

COVID-19 subject 219

2021-04-30

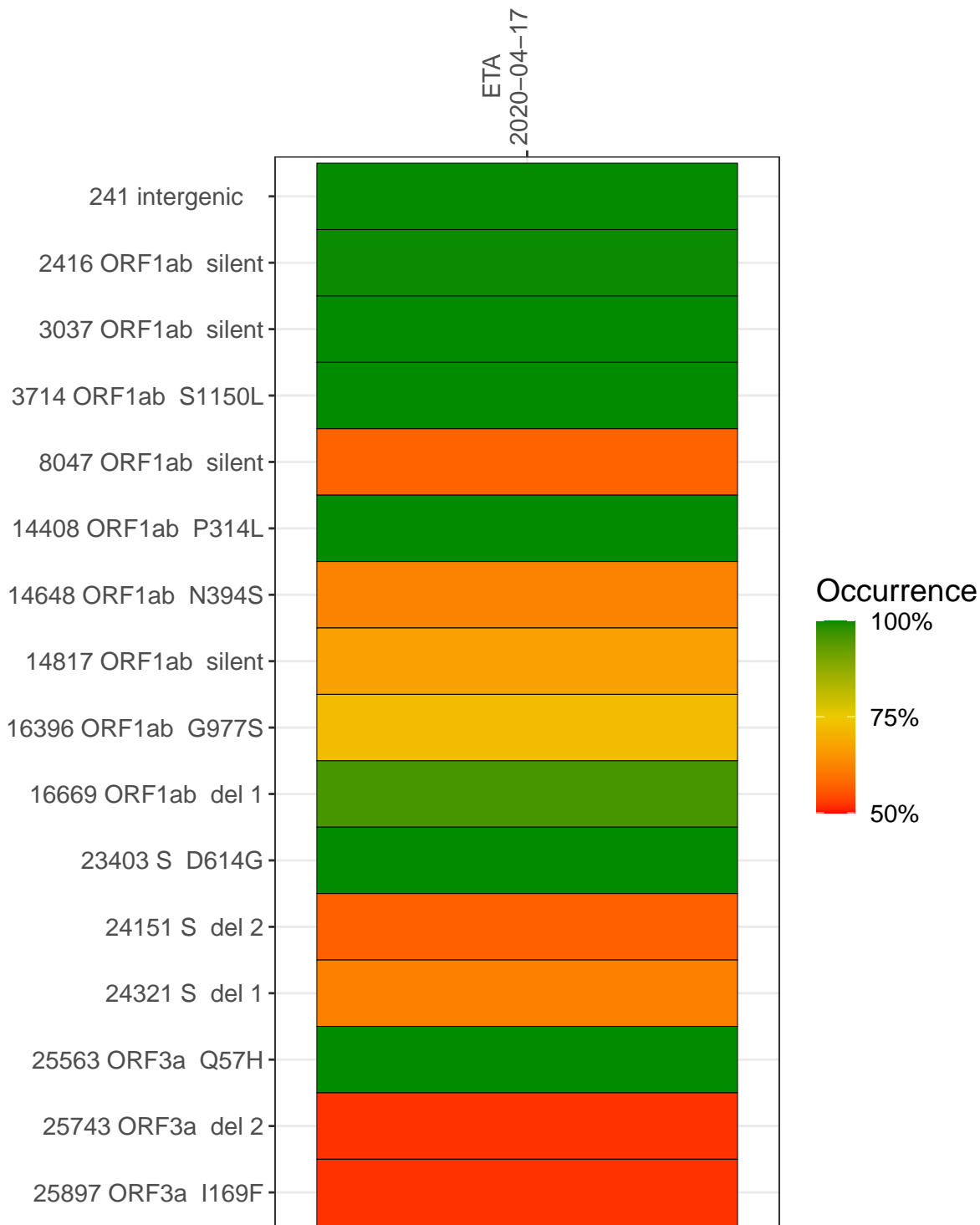
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. Lineages are called with the Pangolin software tool (Rambaut et al 2020) for genomes with > 90% sequence coverage.

Table 1. Sample summary.

Experiment	Type	Genomes	Sample type	Sample date	Largest contig (KD)	Lineage	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0017	composite	NA	ETA	2020-04-17	7.66	NA	93.4%	92.3%
VSP0017-1m	single experiment	NA	ETA	2020-04-17	0.93	NA	39.4%	35.2%
VSP0017-2m	single experiment	NA	ETA	2020-04-17	1.23	NA	47.4%	43.7%
VSP0017-3	single experiment	9350	ETA	2020-04-17	1.53	NA	48.9%	43.1%
VSP0017-4	single experiment	1870	ETA	2020-04-17	1.06	NA	49.3%	47.3%
VSP0017-5	single experiment	1870	ETA	2020-04-17	1.69	NA	44.1%	39.8%
VSP0017-6	single experiment	1870	ETA	2020-04-17	3.40	NA	64.3%	62.3%

Variants shared across samples

The heat map below shows how variants (reference genome /home/everett/projects/SARS-CoV-2-Philadelphia/Wuhan-Hu-1) are shared across subject samples where the percent variance is colored. Variants are called if a variant position is covered by 5 or more reads, the alternative base is found in > 50% of read pairs and the variant yields a PHRED score > 20. Gray tiles denote positions where the variant was not the major variant or no variants were found. The relative base compositions of each experiment used to calculate tiles are shown in the following plot where the total number of position reads are shown atop of each plot.



