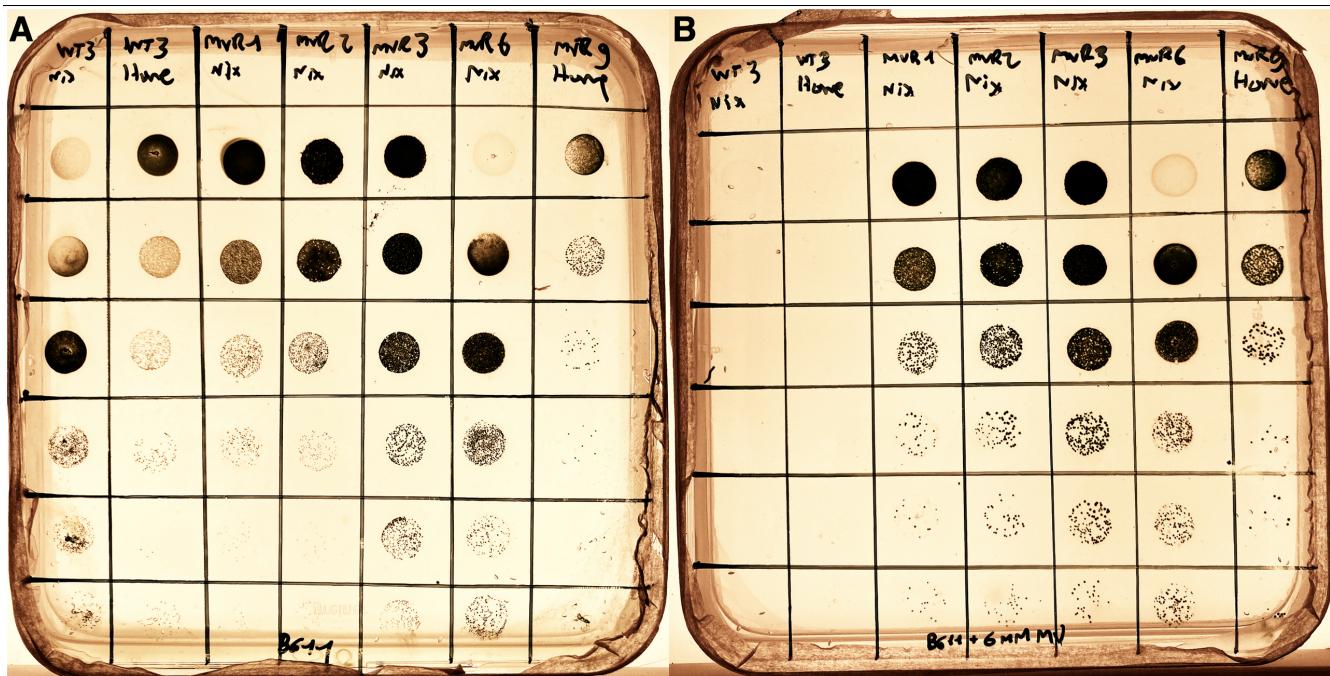


Figure 4: Spot plates of 2 unadapted (first two columns on each plates) and 5 adapted (columns 3 to 7) cultures of *Synechocystis* sp. PCC 6803 in BG11 plates without methyl viologen (A) and with the addition of $6 \mu\text{M}$ of methyl viologen (B). Rows correspond to consecutive 10-fold serial dilutions of culture inocula (all starting at on $\text{OD}_{750} = 1$). Plates were grown at 30° under constant illumination (intensity = $100 \mu\text{E m}^{-2}\text{s}^{-1}$) for 10 days.

4 Chronoamperometries strain by strain

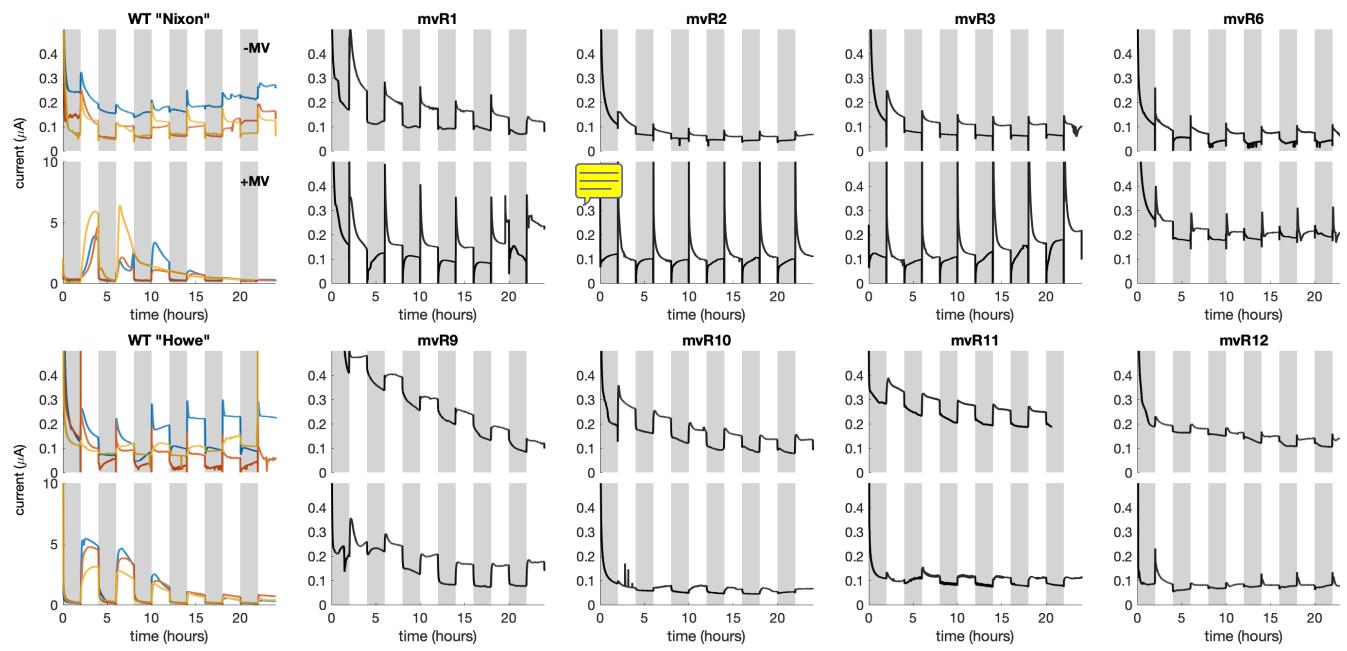


Figure 5: Chronoamperometry of all strains in the presence (bottom subtiles) and absence of $6 \mu\text{M}$ MV. White and gray panels represent periods of illumination and darkness respectively.

