Characterisation of SARS-CoV-2 genomic variations in response to molnupiravir treatment in the AGILE Phase IIa clinical trial.

Thomas Fletcher (tom.fletcher@lstmed.ac.uk)
Liverpool School of Tropical Medicine

I'ah Donovan-Banfield
Department of Infection Biology and Microbiomes, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool https://orcid.org/0000-0002-5124-2427

Rebekah Penrice-Randal
Department of Infection Biology and Microbiomes, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool

Hannah Goldswain
Department of Infection Biology and Microbiomes, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool https://orcid.org/0000-0003-4194-8714

Aleksandra Rzeszutek
Department of Evolution, Ecology and Behaviour, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool

Jack Pilgrim
Department of Evolution, Ecology and Behaviour, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool

Katie Bullock
GCPLab Facility, Institute of Systems, Molecular and Integrative Biology, University of Liverpool

Geoffrey Saunders
Southampton Clinical Trials Unit, University of Southampton

Josh Northey
Southampton Clinical Trials Unit, University of Southampton

Xiaofeng Dong
Department of Infection Biology and Microbiomes, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool

Yan Ryan
Department of Infection Biology and Microbiomes, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool

Helen Reynolds

6Department of Molecular and Clinical Pharmacology, Institute of Translational Medicine, University of Liverpool

Michelle Tetlow

6Department of Molecular and Clinical Pharmacology, Institute of Translational Medicine, University of Liverpool

Lauren Walker

NIHR Royal Liverpool and Broadgreen Clinical Research Facility,

Richard FitzGerald

NIHR Royal Liverpool and Broadgreen Clinical Research Facility

Colin Hale

NIHR Royal Liverpool and Broadgreen Clinical Research Facility

Rebecca Lyon

NIHR Royal Liverpool and Broadgreen Clinical Research Facility

Christie Woods

NIHR Royal Liverpool and Broadgreen Clinical Research Facility

Shazaad Ahmad

NIHR Manchester Clinical Research Facility, Manchester University NHS Foundation Trust

Dennis Hadjiyiannakis

NIHR Lancashire Clinical Research Facility, Lancashire Teaching Hospitals NHS Foundation Trust

Jimstan Periselneris

NIHR Kings Clinical Research Facility, King’s College Hospital NHS Foundation Trust

Emma Knox

Southampton Clinical Trials Unit, University of Southampton

Calley Middleton

Southampton Clinical Trials Unit, University of Southampton

Lara Lavelle-Langham

GCPLab Facility, Institute of Systems, Molecular and Integrative Biology, University of Liverpool

Victoria Shaw

GCPLab Facility, Institute of Systems, Molecular and Integrative Biology, University of Liverpool

William Greenhalf

Royal Liverpool University Hospital

Thomas Edwards

Centre for Drugs and Diagnostics, Liverpool School of Tropical Medicine

David Lalloo

Liverpool School of Tropical Medicine

Christopher Edwards

NIHR Southampton Clinical Research Facility

Alistair Darby

Department of Infection Biology and Microbiomes, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool

Miles Carroll
Oxford University

Gareth Griffiths
Southampton Clinical Trials Unit, University of Southampton

Saye Khoo
Department of Molecular and Clinical Pharmacology, Institute of Translational Medicine, University of Liverpool

Julian Hiscox
University of Liverpool

Brief Communication

Keywords:

Posted Date: July 18th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1835695/v1

Additional Declarations: There is **NO** Competing Interest.

Version of Record: A version of this preprint was published at Nature Communications on November 26th, 2022. See the published version at https://doi.org/10.1038/s41467-022-34839-9.

Characterisation of SARS-CoV-2 genomic variations in response to molnupiravir treatment in the AGILE Phase IIa clinical trial.