ETA 2020-04-17						
ORF3ε ORF3α ORF3β ORF3γ ORF3δ ORF3ε ORF3ζ ORF3η ORF3θ ORF3ι ORF3κ ORF3λ ORF3μ ORF3ν ORF3ξ ORF3ο ORF3π ORF3ρ ORF3σ ORF3τ ORF3υ ORF3φ ORF3χ ORF3ψ ORF3ω ORF3x ORF3y ORF3z ORF3aa ORF3ab ORF3ac ORF3ad ORF3ae ORF3af ORF3ag ORF3ah ORF3ai ORF3aj ORF3ak ORF3al ORF3am ORF3an ORF3ao ORF3ap ORF3aq ORF3ar ORF3as ORF3at ORF3au ORF3av ORF3aw ORF3ax ORF3ay ORF3az ORF3ba ORF3bb ORF3bc ORF3bd ORF3be ORF3bf ORF3bg ORF3bh ORF3bi ORF3bj ORF3bk ORF3bl ORF3bm ORF3bn ORF3bo ORF3bp ORF3bq ORF3br ORF3bs ORF3bt ORF3bu ORF3bv ORF3bw ORF3bx ORF3by ORF3bz ORF3ca ORF3cb ORF3cc ORF3cd ORF3ce ORF3cf ORF3cg ORF3ch ORF3ci ORF3cj ORF3ck ORF3cl ORF3cm ORF3cn ORF3co ORF3cp ORF3cq ORF3cr ORF3cs ORF3ct ORF3cu ORF3cv ORF3cw ORF3cx ORF3cy ORF3cz ORF3da ORF3db ORF3dc ORF3dd ORF3de ORF3df ORF3dg ORF3dh ORF3di ORF3dj ORF3dk ORF3dl ORF3dm ORF3dn ORF3do ORF3dp ORF3dq ORF3dr ORF3ds ORF3dt ORF3du ORF3dv ORF3dw ORF3dx ORF3dy ORF3dz ORF3ea ORF3eb ORF3ec ORF3ed ORF3ee ORF3ef ORF3eg ORF3eh ORF3ei ORF3ej ORF3ek ORF3el ORF3em ORF3en ORF3eo ORF3ep ORF3eq ORF3er ORF3es ORF3et ORF3eu ORF3ev ORF3ew ORF3ex ORF3ey ORF3ez ORF3fa ORF3fb ORF3fc ORF3fd ORF3fe ORF3ff ORF3fg ORF3fh ORF3fi ORF3fj ORF3fk ORF3fl ORF3fm ORF3fn ORF3fo ORF3fp ORF3fq ORF3fr ORF3fs ORF3ft ORF3fu ORF3fv ORF3fw ORF3fx ORF3fy ORF3fz ORF3ga ORF3gb ORF3gc ORF3gd ORF3ge ORF3gf ORF3gg ORF3gh ORF3gi ORF3gj ORF3gk ORF3gl ORF3gm ORF3gn ORF3go ORF3gp ORF3gq ORF3gr ORF3gs ORF3gt ORF3gu ORF3gv ORF3gw ORF3gx ORF3gy ORF3gz ORF3ha ORF3hb ORF3hc ORF3hd ORF3he ORF3hf ORF3hg ORF3hi ORF3hj ORF3hk ORF3hl ORF3hm ORF3hn ORF3ho ORF3hp ORF3hq ORF3hr ORF3hs ORF3ht ORF3hu ORF3hv ORF3hw ORF3hx ORF3hy ORF3hz ORF3ia ORF3ib ORF3ic ORF3id ORF3ie ORF3if ORF3ig ORF3ih ORF3ii ORF3ij ORF3ik ORF3il ORF3im ORF3in ORF3io ORF3ip ORF3iq ORF3ir ORF3is ORF3it ORF3iu ORF3iv ORF3iw ORF3ix ORF3iy ORF3iz ORF3ja ORF3jb ORF3jc ORF3jd ORF3je ORF3jf ORF3jg ORF3jh ORF3ji ORF3jj ORF3jk ORF3jl ORF3jm ORF3jn ORF3jo ORF3jp ORF3jq ORF3jr ORF3js ORF3jt ORF3ju ORF3jv ORF3jw ORF3jx ORF3jy ORF3jz ORF3ka ORF3kb ORF3kc ORF3kd ORF3ke ORF3kf ORF3kg ORF3kh ORF3ki ORF3kj ORF3kk ORF3kl ORF3km ORF3kn ORF3ko ORF3kp ORF3kq ORF3kr ORF3ks ORF3kt ORF3ku ORF3kv ORF3kw ORF3kx ORF3ky ORF3kz ORF3la ORF3lb ORF3lc ORF3ld ORF3le ORF3lf ORF3lg ORF3lh ORF3li ORF3lj ORF3lk ORF3ll ORF3lm ORF3ln ORF3lo ORF3lp ORF3lq ORF3lr ORF3ls ORF3lt ORF3lu ORF3lv ORF3lw ORF3lx ORF3ly ORF3lz ORF3ma ORF3mb ORF3mc ORF3md ORF3me ORF3mf ORF3mg ORF3mh ORF3mi ORF3mj ORF3mk ORF3ml ORF3mm ORF3mn ORF3mo ORF3mp ORF3mq ORF3mr ORF3ms ORF3mt ORF3mu ORF3mv ORF3mw ORF3mx ORF3my ORF3mz ORF3na ORF3nb ORF3nc ORF3nd ORF3ne ORF3nf ORF3ng ORF3nh ORF3ni ORF3nj ORF3nk