5 Addition of methyl viologen at various growth phases

To investigate whether adaptation to methyl viologen depends on the growth stage at which cells are at during the treatment, an additional growth curve was performed during which methyl viologen treatment was added at different cell phases. As shown in Figure 6, in this case only one culture managed to adapt to the methyl viologen treatment. This culture was treated during exponential phase. The higher the number of cells during the MV treatment, the more likely it would be that in the population there are some resistant cells already which then get enriched. Perhaps what is happening is that during exponential phase, genome replication rate is at its highest and this - together also with the mutagenic effects of the ROS produced by MV - might lead to an enhanced mutation rate in the cells and thus evolution of adapted strains.

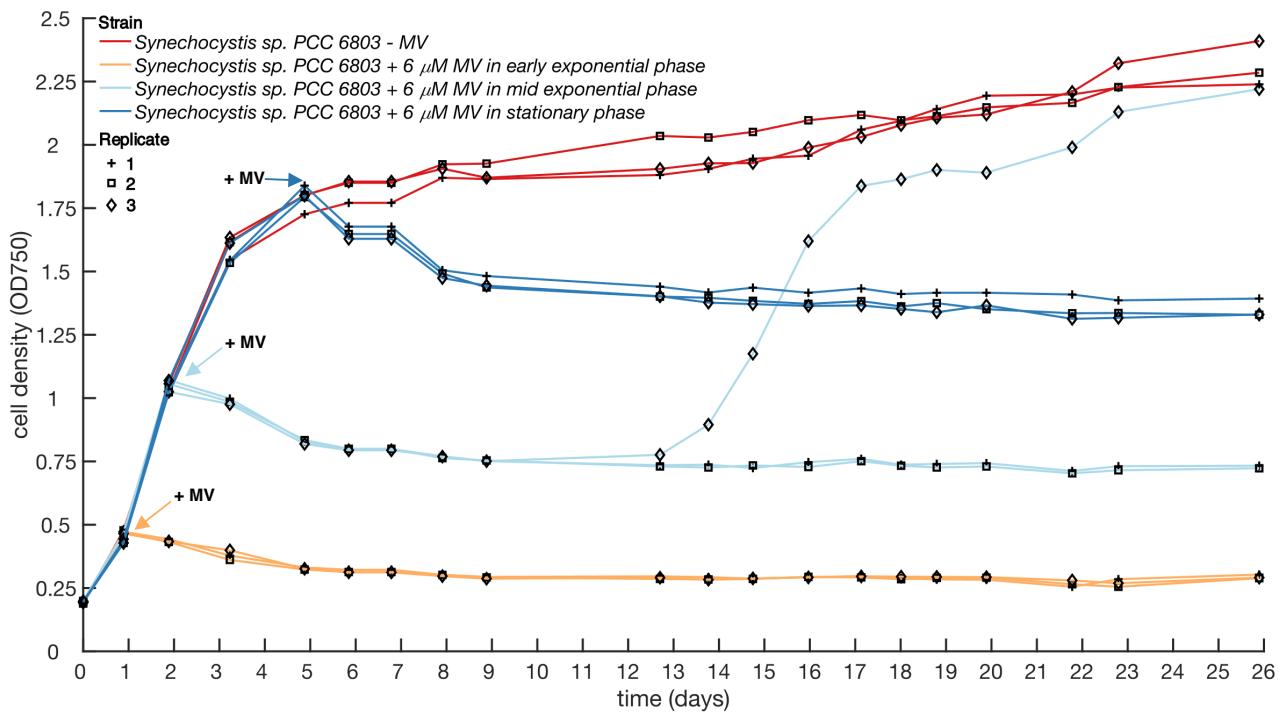


Figure 6: Growth curves of *Synechocystis* sp. PCC 6803 strains grown at 30 ° under constant illumination (intensity = $40 \mu Em^{-2}s^{-1}$) in the absence (red trace) and upon addition of 6 μM of methyl viologen during various growth stages (as indicated by the arrows colored according to the legend on top right).

6 The gene *prqA* is not essential for spontaneous evolution of MV resistance

Previous findings have identified that a mutation in the DNA-binding domain of the PrqR repressor led to resistance to methyl viologen. Because PrqR was found to repress the multidrug and toxin extrusion (MATE) protein PrqA, it was suggested that this protein was involved in the detoxification of methyl viologen via extrusion mechanisms. In order to validate whether PrqA is involved in the resistance against methyl viologen, wild-type, mutant and complemented *prqA* strains were subjected to the same growth profile as above to observe if there were any differences regarding the spontaneous evolution against methyl viologen for the different strains.