I'ah Donovan-Banfield[1,2], Rebekah Penrice-Randal[1], Hannah Goldswain[1], Aleksandra M. Rzeszutek[3], Jack Pilgrim[3], Katie Bullock[4], Geoffrey Saunders[5], Josh Northey[5], Xiaofeng Dong[1], Yan Ryan[1], Helen Reynolds[6], Michelle Tetlow[6], Lauren E. Walker[6,7], Richard FitzGerald[6,7], Colin Hale[7], Rebecca Lyon[7], Christie Woods[7], Shazaad Ahmed[8], Dennis Hadjiyiannakis[9], Jimstan Periselneris[10], Emma Knox[5], Calley Middleton[5], Lara Lavelle-Langham[4], Victoria Shaw[4,11], William Greenhalf[4], Thomas Edwards[12], David G. Lalloo[13], Christopher J. Edwards[14,15], Alistair C. Darby[1,16], Miles W. Carroll[2,17], Gareth Griffiths[5], Saye Khoo[6], Julian A. Hiscox*[1,2,18], Thomas Fletcher*[2,12].

[1]Department of Infection Biology and Microbiomes, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, UK.

[2]NIHR Health Protection Research Unit in Emerging and Zoonotic Infections, Liverpool, UK.

[3]Department of Evolution, Ecology and Behaviour, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, UK.

[4]GCPLab Facility, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, UK.

[5]Southampton Clinical Trials Unit, University of Southampton, UK

[6]Department of Molecular and Clinical Pharmacology, Institute of Translational Medicine, University of Liverpool, UK

[7]NIHR Royal Liverpool and Broadgreen Clinical Research Facility, Liverpool, UK

[8]NIHR Manchester Clinical Research Facility, Manchester University NHS Foundation Trust, Manchester, UK

[9]NIHR Lancashire Clinical Research Facility, Lancashire Teaching Hospitals NHS Foundation Trust, Preston, UK

[10]NIHR Kings Clinical Research Facility, King's College Hospital NHS Foundation Trust, London, UK

[11]The Clinical Directorate, University of Liverpool, UK

[12]Centre for Drugs and Diagnostics, Liverpool School of Tropical Medicine, Liverpool, UK

[13]Liverpool School of Tropical Medicine, Liverpool, UK

[14]Human Development and Health School, University of Southampton, Southampton, UK

[15]NIHR Southampton Clinical Research Facility, University Hospital Southampton NHS Foundation Trust, Southampton, UK

[16] NIHR Health Protection Research Unit in Gastrointestinal Infections, Liverpool, UK.

[17]Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, UK

[18]A*STAR Infectious Diseases Laboratories (A*STAR ID Labs), Agency for Science, Technology and Research (A*STAR), Singapore.

*Corresponding authors.

Word count

Abstract 124 words

Main 1923 words

Abstract

Molnupiravir is an antiviral approved for treating COVID-19, which is thought to drive lethal error catastrophe. How this drug-induced mechanism of action impacts the emergence of resistance mutations is unclear. AGILE Candidate Specific Trial (CST)-2 is a phase IIa trial randomising 180 adult outpatients with SARS-COV-2 infection within five days of symptom onset to molnupiravir or placebo, with rich serial sampling of nasopharyngeal swabs over 29 days. Viral sequences, that passed genome quality control criteria, from subjects who received molnupiravir (n=59) or a placebo (n=65) were analysed by high-throughput amplicon sequencing. We found evidence that molnupiravir significantly increased the transition/transversion frequency in SARS-CoV-2 in patients, a hallmark of molnupiravir treatment. Over the course of treatment, no consistent, accumulated mutations were identified in either arm.

Main

The roll-out of oral, directly acting antivirals (DAAs) to treat SARS-CoV-2 needs to be accompanied by careful monitoring for development of treatment-emergent resistance mutations in current and future circulating variants, as this may limit the public health impact of therapy. DAAs are small molecules which target key stages of the SARS-CoV-2 life cycle. As with HIV, their genetic barrier to resistance will likely differ between drugs, according to their mechanism of action. The activity of DAAs is expected to be less impacted by different SARS-CoV-2 variants compared with monoclonal antibodies, however clinical data are lacking.

Three small molecule DAAs have received early use authorisation for treating COVID-19: remdesivir, molnupiravir (both nucleoside analogues which target viral nucleic acid synthesis) and nirmatrelvir (which targets the main viral protease). Unlike remdesivir, molnupiravir is orally administered and thus more readily deployed for treatment in the community. Both remdesivir and molnupiravir are prodrugs, with their active triphosphate metabolites incorporated by the RNA-dependent RNA-polymerase (RdRp) (NSP12) which is the catalytic core of the replication complex for viral RNA synthesis[1,2]. This encompasses two major processes: 1) replication of the genome involving synthesis of a negative strand template for direct copying of new genomes and 2) discontinuous transcription of sub-genomic messenger RNAs (sgmRNAs). Directly inhibiting the function of the proteins involved in viral RNA synthesis or interfering with RNA synthesis itself will reduce viral replication and ultimately viral load.