ORF3nl ORF3nm ORF3nn ORF3no ORF3np ORF3nq ORF3nr ORF3ns ORF3nt ORF3nu ORF3nv ORF3nw ORF3nx ORF3ny ORF3nz ORF3oa ORF3ob ORF3oc ORF3od ORF3oe ORF3of ORF3og ORF3oh ORF3oi ORF3oj ORF3ok ORF3ol ORF3om ORF3on ORF3oo ORF3op ORF3oq ORF3or ORF3os ORF3ot ORF3ou ORF3ov ORF3ow ORF3ox ORF3oy ORF3oz ORF3pa ORF3pb ORF3pc ORF3pd ORF3pe ORF3pf ORF3pg ORF3ph ORF3pi ORF3pj ORF3pk ORF3pl ORF3pm ORF3pn ORF3po ORF3pp ORF3pq ORF3pr ORF3ps ORF3pt ORF3pu ORF3pv ORF3pw ORF3px ORF3py ORF3pz ORF3qa ORF3qb ORF3qc ORF3qd ORF3qe ORF3qf ORF3qg ORF3qh ORF3qi ORF3qj ORF3qk ORF3ql ORF3qm ORF3qn ORF3qo ORF3qp ORF3qq ORF3qr ORF3qs ORF3qt ORF3qu ORF3qv ORF3qw ORF3qx ORF3qy ORF3qz ORF3ra ORF3rb ORF3rc ORF3rd ORF3re ORF3rf ORF3rg ORF3rh ORF3ri ORF3rj ORF3rk ORF3rl ORF3rm ORF3rn ORF3ro ORF3rp ORF3rq ORF3rr ORF3rs ORF3rt ORF3ru ORF3rv ORF3rw ORF3rx ORF3ry ORF3rz ORF3sa ORF3sb ORF3sc ORF3sd ORF3se ORF3sf ORF3sg ORF3sh ORF3si ORF3sj ORF3sk ORF3sl ORF3sm ORF3sn ORF3so ORF3sp ORF3sq ORF3sr ORF3ss ORF3st ORF3su ORF3sv ORF3sw ORF3sx ORF3sy ORF3sz ORF3ta ORF3tb ORF3tc ORF3td ORF3te ORF3tf ORF3tg ORF3th ORF3ti ORF3tj ORF3tk ORF3tl ORF3tm ORF3tn ORF3to ORF3tp ORF3tq ORF3tr ORF3ts ORF3tt ORF3tu ORF3tv ORF3tw ORF3tx ORF3ty ORF3tz ORF3ua ORF3ub ORF3uc ORF3ud ORF3ue ORF3uf ORF3ug ORF3uh ORF3ui ORF3uj ORF3uk ORF3ul ORF3um ORF3un ORF3uo ORF3up ORF3uq ORF3ur ORF3us ORF3ut ORF3uu ORF3uv ORF3uw ORF3ux ORF3uy ORF3uz ORF3va ORF3vb ORF3vc ORF3vd ORF3ve ORF3vf ORF3vg ORF3vh ORF3vi ORF3vj ORF3vk ORF3vl ORF3vm ORF3vn ORF3vo ORF3vp ORF3vq ORF3vr ORF3vs ORF3vt ORF3vu ORF3vv ORF3vw ORF3vx ORF3vy ORF3vz ORF3wa ORF3wb ORF3wc ORF3wd ORF3we ORF3wf ORF3wg ORF3wh ORF3wi ORF3wj ORF3wk ORF3wl ORF3wm ORF3wn ORF3wo ORF3wp ORF3wq ORF3wr ORF3ws ORF3wt ORF3wu ORF3wv ORF3ww ORF3wx ORF3wy ORF3wz ORF3xa ORF3xb ORF3xc ORF3xd ORF3xe ORF3xf ORF3xg ORF3xh ORF3xi ORF3xj ORF3xk ORF3xl ORF3xm ORF3xn ORF3xo ORF3xp ORF3xq ORF3xr ORF3xs ORF3xt ORF3xu ORF3xv ORF3xw ORF3xx ORF3xy ORF3xz ORF3ya ORF3yb ORF3yc ORF3yd ORF3ye ORF3yf ORF3yg ORF3yh ORF3yi ORF3yj ORF3yk ORF3yl ORF3ym ORF3yn ORF3yo ORF3yp ORF3yq ORF3yr ORF3ys ORF3yt ORF3yu ORF3yv ORF3yw ORF3yx ORF3yy ORF3yz ORF3za ORF3zb ORF3zc ORF3zd ORF3ze ORF3zf ORF3zg ORF3zh ORF3zi ORF3zj ORF3zk ORF3zl ORF3zm ORF3zn ORF3zo ORF3zp ORF3zq ORF3zr ORF3zs ORF3zt ORF3zu ORF3zv ORF3zw ORF3zx ORF3zy ORF3zz	VSP0017-1m	VSP0017-2m	VSP0017-3	VSP0017-4	VSP0017-5	VSP0017-6
	40	610	26	367	0	475
	62	193	0	0	0	476
	0	1	51	75	0	0
	0	0	0	0	35	0
	0	0	0	0	512	689
	52	236	0	0	0	0
	0	0	0	379	0	622
	0	0	0	334	0	688
	29	84	0	0	0	299
	0	0	0	0	580	0
	22	776	0	0	0	714
	83	324	6	267	2	0
	112	431	0	295	0	0
	72	167	0	308	485	457
	16	708	0	611	0	0
	6	352	0	329	0	0