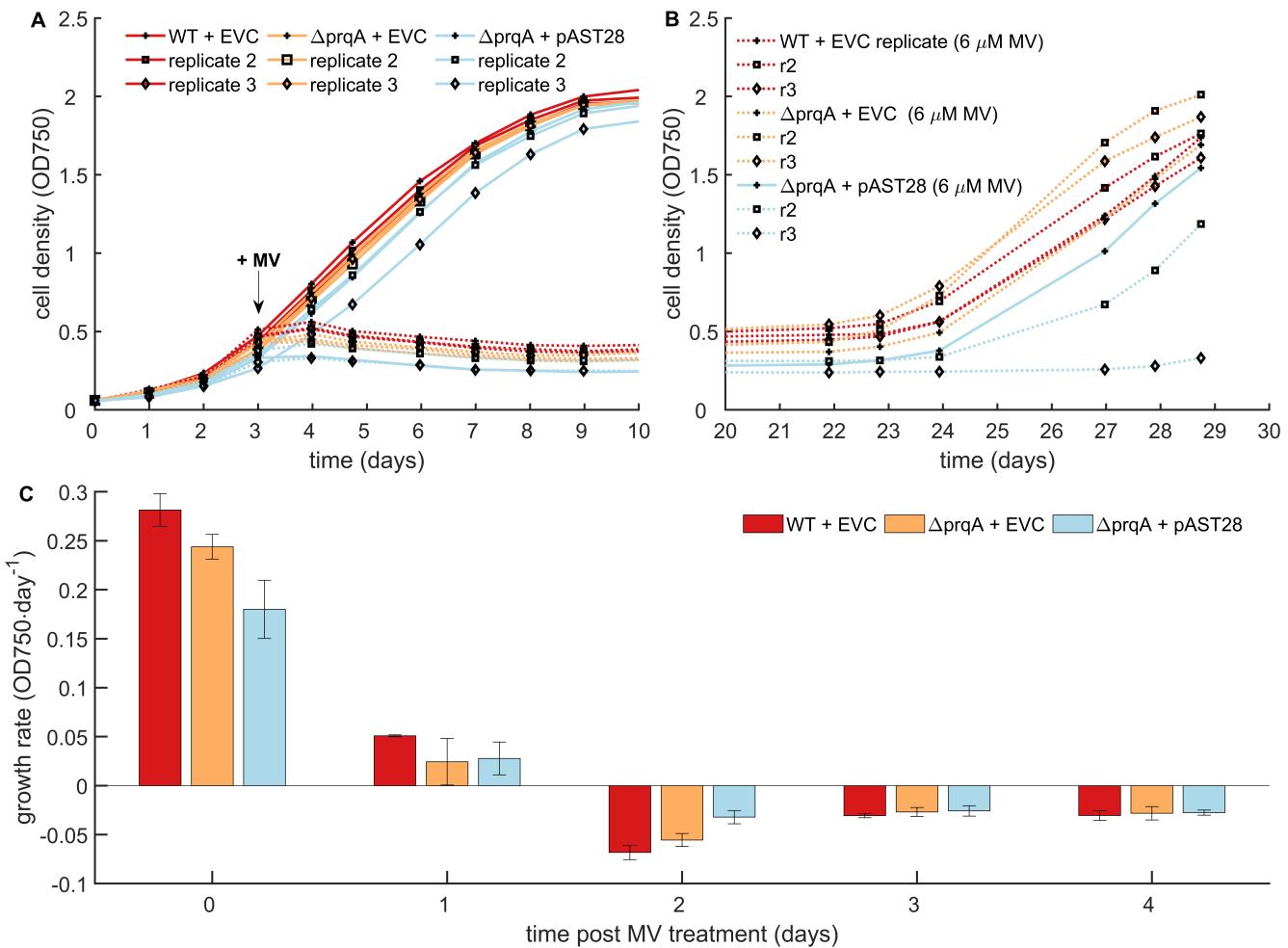


Figure 7: Growth curves of wild-type, Δ prqA and complemented Δ prqA strains of *Synechocystis* sp. PCC 6803 grown at 30 ° under constant illumination (intensity = $100 \mu Em^{-2}s^{-1}$) in the absence (green trace) and upon addition of $6 \mu M$ of methyl viologen during exponential growth (purple trace).

As demonstrated in Figure 7, the presence, absence or heterologous expression of PrqA, did not affect the spontaneous evolution of methyl viologen resistance. This suggests that PrqA is not necessary to develop resistance against MV and therefore may not be involved in its extrusion.

7 Whole genome sequencing and bioinformatic analyses

The genomes of wild-type and MV-resistant *Synechocystis* strains were sequenced to elucidate the genetic determinants potentially underpinning MV resistance.

7.1 Genome sequencing QC report



Sample	Concentration (ng/ μ l)	Volume (μ l)	Effective (%)	Error (%)	Q20 (%)	Q30 (%)	GC (%)
WT_Nixon	31.80	92	99.82	0.03	97.50	92.73	55.06
WT_Howe	27.80	98	99.78	0.03	96.75	90.99	46.50
mvR01_Nixon	36.40	95	99.77	0.03	97.43	92.60	52.66
mvR02_Nixon	25.00	97	99.78	0.03	97.34	92.30	47.98
mvR03_Nixon	30.00	101	99.81	0.03	97.65	93.02	48.64
mvR06_Nixon	44.60	98	99.80	0.03	97.19	91.96	51.59
mvR09_Howe	20.60	95	99.78	0.03	97.28	92.16	47.39
mvR10_Howe	20.20	95	99.86	0.03	97.66	93.39	47.83
mvR11_Howe	27.80	92	99.87	0.03	97.55	93.14	46.53
mvR12_Howe	19.60	98	99.86	0.03	97.65	93.31	47.50

Alignment of the trimmed and paired reads from whole genome sequencing experiments was carried out against several reference genome sequences of *Synechocystis* substrains (Figures 9 and 9). 8 shows the results of the type and number of mutations of both wild types as compared to all *Synechocystis* substrain reference genomes available in the NCBI datatase. For both wild-types, the GT-Kazusa reference genome was the one that resulted in the least amount of SNPs.

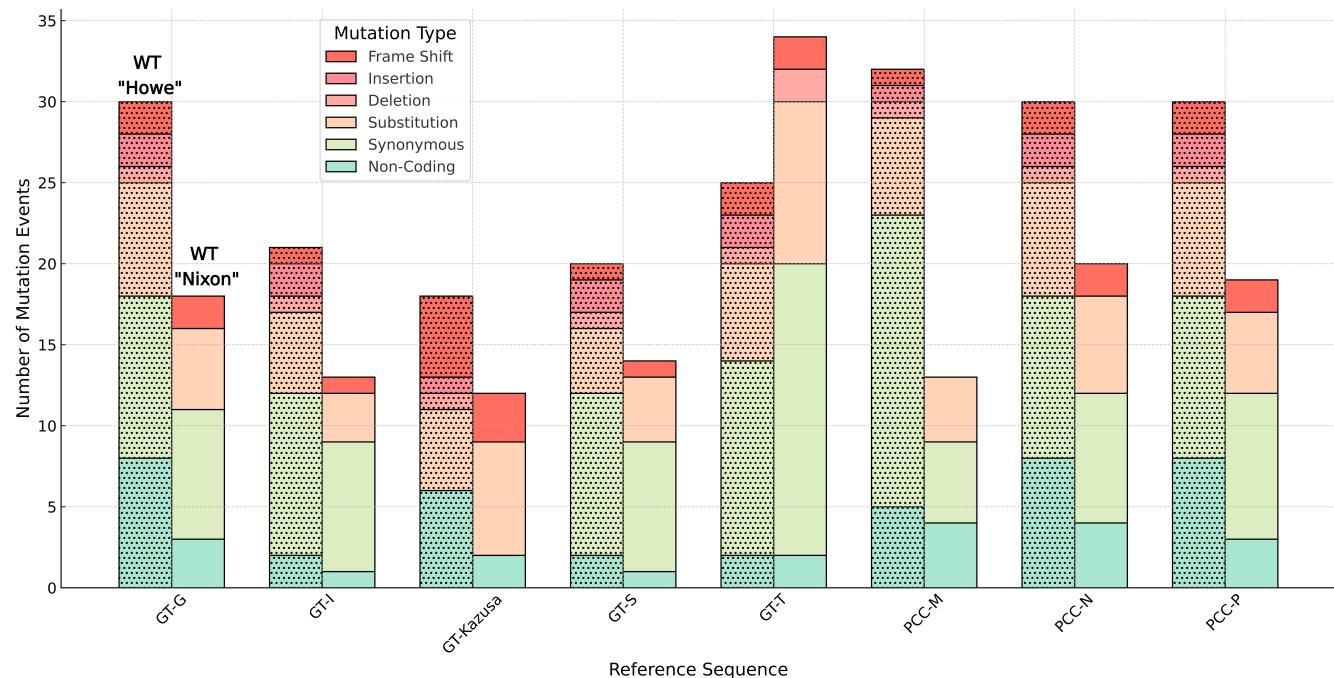


Figure 8: Overview of the variant analysis results.

The mutations were then filtered to identify only non-synonymous ones which results in mutated proteins.