Molnupiravir has a different mechanism of action to remdesivir[2,3]. In human airway cultures and mouse models of disease, molnupiravir inhibits SARS-CoV-2 RNA synthesis by inducing G $\rightarrow$ A and C $\rightarrow$ U transition mutations, causing lethal mutagenesis[4]. The MOVe-OUT phase III

double-blinded clinical trial reported that early treatment with molnupiravir reduced the risk of hospitalisation or death in at-risk, unvaccinated adults with COVID-19[5].

AGILE is the UK early-phase trial platform for COVID-19 antivirals[6] conducted by the Southampton Clinical Trials Unit, University of Liverpool, Liverpool School of Tropical Medicine, the NIHR Royal Liverpool and Broadgreen Clinical Research Facility (CRF) and the CRF network. Following the establishment of a recommended phase II dose[7] the AGILE CST-2 phase II randomised 180 adult outpatients with confirmed SARS-CoV-2 infection within five days of symptom onset to receive molnupiravir (800 mg twice daily for 5 days) or placebo (clinical trial number: NCT04746183). Here, we sequenced serial nasopharyngeal samples from those subjects to characterise drug-induced viral adaptation (Figure 1a(i)). An amplicon-based deep sequencing approach was used to determine the SARS-CoV-2 genome to high sequence read depth such that both lineage assignment and minor genomic variant information could be generated to enable identification of the mechanism of action (Figure 1a(ii), Extended Data Table 1). Patients were included in the minor variant analysis if all three of their samples met the following criteria: 1) the consensus genome had a minimum 90% consensus called and 2) 90% of genome positions had a minimum coverage of 200X. Using these criteria, longitudinal samples from 65 patients receiving placebo and 59 patients treated with molnupiravir were identified for SARS-CoV-2 genomic analysis.

Molnupiravir was predicted to increase the number of mutations in the genome of SARS-CoV-2 (Figure 1b) and that this would manifest as an increase in the transition/transversion (Ts/Tv) ratio[8]. The sequencing data indicated that transition mutations were significantly increased in viral RNA from molnupiravir treated patients at Day 3 or Day 5 compared to patients given a placebo (Figure 1c). The frequency of C $\rightarrow$ U mutations were higher than those for G $\rightarrow$ A

(Figure 1d). U → C mutations were also significantly increased. All other base changes showed no increase over time in either group (Supplemental Figure 1).

The implications of greater viral diversity in response to molnupiravir treatment are currently unknown, but it could potentially influence the genetic barrier to resistance. To address this, SARS-CoV-2 sequence was translated *in silico* at both the dominant and minor variant genome level and treatment-emergent mutations were analysed to assess preferential enrichment of mutations (i.e., is there a greater chance of mutations arising during treatment and then persisting in these regions thereafter). Given the mechanism of action of molnupiravir, the two most obvious genes under selection pressure would be *nsp12* (the RNA dependent RNA polymerase; RdRp) and *nsp14* (the exonuclease). Incorporation of molnupiravir in the nascent template would likely affect either NSP12 or NSP14, potentially triggering stalling or back-tracking for excision by the exonuclease. Two previous studies on the incorporation of molnupiravir into the nascent template found that molnupiravir did not cause polymerase stalling, but one of the studies demonstrated that molnupiravir was capable of inducing chain termination[1,3]. If chain termination occurred, this may have placed selection pressure on both the RdRp and the exonuclease to be able to counter the effects of molnupiravir. In our study, the data indicated that there was no change in the predicted amino acid sequence of NSP12 and NSP14 at the dominant genome level over the first five days of molnupiravir treatment (Figure 2b and c).

