**Simulated Dataset results:**

**Table: Summarized results of simulated reads annotation.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sample\_ID** | **Contigs** | **Bases** | **CDS** | **rRNA** | **tRNA** | **tmRNA** |
| 24820\_2209 | 36 | 2116537 | 2117 | 5 | 50 | 1 |
| 24820\_2267 | 59 | 2173452 | 2212 | 5 | 46 | 1 |
| 11822\_158 | 27 | 2090334 | 2088 | 6 | 48 | 1 |
| 15682\_350 | 36 | 2188217 | 2191 | 3 | 46 | 1 |

**Table 2: Summary of Panaroo results for simulated sequences**

|  |  |
| --- | --- |
| **Types of Gene** | **Counts** |
| Core genes (99% <= strains <= 100%) | 1879 |
| Soft core genes (95% <= strains < 99%) | 0 |
| Shell genes (15% <= strains < 95%) | 526 |
| Cloud genes (0% <= strains < 15%) | 0 |
| Total genes (0% <= strains <= 100%) | 2405 |

**Table 2: Summary of Snippy results for simulated sequences**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sample\_ID** | **Complex Variants** | **Deletions**  **(DEL)** | **Insertions**  **(INS)** | **Multiple Nucleotide Polymorphism (MNP)** | **Single Nucleotide Polymorphism (SNP)** | **Total Variants** |
| 24820\_2209 | 881 | 34 | 28 | 1 | 5128 | 6072 |
| 24820\_2267 | 839 | 27 | 22 | - | 4950 | 5838 |
| 11822\_158 | 258 | 34 | 23 | - | 1544 | 1859 |
| 15682\_350 | 816 | 29 | 28 | 1 | 4814 | 5688 |

**Table 3: Summary of desired genes detected in simulated reads.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Gene Name | Position | Reference | Alternate | Quality | Depth (DP) | Allele Count (AO) | Allele Quality (QA) | Genotype Likelihood (GL) |
| gla | 52 | ATT | TAA | 370.63 | 11 | 11 | 436 | -39.5964,-3.31133,0 |
| rsmG | 53 | CGG | AAA | 475.44 | 14 | 14 | 553 | -50.1157,-4.21442,0 |
| lemA | 52 | GTT | TAA | 402.593 | 12 | 12 | 474 | -43.0126,-3.61236,0 |
| pgsA | 52 | AT | TA | 450.011 | 14 | 14 | 546 | -49.4851,-4.21442,0 |