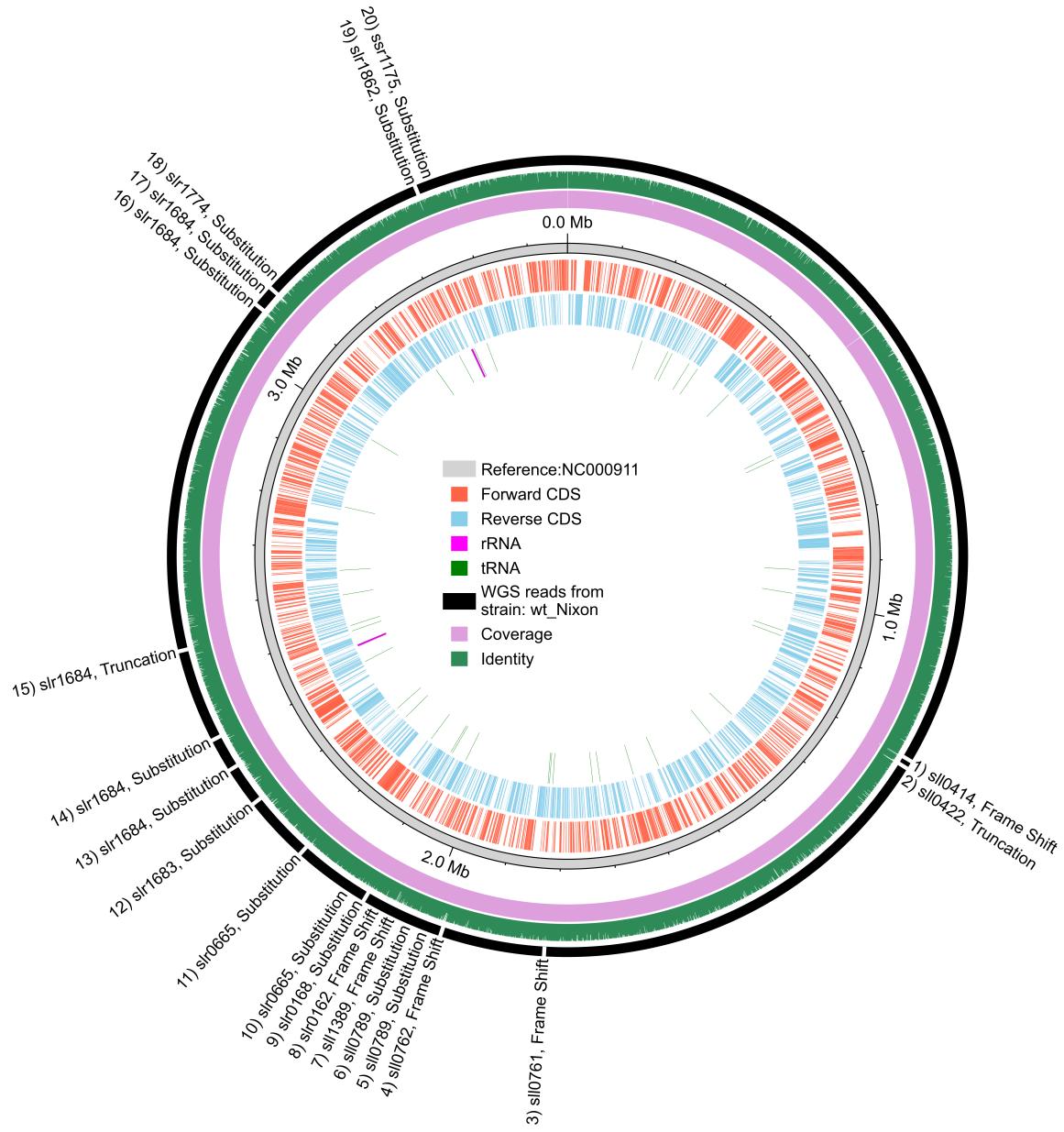


Figure 9: Trimmed and paired reads from whole genome sequencing experiments of *Synechocystis* "Nixon" wild-type aligned to reference "Kazusa" strain (NC000911). Numbered are mutation events found within coding sequences of the reference genome.

For both wild-type strains, the aligned reads covered 100% of the reference genome sequences, indicating that the genomic sequencing was successful. Approximately 20 mutations in coding regions were identified for both the Howe and Nixon wild-type strains. Many mutations were found in transposases, which is unsurprising given the highly evolvable nature of these mobile genetic elements. Many mutations also had moderately small variant frequency values, indicating that such

mutations were either artifacts of sequencing or not fully segregated mutations. Cyanobacteria, have a high degree of poliploidy in their genomes so this might explain why.