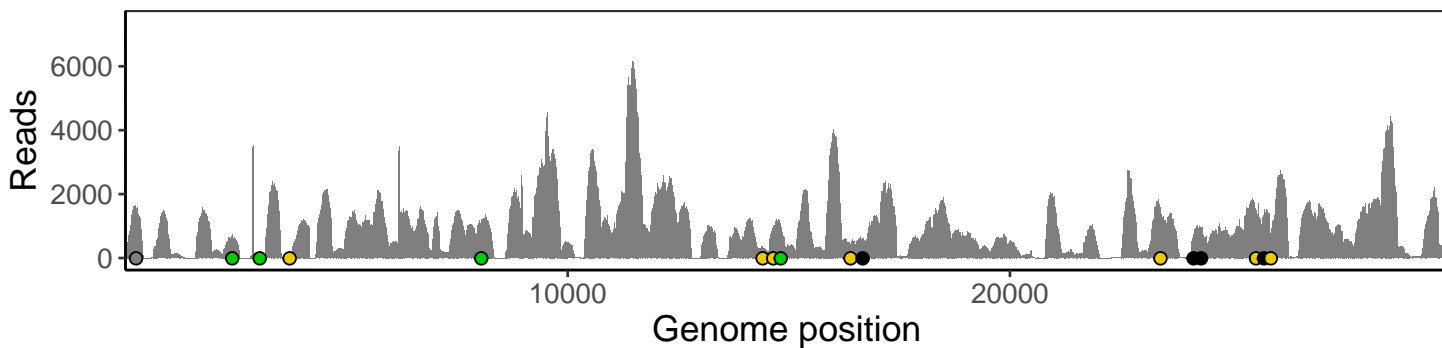
Base change

- Expected
- A
- T
- C
- G
- N
- Ins/Del
- No data

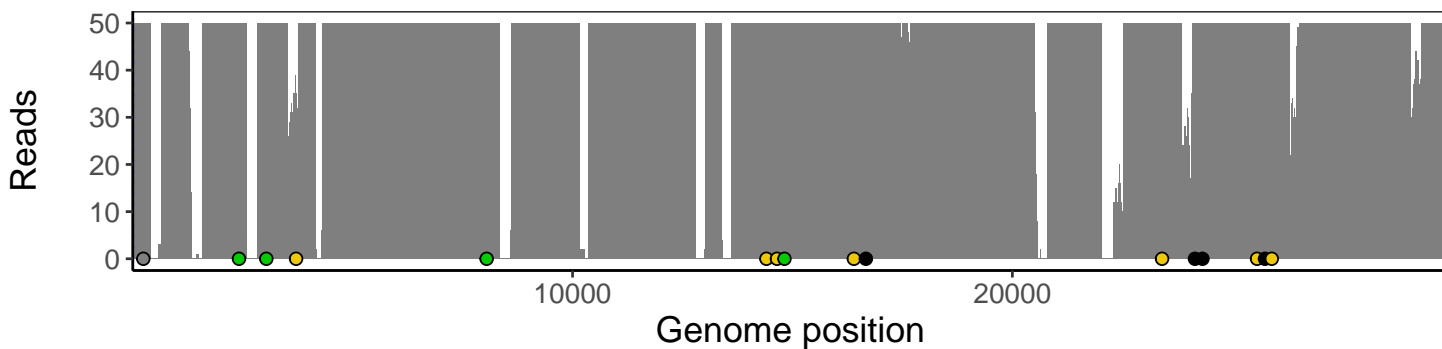
Analyses of individual experiments and composite results

VSP0017 | 2020-04-17 | ETA | 219-tri | composite result

The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according to variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



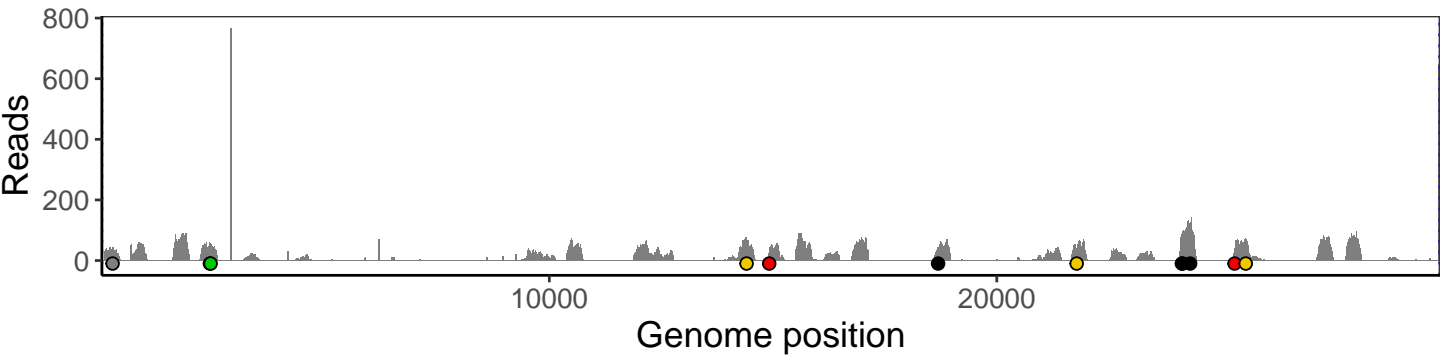
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



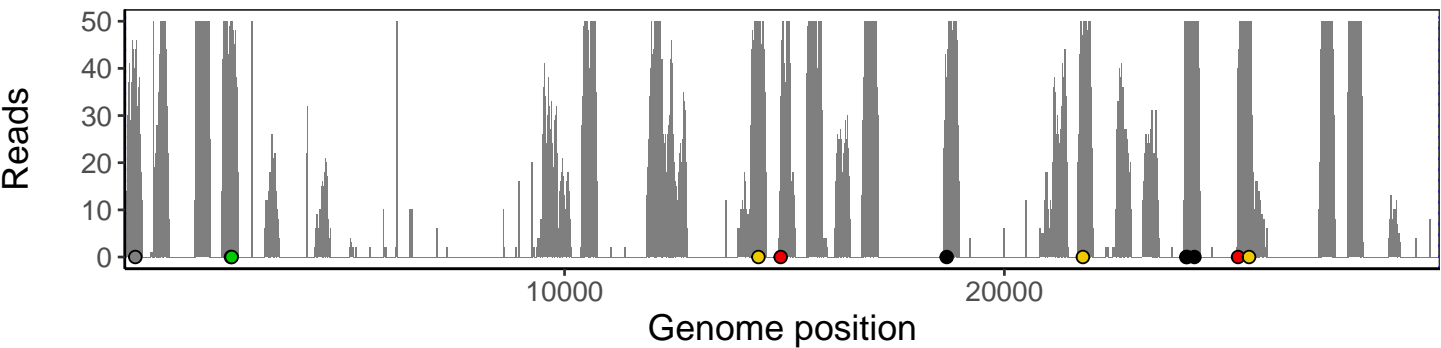
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



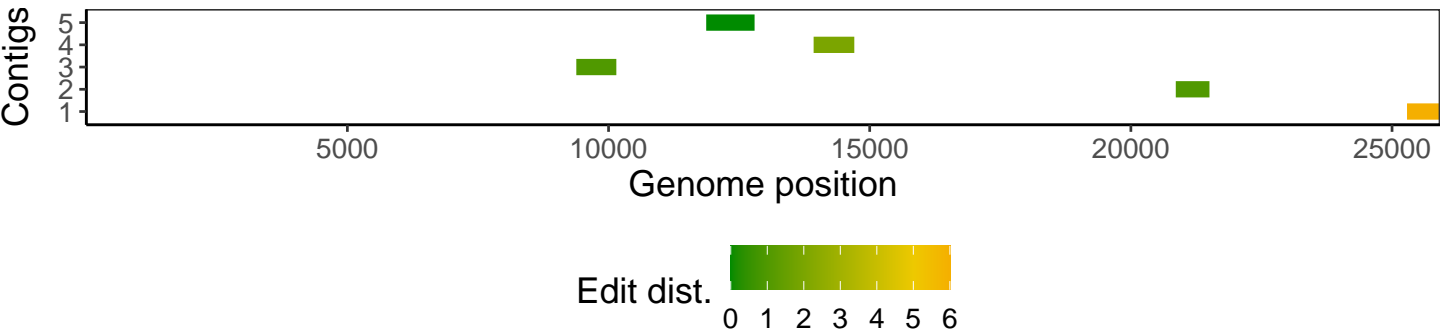
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