Reflecting the change in the Ts/Tv ratio, the diversity of the predicted amino acid sequence increased over the course of infection in both treatment groups. The spread of diversity was reflected across the genome, with a slight bias towards the 3' end. More diversity observed in the Day 5 samples from the molnupiravir-treated group compared to the placebo group

(Figure 2 - with data from patients infected with Delta variant of concern (VoC) viruses as an example). A similar pattern was found in patients infected with other VoCs (Extended Data Figure 2).

Curiously, two positions in NSP14 had a slightly increased diversity (199 and 202) that were present in samples from both treated and placebo groups but may represent a persistent sub-population (Figure 2c). To understand any risks of combining molnupiravir with monoclonal antibody treatment, we also evaluated amino acid substitutions in the spike protein. Two of the positions (19 and 95), which are known lineage-defining mutation sites in all Delta sub-lineages, were variable in patients from both treated and placebo control groups. However, reflecting the mechanism of action of molnupiravir, this appeared to be more diverse in the treated group (Figure 2d and Extended Data Figures 3-5 for other lineages).

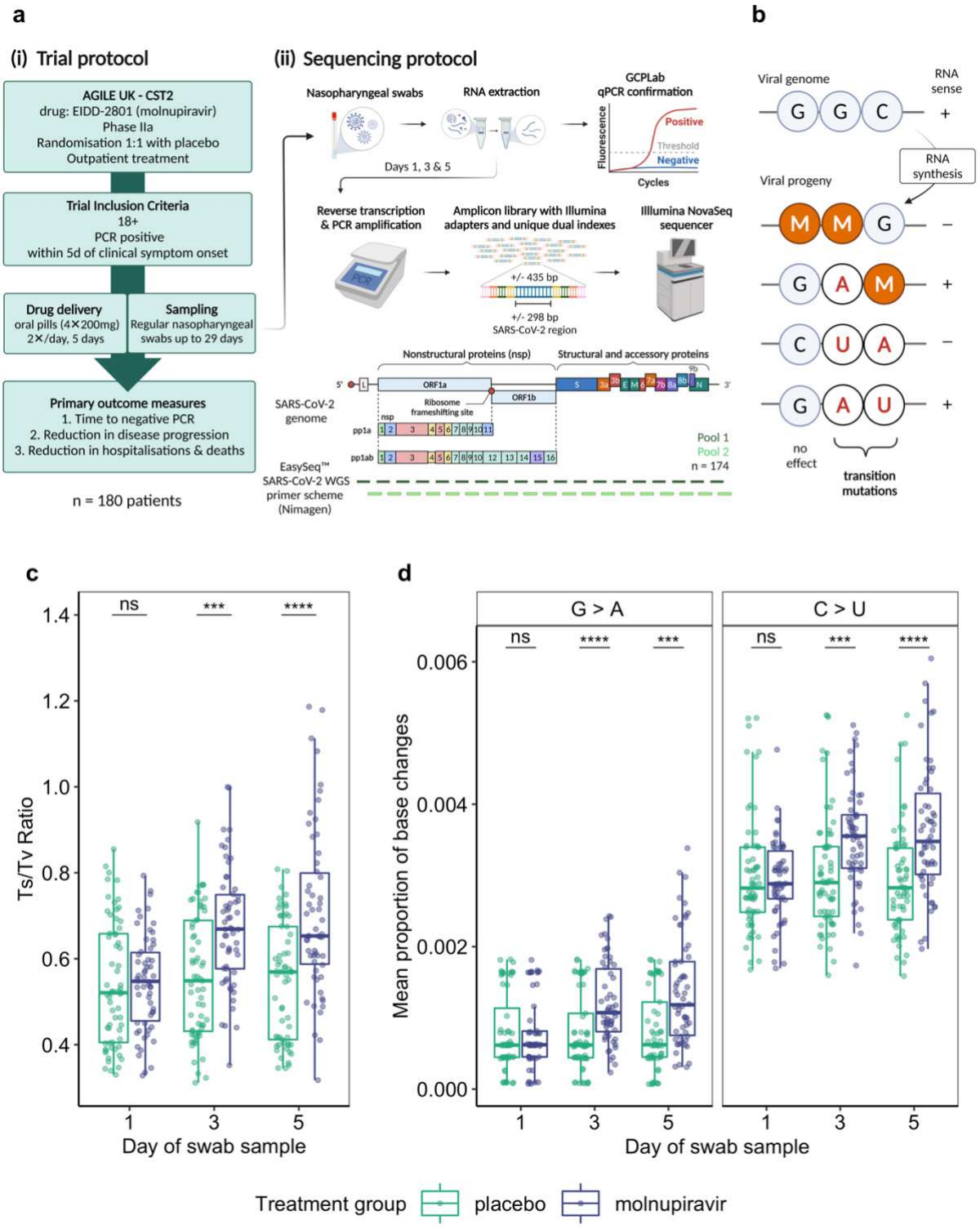
To our knowledge, this is the first confirmation of the mechanism of molnupiravir on viral replication in humans infected with SARS-CoV-2, following an approved dosing regimen. In the molnupiravir treated group, the Ts/Tv mutation ratio was higher than in the placebo group. This corresponded with higher C $\rightarrow$ U and G $\rightarrow$ A mutations than other combinations. The increase in this ratio corresponded to the length of treatment, with the greatest diversity seen on Day 5. There were no amino acid substitutions in SARS-CoV-2 that were enriched consistently at specific sites in the molnupiravir-treated group at any of the sampled times, including in the genes encoding NSP12 and NSP14. This suggests, that over the course of treatment assessed in this study, no drug-induced adaptations emerged due to molnupiravir treatment.