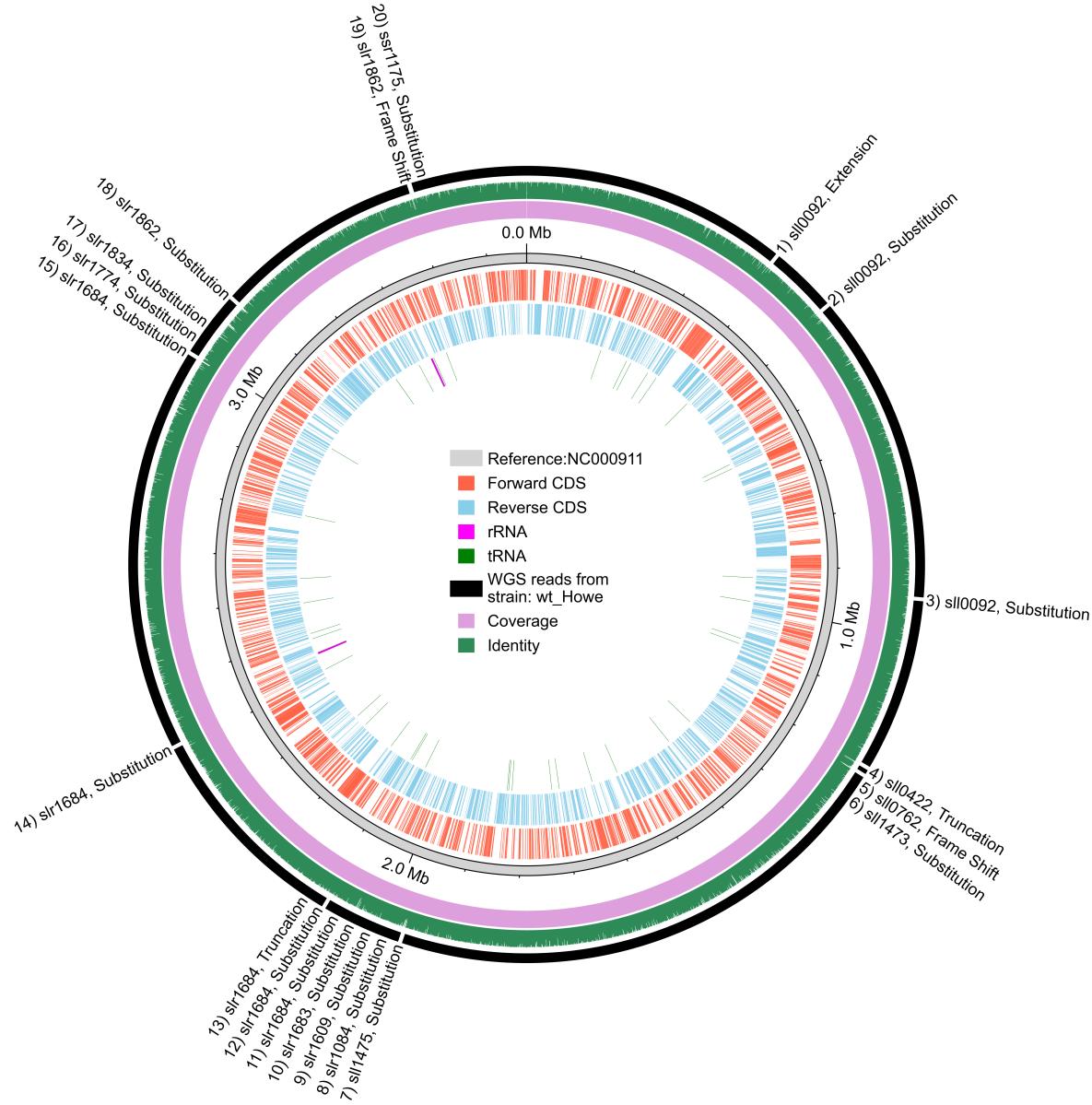


Figure 10: Trimmed and paired reads from whole genome sequencing experiments of *Synechocystis* "Howe" wild-type aligned to reference "Kazusa" strain (NC000911). Numbered are mutation events found within coding sequences of the reference genome.

The Nixon wild-type strain appeared to have a mutation in the type II secretion system F family protein (pilC). Such mutation corresponded to a deletion from (G)9 to (G)8, leading to a Frame Shift. Such mutation has been identified previously as an artefact due to repetitive G nucleotides and

misannotations in the database. Analogously, the Howe substrain exhibited an - at first surprising - substitution in photosystem I. Such mutation has also been identified previously as a misannotation of the reference genome sequence NC000911. In order to identify genotypic differences between the "Howe" and "Nixon" substrains, all variants identified were filtered to retain only the ones that were unique between the two substrains. This was performed for each of the reference sequence that the reads were aligned to. The analysis revealed that both strains have mutations in transposase genes, suggesting potential genomic instability or changes in the frequency of genomic rearrangements. "Nixon" has mutations affecting sugar transport and secretion systems, potentially altering its carbohydrate metabolism and interactions with the environment. "Howe" has mutations affecting motility and respiratory processes, suggesting potential changes in movement and energy metabolism.

	Gene	Product	CDS Position	Change	Protein Effect	Amino Acid Change	Variant Frequency	Reference
1	sll0020	ATP-dependent Clp protease ATP-binding subunit	580	T > G	Substitution	K > Q	0.997	GT-T
2	sll0401	citrate synthase	434	A > C	Substitution	M > R	1	PCC-M
3	sll0414	DUF4335 domain-containing protein	334	+C	Frame Shift		0.955	GT-Kazusa
4	sll0761	hypothetical protein	302	-C	Frame Shift		0.952	GT-Kazusa
5	sll0762	Ig-like domain-containing protein	8924	+C	Frame Shift		0.957	GT-Kazusa
6	sll0771	sugar porter family MFS transporter	85	(C)8 > (C)7	Frame Shift		0.919	GT-T
7	sll0789	two-component system regulator RppA	403	CCA > TCG	Substitution	W > R	0.517 + 0.003	GT-Kazusa, GT-I, GT-S, GT-T, PCC-N, PCC-P, GT-G
8	sll0789	two-component system regulator RppA	397	TTG > GTC	Substitution	Q > D	0.550 + 0.0005	GT-Kazusa, PCC-P, PCC-N, GT-T, GT-S, GT-I, GT-G
9	sll0798	sensor histidine kinase RppB	1339	A > C	Substitution	F > V	0.501	PCC-M
10	sll0798	sensor histidine kinase RppB	1336	T > G	Substitution	I > L	0.516	PCC-M
11	sll1201	Rpn family nuclelease/putative transposase	340	C > A	Substitution	V > L	0.509	PCC-M
12	sll1389	hypothetical protein	430	(G)8 > (G)7	Frame Shift		0.89 + 0.012	GT-Kazusa, PCC-P, PCC-N, PCC-M, GT-T, GT-S, GT-I, GT-G
13	sll1575	serine/threonine-protein kinase	308	+T	Frame Shift		0.972 + 0	PCC-N, PCC-P
14	sll1732	NAD(P)H-quinone oxidoreductase subunit F	1766	C > T	Substitution	R > Q	0.997	GT-T
15	sll1895	EAL domain-containing protein	1134	(T)3 > (T)2	Frame Shift		0.965	GT-G
16	sll1968	anti-sigma regulatory factor	124	T > C	Substitution	K > E	1	GT-I
17	sir0162	type II secretion system F protein	420	(G)9 > (G)8	Frame Shift		0.835	GT-Kazusa
18	sir0168	DUF4114 domain-containing protein	1207	A > G	Substitution	K > E	1	GT-Kazusa
19	sir0611	solanoyl diphosphate synthase	205	+ACGGCG	Insertion	-> TA	0.554 + 0	GT-G, GT-I, GT-S, GT-T, PCC-N, PCC-P
20	sir0665	bifunctional aconitate hydratase	31	G > A	Substitution	D > N	0.515 + 0	GT-Kazusa, PCC-P, PCC-N, GT-T, GT-S, GT-I, GT-G
21	sir0665	bifunctional aconitate hydratase	26	T > C	Substitution	V > A	0.518 + 0	GT-Kazusa, PCC-P, PCC-N, GT-T, GT-S, GT-I, GT-G
22	sir0667	IS5 family transposase	443	+CATGGA	Insertion	G > VHG	0.836 + 0.01	GT-Kazusa, PCC-P, PCC-N, PCC-M, GT-T, GT-S, GT-I, GT-G
23	sir1085	glycosyltransferase family 4 protein	67	-GAACT GTCCATC	Deletion	ELSI >	0.74 + 0.043	GT-Kazusa, PCC-P, PCC-N, PCC-M, GT-T, GT-S, GT-I, GT-G
24	sir2031	PP2C family protein-serine/threonine phosphatase	25	A > T	Substitution	S > C	1 + 0	GT-G, PCC-N, PCC-P
25	sir2031	PP2C family protein-serine/threonine phosphatase	27	CTT > AAA	Substitution	SL > RK	0.645 + 0.005	GT-G
26	sir2031	PP2C family protein-serine/threonine phosphatase	42	T > A	Substitution	D > E	0.6 + 0	GT-G, PCC-N, PCC-P
27	sir2122	acylneuraminate cytidylyltransferase	318	C > G	Substitution	S > R	1	PCC-M
28	ssl2749	nucleotidyltransferase family protein	74	T > C	Substitution	Q > R	0.534	PCC-M

Table 1: Unique mutations found in WT "Nixon"

Interestingly the "Nixon" substrains harbours a deletion in the glycosyltransferase family 4 protein, and a frame shift in the sugar porter family MFS transporter protein (sll0771).