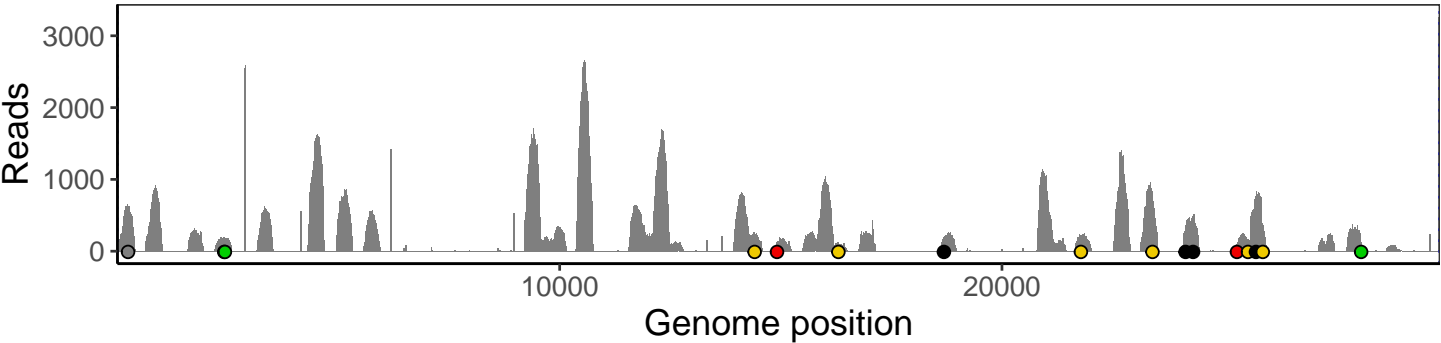
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



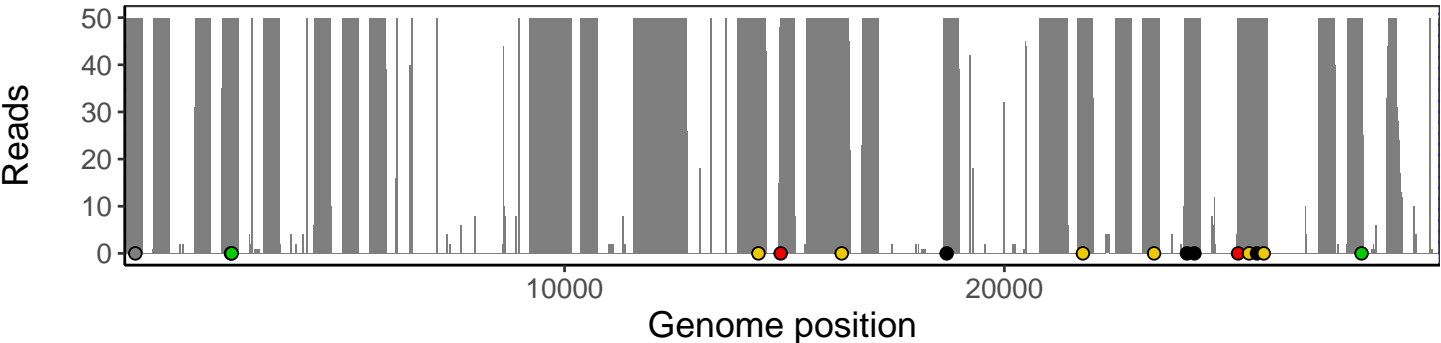
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



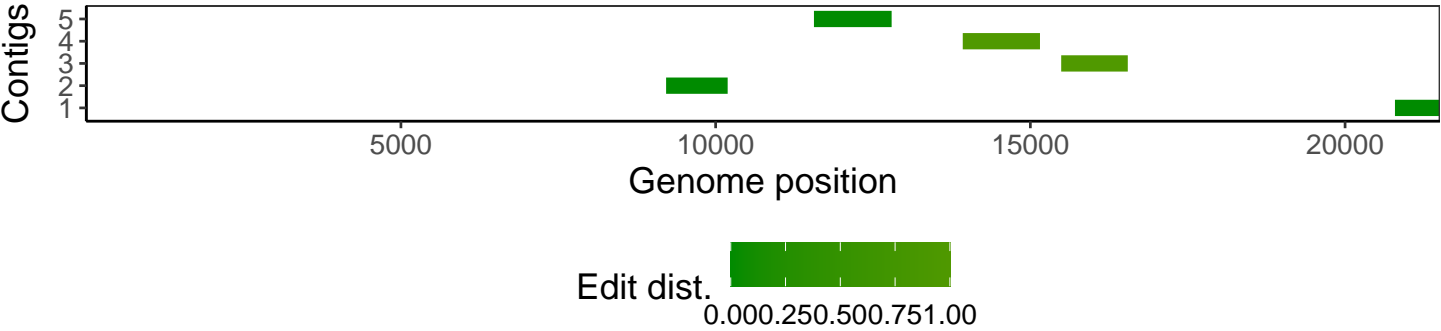
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



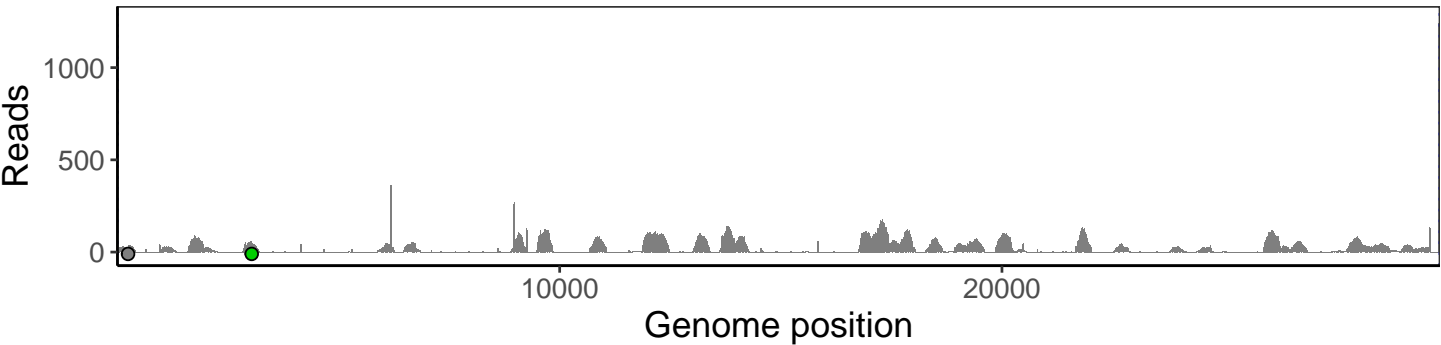
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