**Table : Details of detected premature stop codons in simulated reads**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sample\_ID** | **Gene** | **Premature Stop Codon Position** | **Product** | **Sequence** |
| 24820\_2209 | lemA | 52 | Protein lemA | ATGACTTGGATTATTCTTGGAGTTATCGCTCTTATTGTTATTTTTGTGATT**TAA**AGCTATAACGGTTTGGTTAAAAATCGTATGCAAACCAAGGAGGCTTGGAGTCAGATTGATGTTCAGTTGAAACGTCGAAATGACCTCTTGCCGAACTTGATTGAGACTGTAAAAGGTTATGCCAAATATGAAGGTTCTACCCTTGAAAAGGTGGCAGAACTACGTAACCAAGTGGCGGCAGCGACTTCACCAGCAGAAGCTATGAAAGCCAGTGATGCCCTCACTCGTCAGGTTTCAGGTATTTTTGCAGTTGCAGAAAGCTATCCAGATTTGAAAGCTAGTGCTAACTTTGTTAAATTGCAAGAGGAGTTGACAAACACAGAAAATAAAATTTCTTACTCTCGTCAACTCTATAACAGTGTTGTCAGCAACTACAATGTAAAATTAGAAACTTTCCCGAGCAATATTATCGCTGGAATGTTTGGATTTAAAGCGGCAGATTTCCTTCAAACACCTGAAGAGGAAAAGTCGGTTCCTAAAGTTGATTTTAGCGGTTTAGGTGACTAA |
| 24820\_2267 | pgsA | 52 | CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase | ATGAAAAAAGAACAAATTCCCAATCTCTTAACAATAGGTCGAATTCTCTTT**TAA**CCTATTTTTATCTTTATTTTAACGATAGGAAATTCGATAGAGAGTCATATAGTTGCAGCTATTATCTTTGCTGTTGCCAGTATTACCGACTATTTAGATGGATATTTAGCTCGTAAATGGAATGTGGTCAGTAATTTTGGTAAATTTGCAGATCCTATGGCGGATAAGTTACTAGTTATGTCGGCTTTTATTGTGTTGATTGAGTTAGGTATGGCTCCGGCTTGGATTGTTGCAGTGATTATCTGTCGTGAGTTAGCTGTGACAGGTTTAAGGCTTTTATTGGTTGAAACTGGTGGAACAATTTTAGCAGCAGCAATGCCTGGAAAAATTAAAACTTTTAGTCAGATGTTTGCCATTATTTTCTTGCTATTACATTGGACTTTGCTTGGTCAAGTTCTACTTTATGTAGCCTTATTTTTCACTATCTACTCTGGCTATGACTATTTCAAGGGTAGTGCCTATGTATTTAAAGGGACATTTGGTTCGAAATGA |
| 11822\_158 | gla | 52 | Glycerol facilitator-aquaporin | ATGGATTTCACATGGGCACTGAAGTATGCCACTGAATTTTTGGGAACTGCC**TAA**TTGATCATTCTTGGGAATGGTGCAGTTGCCAACGTTGAACTTAAAGGTACGAAAGGTCACCAAAGTGGCTGGATCGTCATCGCTGTTGGTTATGGTATGGGGGTTATGATCCCTGCCTTGATGTTTGGTAACGTATCTGGGAATCACATCAACCCTGCTTTCACTCTAGGGCTTGCAGTTAGCGGTCTTTTCCCTTGGGCACAAGTGGTACCTTATATTATCGCGCAAGTCTTGGGGGCTATCTTTGGCCAAGCCTTAGTTGTGGCAACATACCGTCCATTCTACTTGAAAACTGAAAACCCAAATAACATCTTGGGAACTTTCTCAACTATTTCAAGTATTGACCATGGTACAAAAGAAAGTCGCTATGCAGCAACTGTCAATGGTTTGATTAATGAGTTTGTTGGTTCATTTGTTTTGTTCTTTGCAGCTCTTGGTTTGACTAAAAACTTCTTTGGTGCTGAAGTGCTTCAATTCATGAAACAAAAGGCAATAGAAGCAGGACAAACAGTTGATTTTTCTGACTTGGCTATTAAAGCACAGGTGGCTCCACACACTGCTTCAGGATTTTCTGTGGCTCACTTGGCACTTGGATTCCTCGTTATGGCTTTGGTAACATCACTTGGAGGACCTACAGGACCTGCCTTGAACCCAGCCCGTGACTTGGGACCACGTCTCCTTCATGCTTTCCTTCCAAAATCAGTTCTTGGTGAGCATAAAGGCGATTCAAAATGGTGGTATTCTTGGGTACCAGTAGTAGCACCTATCGCAGCAGCAATTGCGGCAATAGCTGTATTCAAATTCCTTTATCTCTAA |
| 15682\_350 | rsmG | 52 | Ribosomal RNA small subunit methyltransferase G | ATGAAACCAGAAACATTTTACAACTTGCTTGCCGAGCAGAATCTTCCACTT**TAA**AACCAGCAAAAAGAACAATTTGAACGTTATTTTGAGCTCTTGGTCGAGTGGAATGAGAAGATTAATTTGACGGCGATTACGGACAAGGAAGAAGTTTATCTCAAACATTTTTACGATTCGATTGCACCCATTCTTCAAGGTTTGATTCCCAATGAACCTATCAAACTTCTTGATATCGGGGCTGGGGCAGGATTTCCTAGTCTACCAATGAAAATTCTCTATCCGGAGTTAGATGTGACCATTATTGATTCACTCAATAAGCGCATCAACTTCCTACAACTCTTGGCTCAAGAACTGGATTTGAACGGAGTTCATTTCTACCACGGACGTGCCGAAGATTTTGCCCAAGACAAGAACTTCCGTGCTCAATATGATTTTGTAACAGCTCGTGCGGTTGCCCGTATGCAGGTCCTATCTGAATTGACTATTCCCTACCTTAAGGTTGGTGGCAAACTATTGGCACTCAAGGCCAGCAATGCGCCTGAGGAATTATTAGAAGCTAAGAATGCCCTCAATCTCCTTTTTAGTAAGGTCGAAGACAATCTCAGCTACGCCCTACCGAATAGAGATCCGCGCTATATCACAGTGGTAGAAAAGAAAAAAGAAACACCAAATAAATATCCACGTAAGGCTGGTATGCCAAATAAACGCCCACTTTAA |