	Gene	Product	CDS Position	Change	Protein Effect	Amino Acid Change	Variant Frequency	Reference
1	sll0092	transposase	232	A -> G	Substitution	Y -> H	0.557 + 0	GT-Kazusa, PCC-P, PCC-N, GT-T, GT-S, GT-I, GT-G
2	sll0092	transposase	205	G -> A	Substitution	P -> S	0.523 + 0	GT-Kazusa, GT-G, GT-I
3	sll0092	transposase	269	T -> C	Extension		0.614 + 0	GT-Kazusa, PCC-P, PCC-N, GT-T, GT-S, GT-I, GT-G
4	sll0762	hypothetical protein	302	-C	Frame Shift		0.928	GT-Kazusa
5	sll1473	PAS domain S-box protein	1397	G -> A	Substitution	A -> V	0.516	GT-Kazusa
6	sll1475	ATP-binding protein	15	A -> T	Substitution	D -> E	0.5	GT-Kazusa
7	sll1533	type IV pilus twitching motility protein PilT	674	-TGTTGA	Deletion	INK -> K	0.83	GT-T
8	sll1774	Rpn family nuclease/putative transposase	340	TAA -> CAC	Substitution	L -> V	0.552	PCC-M
9	sll0611	solanosyl diphosphate synthase	205	+CGGCG	Frame Shift		0.503 + 0	GT-G, GT-I, GT-S, GT-T, PCC-N, PCC-P
10	sll09240	UPF0175 family protein	200	C -> A	Truncation		0.508	PCC-M
11	sll1084	hypothetical protein	230	T -> A	Substitution	V -> D	0.547 + 0	GT-Kazusa, GT-S, GT-T
12	sll1274	type IV pilus assembly protein PilM	266	T -> A	Substitution	V -> E	1	GT-T
13	sll1510	phosphate acyltransferase PlsX	263	G -> T	Substitution	G -> V	0.998	PCC-N
14	sll1609	AMP-binding protein	764	T -> G	Substitution	F -> C	1 + 0	GT-Kazusa, PCC-P, PCC-N, GT-T, GT-S, GT-I, GT-G
15	sll1712	Rpn family nuclease/putative transposase	559	CTG -> ATT	Substitution	L -> I	0.611	PCC-M
16	sll1712	Rpn family nuclease/putative transposase	550	G -> T	Substitution	A -> S	0.641	PCC-M
17	sll1753	CHAT domain-containing protein	3013	A -> C	Substitution	T -> P	0.521	PCC-M
18	sll1834	photosystem I core protein PsA	1810	G -> A	Substitution	V -> I	0.994	GT-Kazusa
19	sll1862	hypothetical protein	454	+C	Frame Shift		0.538	GT-Kazusa
20	sll2031	PP2C family protein-serine/threonine phosphatase	25	AGC -> TGG	Substitution	S -> W	0.528 + 0.002	GT-G, PCC-N, PCC-P
21	sll2031	PP2C family protein-serine/threonine phosphatase	29	T -> A	Truncation		0.543 + 0	GT-G, PCC-N, PCC-P
22	sll2031	PP2C family protein-serine/threonine phosphatase	42	T -> CGA	Frame Shift		0.51 + 0	GT-G, PCC-N, PCC-P
23	ssl0172	transposase	205	G -> A	Substitution	P -> S	0.523 + 0	GT-S, GT-T, PCC-N, PCC-P

Table 2: Unique mutations found in WT "Howe"

In the "Howe" substrain a deletion in the type IV pilus twitching motility protein PilT and a substitution in in PilM were indentified. These might underlye the non-motile phenotype of the "Howe" substrains. The genetic differences identified between the Nixon and Howe substrains provide a foundation for understanding their distinct phenotypic behaviors. Nixon's mutations in sugar transport and metabolism genes align with its inability to grow in the presence of sugar in light conditions. In contrast, Howe's mutations related to motility and respiratory processes correspond with its non-motile phenotype and capability to grow photoheterotrophically.

7.2 Sequenced strains compared to various substrains

To visualise the SNPs similarity between various strains, a principal component analysis (PCA) was performed, where the feature matrix was constructed to represent each strain as a row and each reference genome as a column. The entries in this matrix were the total SNP counts for each strain against each reference genome. Thus, the PCA effectively reduced the multidimensional space spanned by the different reference genomes into a lower-dimensional representation. The PCA plot positions each strain based on its SNP profile across different reference genomes. Strains that are genetically similar in terms of SNP counts against multiple reference genomes would appear closer in the reduced-dimensional space.