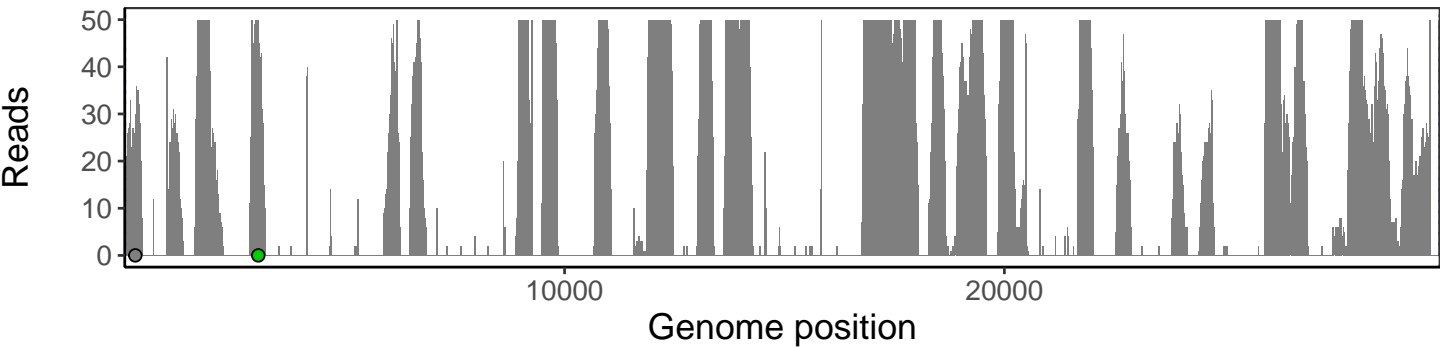
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



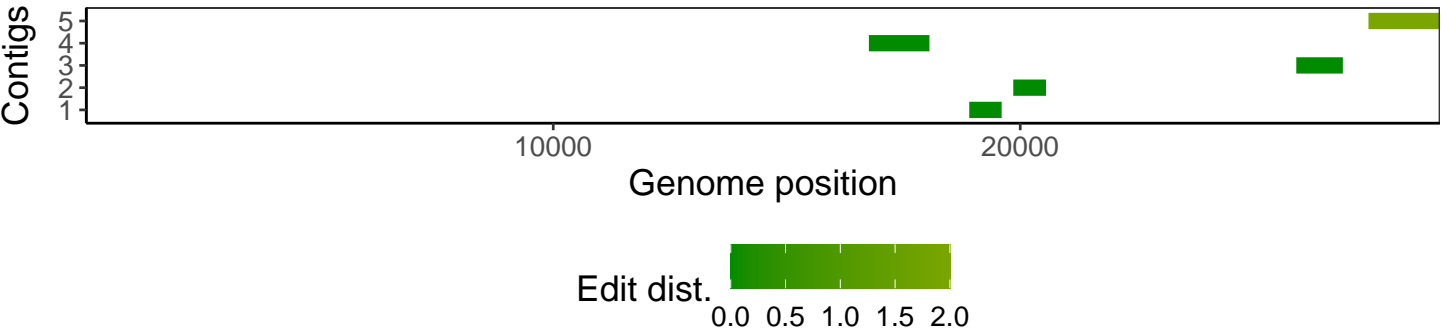
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



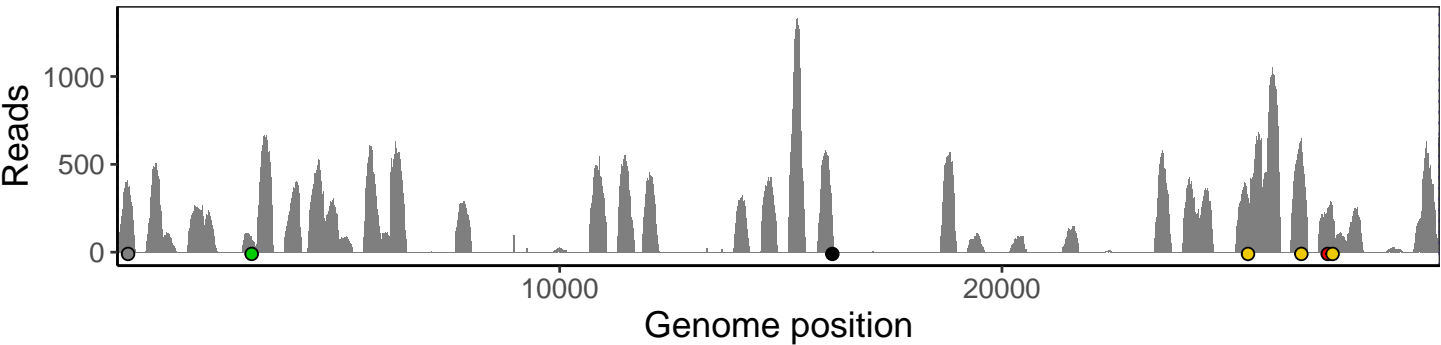
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