During SARS-CoV-2 infection, viral adaptation and neutral mutations occur. Treatment with molnupiravir aims to surmount the threshold of tolerated genetic errors, such that viral

replication is diminished, resulting in a concomitant reduction in viral load. This study revealed the intricacies of this mechanism of action in humans. This study also highlighted the utility of minor genomic variant analysis in examining intra-host virus populations which strengthens the prediction, and surveillance, of treatment-emergent adaptations. A deep-sequencing and bioinformatic pipeline for handling and visualising minor variant data was established and can be used with other antiviral treatments for COVID-19 or similar viral infections. In future, such approaches can be used by regulatory bodies and public health officials to inform approval decisions and surveillance of resistance in the wake of large-scale administration of newly approved drugs. The data described complements the clinical findings and has provided comprehensive information regarding drug effects on viral genomes.

References

1. Kabinger, F., *et al. Nat Struct Mol Biol* **28**, 740-746 (2021).
2. Kokic, G., *et al. Nat Commun* **12**, 279 (2021).
3. Gordon, C.J., *et al. J Biol Chem* **297**(2021).
4. Sheahan, T.P., *et al. Sci Transl Med* **12**(2020).
5. Jayk Bernal, A., *et al. N Engl J Med* **386**, 509-520 (2022).
6. Griffiths, G., *et al. Trials* **21**, 544 (2020).
7. Khoo, S.H., *et al. J Antimicrob Chemother* **76**, 3286-3295 (2021).
8. Agostini Maria, L., *et al. J Virol* **93**, e01348-01319 (2019).
9. Coolen, J.P.M., *et al. J Clinl Virol* **144**, 104993 (2021).
10. O'Toole, Á., *et al. Virus Evol* (2021).
11. Prince, T., *et al. mSphere* **0**, e00913-00921.



174

175 **Figure 1: Protocol overview and the detection of the molecular signatures of molnupiravir**
176 **mechanism of action. a,** (**i**) A simplified AGILE CST-2 Phase IIa trial protocol. Molnupiravir was
177 administered to outpatients as four oral pills (200mg each, 800mg total) every 12 hours for
178 five days. Patients were randomised placebo to drug 1:1, with nasopharyngeal swabs taken
179 for viral load monitoring. **(ii)** Sequencing protocol. RNA extracted from nasopharyngeal
180 swabs, taken at days 1, 3 and 5 post treatment initiation, was used for amplicon library
181 preparation using the EasySeq™ RC-PCR SARS-CoV-2 WGS kit (Nimagen, Netherlands).