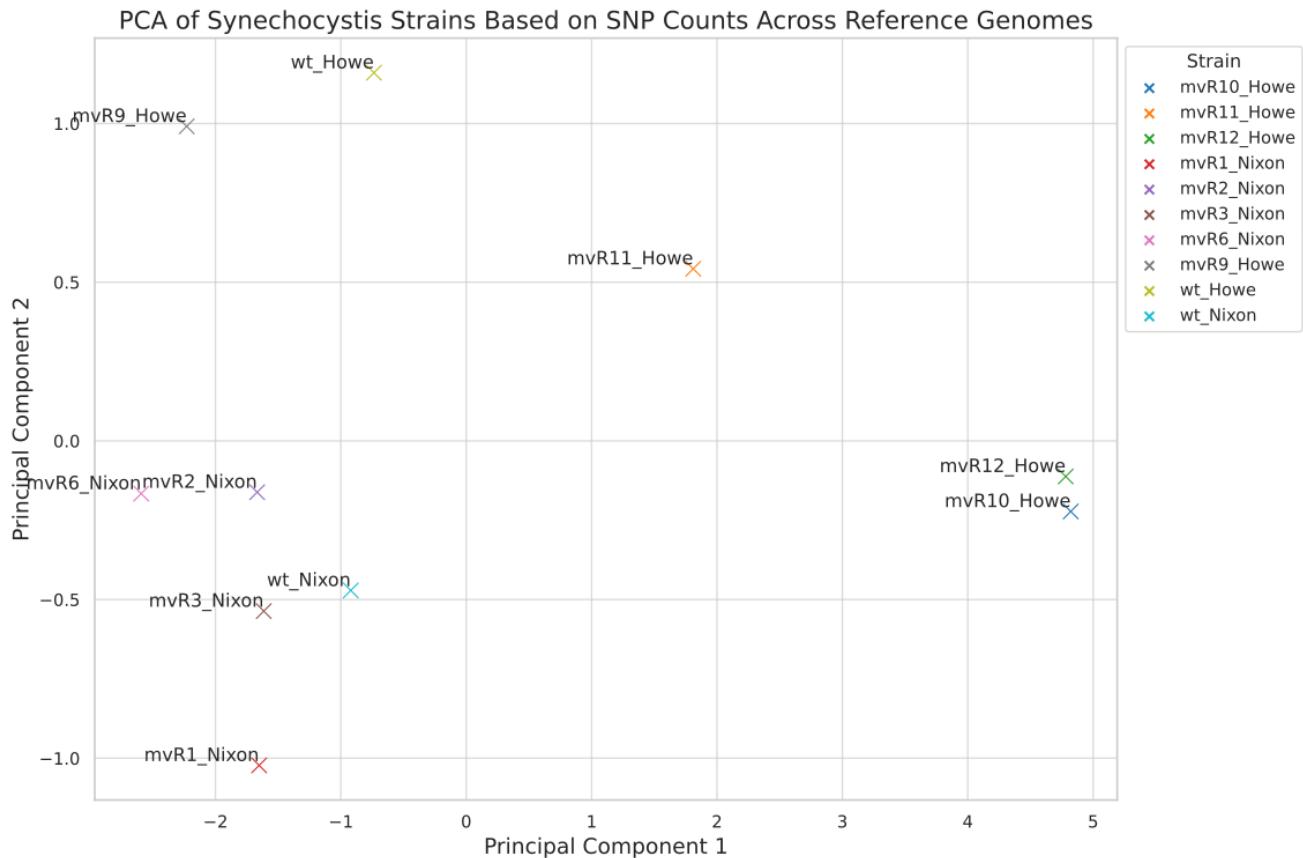


Figure 11: Principal component analysis based on SNPs count of the sequenced genomes

Principal Component 1 (PC1) captured approximately 88.22% of the total variance in the dataset. Principal Component 2 (PC2) accounted for about 5.37% of the total variance. Cumulative Explained Variance: Together, the first two principal components capture approximately 93.58% of the total variance. The first principal component (PC1) alone captures a significant proportion of the dataset's variance, indicating an effective reduction in dimensionality. The high cumulative explained variance (93.58%) implies that the first two principal components provide a statistically robust representation of the dataset's original variability. Therefore PCA analysis indicated that strains with the "Howe" and "Nixon" suffixes tend to cluster with their respective wild-types, corroborating the genomic closeness between mutants and their parental strains.