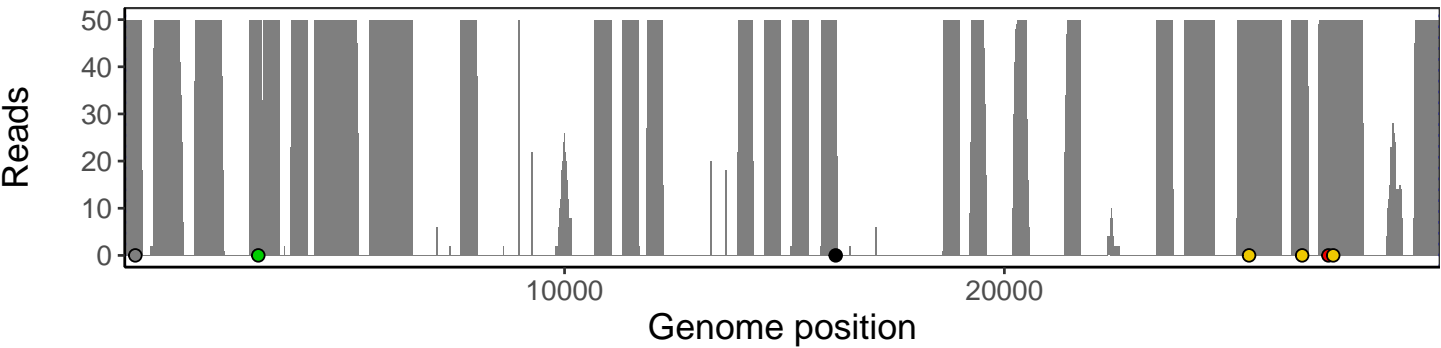
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



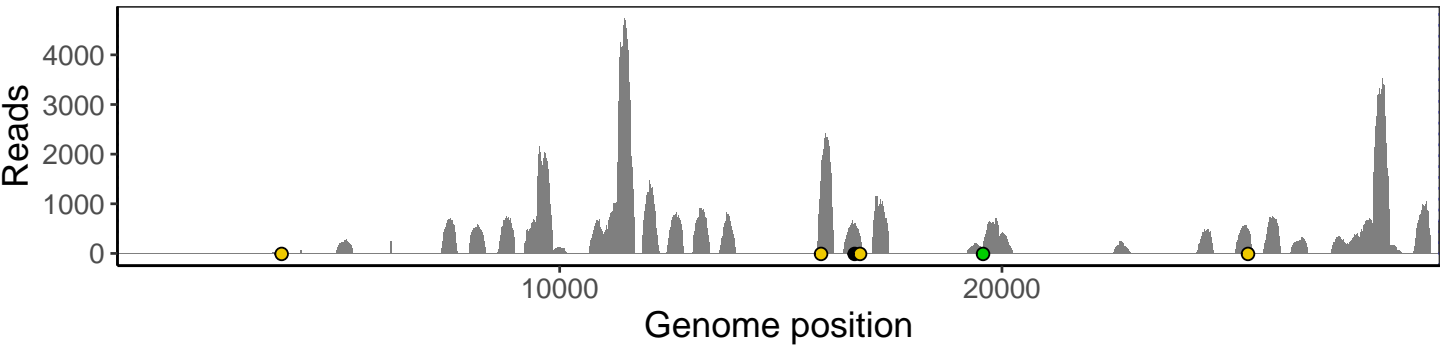
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



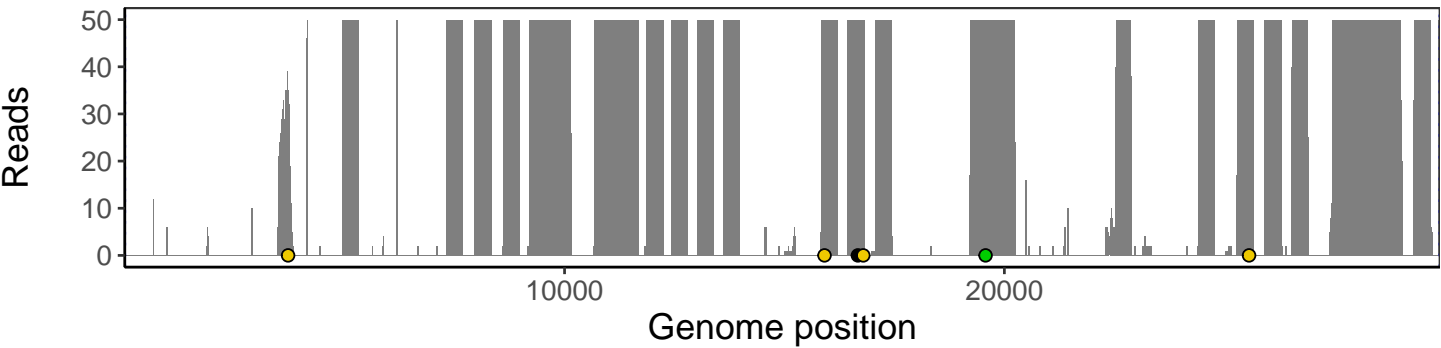
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



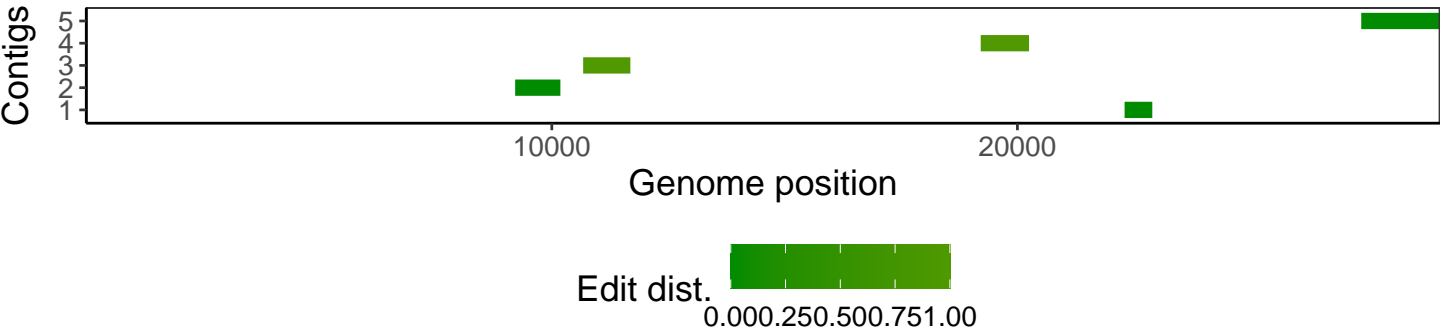
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