7.3 SNPs Polimorphism

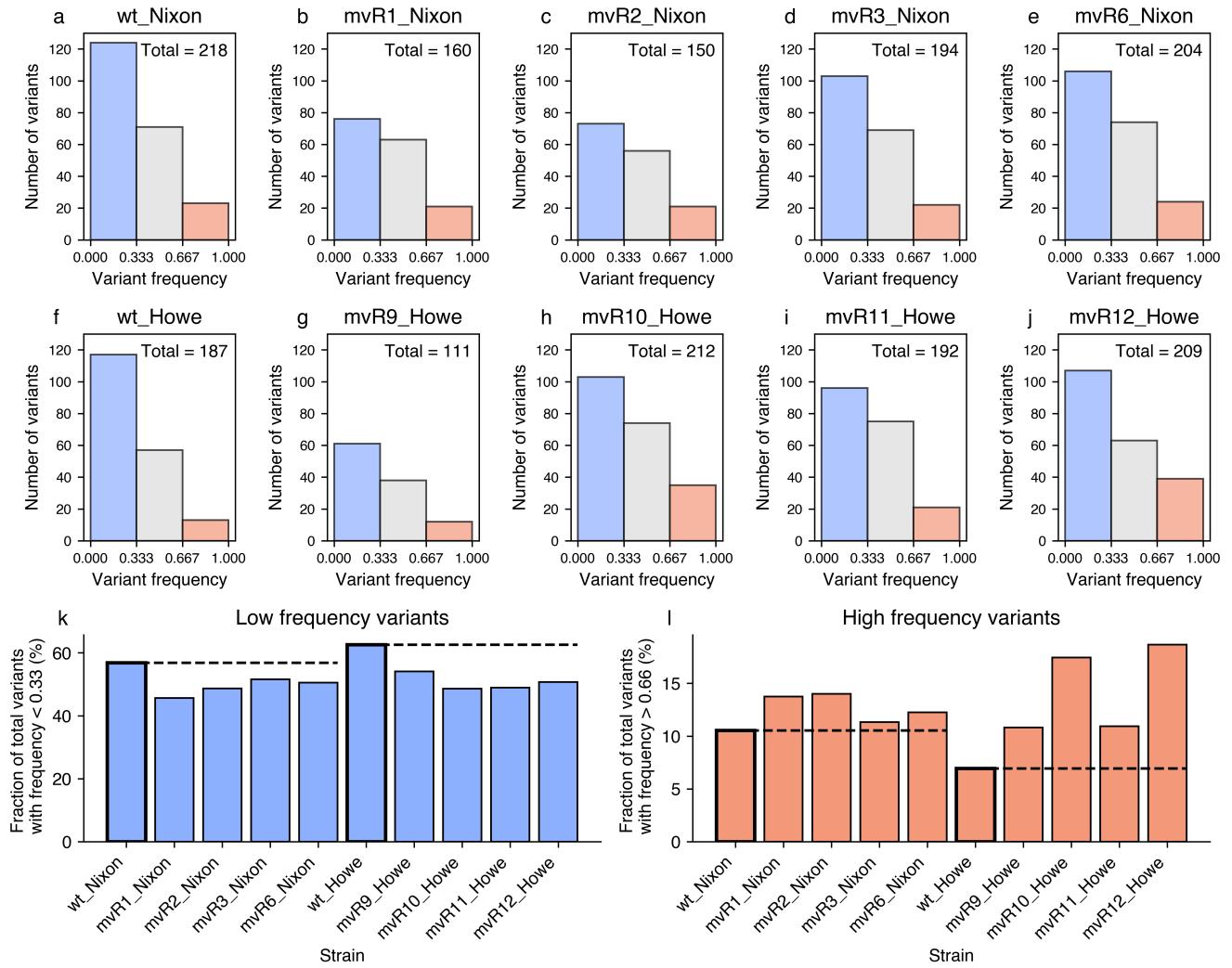


Figure 12: 3 bins, 0.3 and 0.6 as thresholds

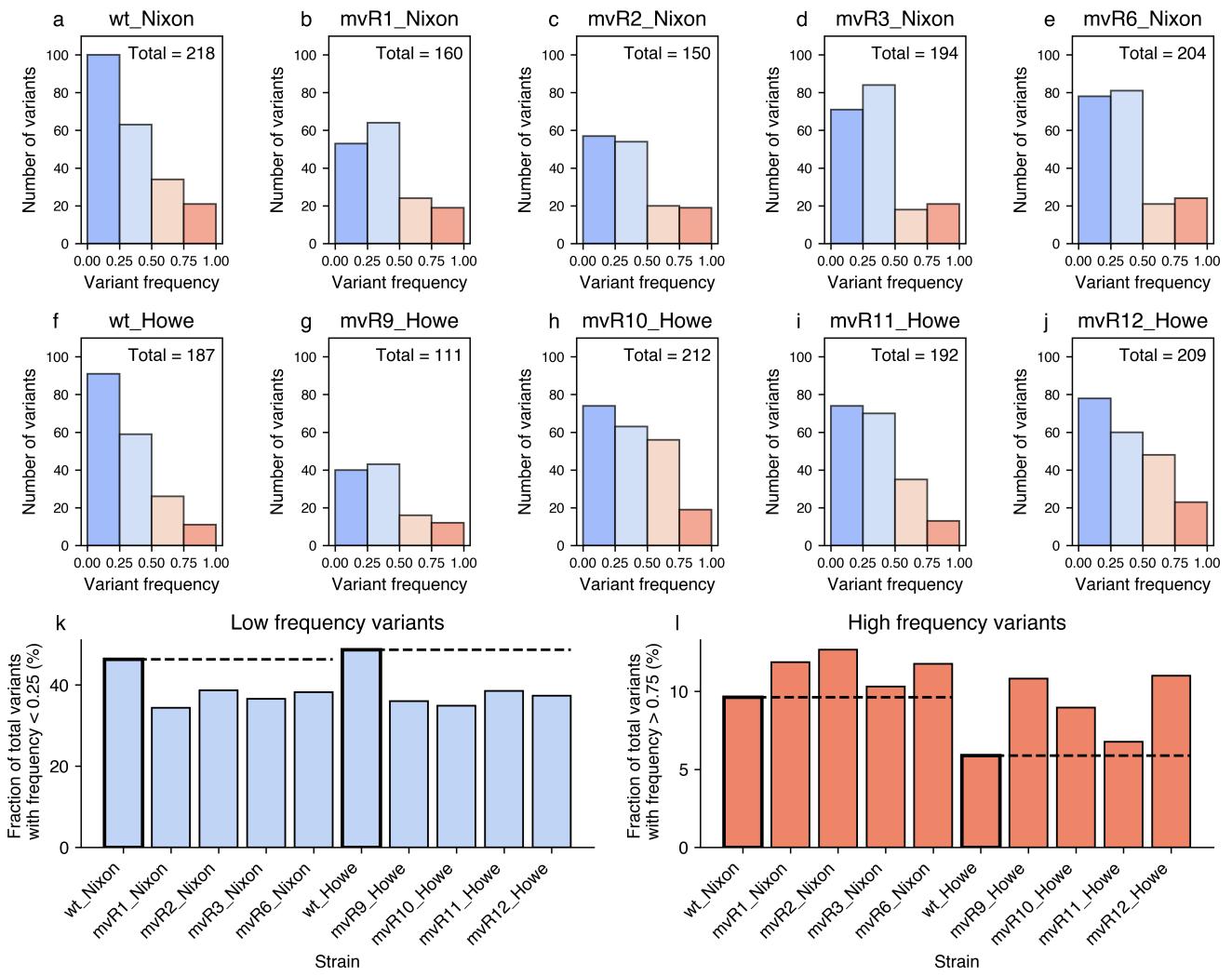


Figure 13: 4 bins, 0.25 and 0.75 as thresholds

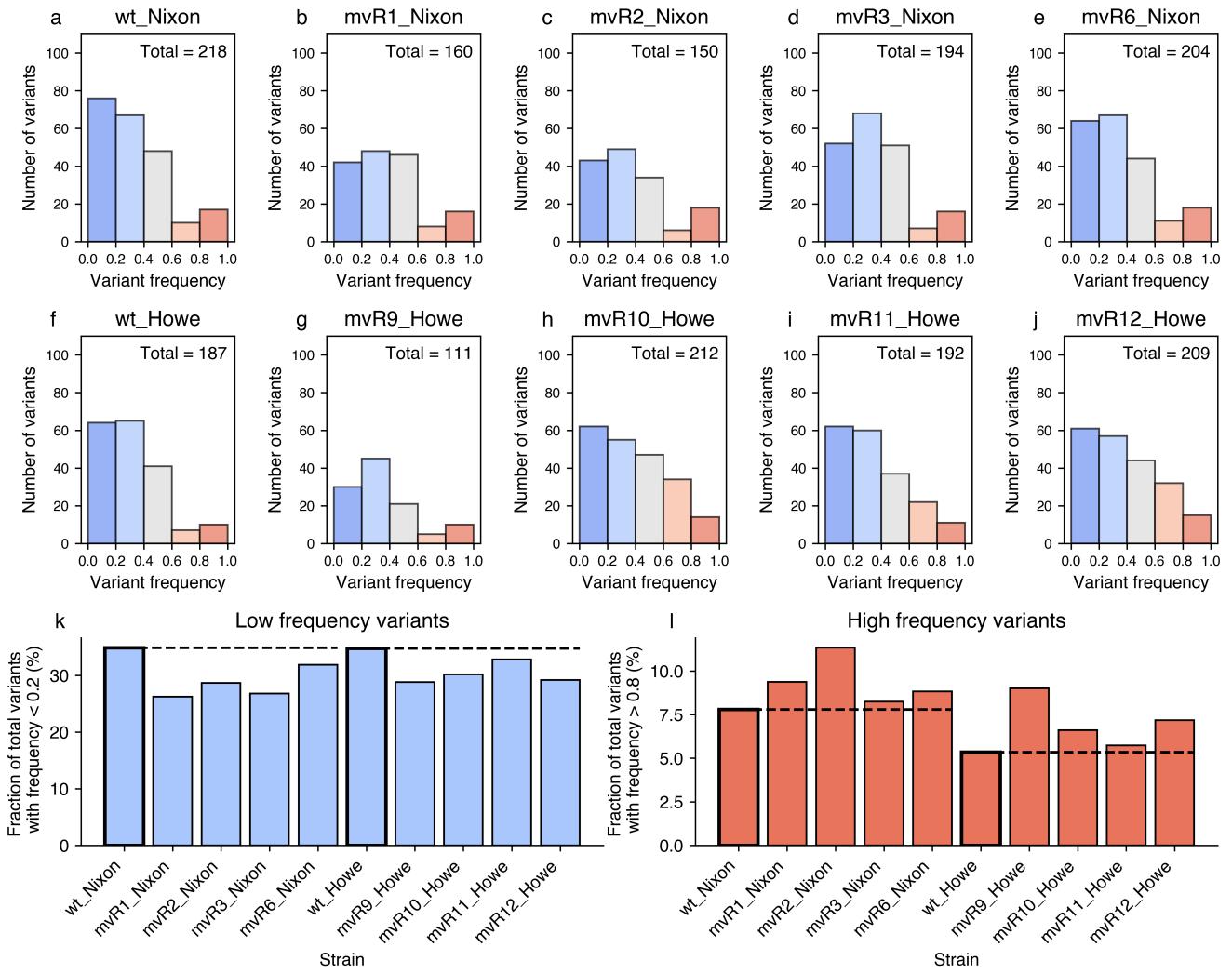


Figure 14: 5 bins, 0.2 and 0.8 as thresholds