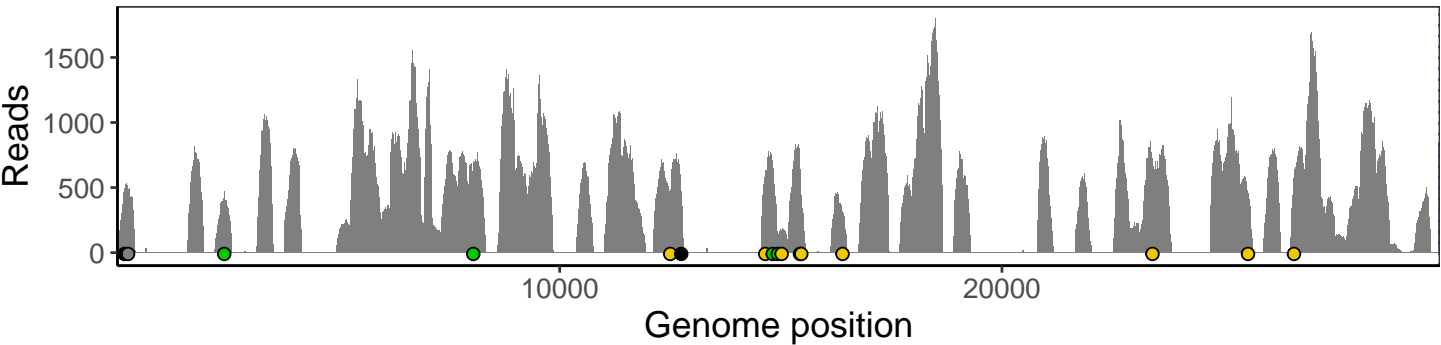
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



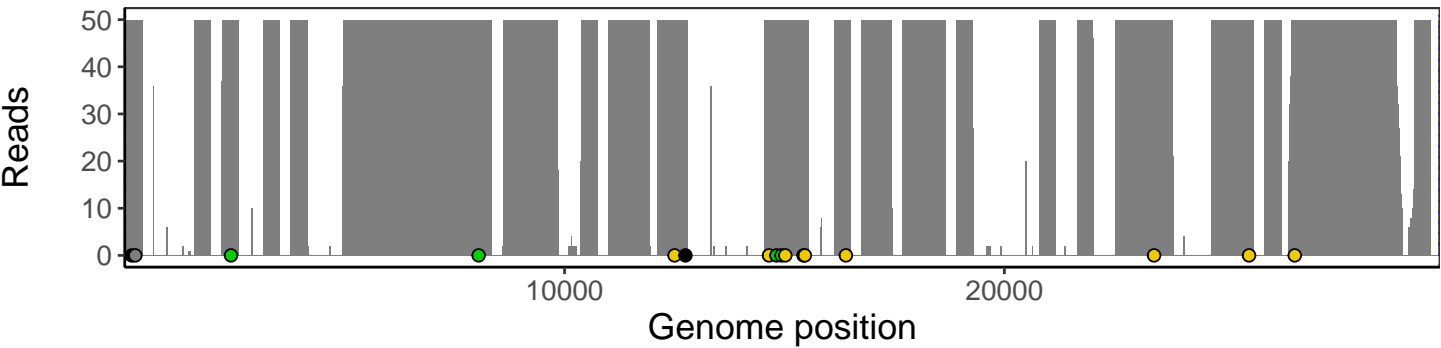
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



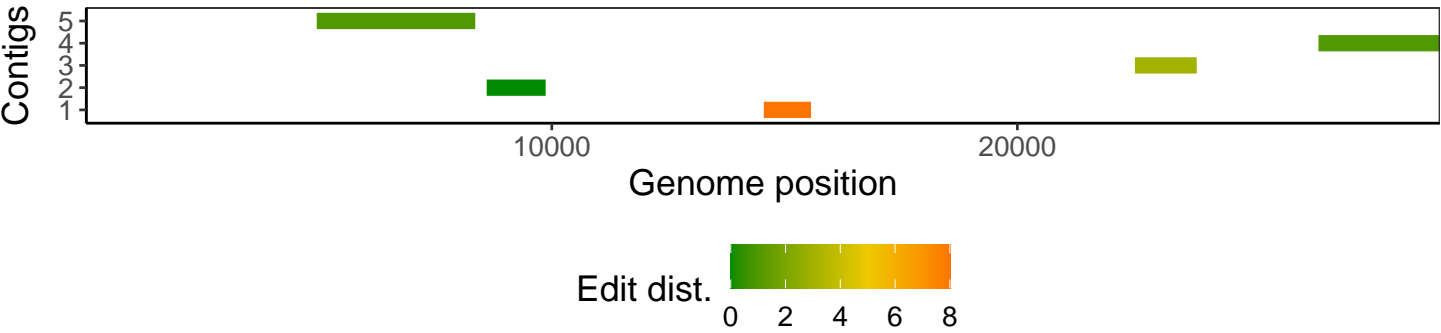
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



Software environment

Software/R package	Version
R	3.4.0
bwa	0.7.17-r1198-dirty
samtools	1.10 Using htlib 1.10
bcftools	1.10.2-34-g1a12af0-dirty Using htlib 1.10.2-57-gf58a6f3
pangolin	2.3.8
genbankr	1.4.0
optparse	1.6.0
forcats	0.3.0
stringr	1.4.0
dplyr	0.8.1
purrr	0.2.5
readr	1.1.1
tidyr	0.8.1
tibble	2.1.2
ggplot2	3.0.0
tidyverse	1.2.1
ShortRead	1.34.2
GenomicAlignments	1.12.2
SummarizedExperiment	1.6.5
DelayedArray	0.2.7
matrixStats	0.54.0
Biobase	2.36.2
Rsamtools	1.28.0
GenomicRanges	1.28.6
GenomeInfoDb	1.12.3
Biostrings	2.44.2
XVector	0.16.0
IRanges	2.10.5
S4Vectors	0.14.7
BiocParallel	1.10.1
BiocGenerics	0.22.1