Resulting sequence reads were mapped to the Wuhan-Hu-1 reference (NC_045512.2). **b,** Molnupiravir mechanism of action *via* the RNA template leads to the accumulation of transition mutations in viral progeny. **c,** Average Ts/Tv ratio values per RNA sample from all patients (placebo n = 65, green; molnupiravir n = 59, blue). SARS-CoV-2 RNA from molnupiravir (blue) patients shows a statistically significant accumulation of transition mutations over time compared to placebo (green). **d**, The same information as in **c** but showing the frequency of individual transition mutations G $\rightarrow$ A and C $\rightarrow$ U. Wilcoxon rank sum test was performed in **c** and **d**; ****$P \leq 0.0001$, ***$P \leq 0.001$, ns = $P > 0.05$. RC-PCR, reverse complement-polymerase chain reaction; WGS, whole genome sequencing; GCPLab, good clinical practice laboratory (University of Liverpool).

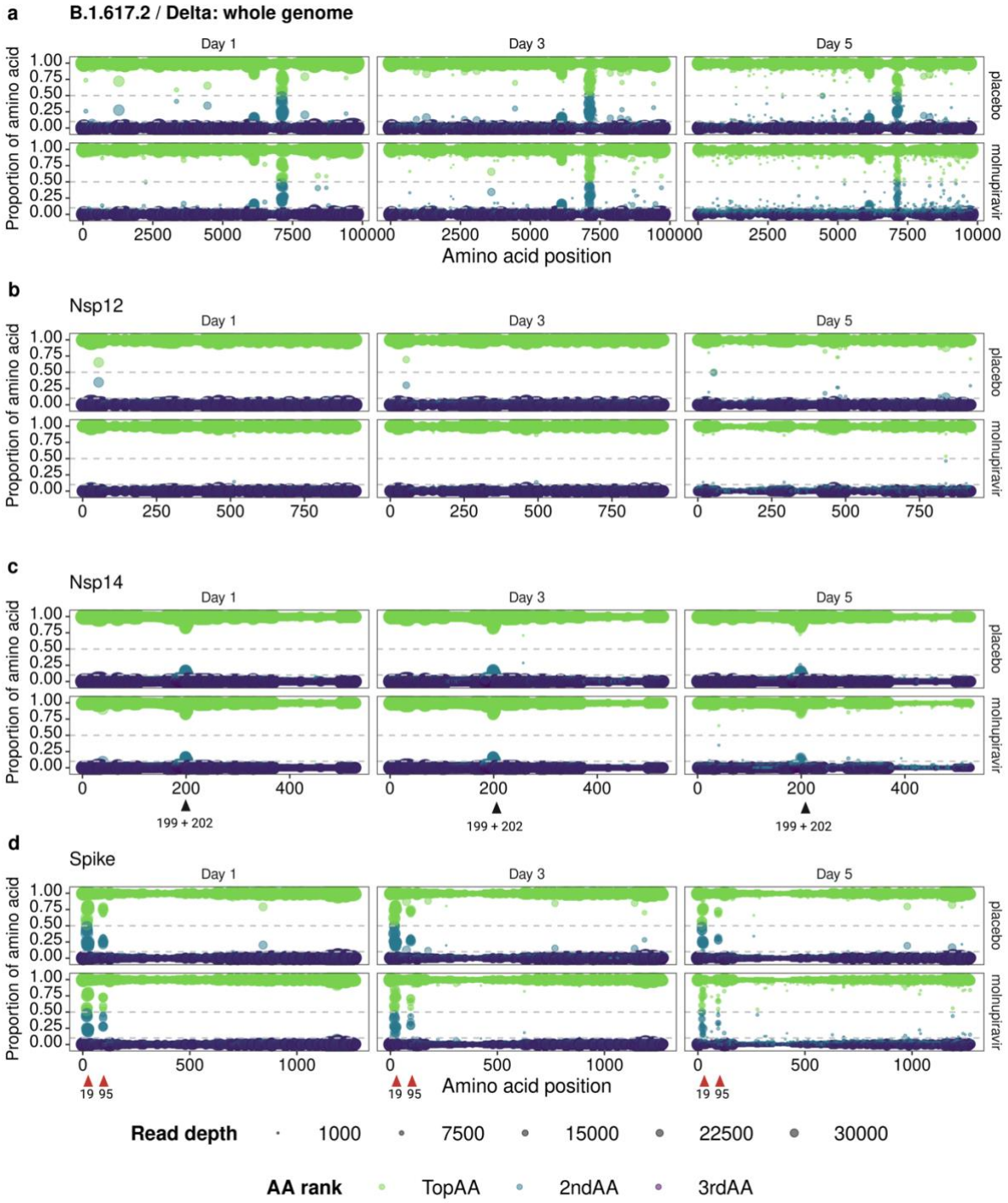


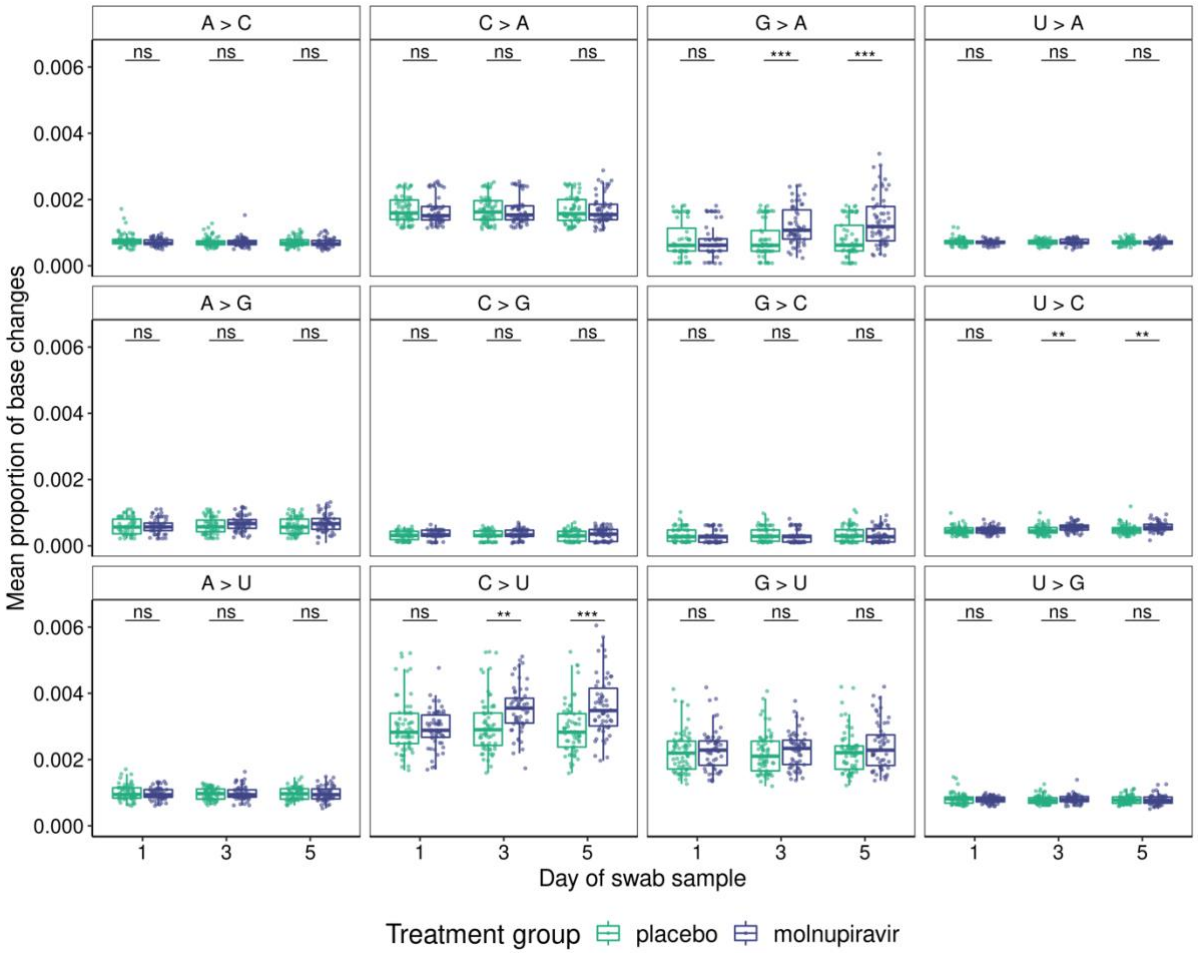
Figure 2: Predicted amino acid variations derived from SARS-CoV-2 RNA in the whole genome, NSP12, NSP14 and Spike sequences. a, Predicted amino acid variation derived from RNA sequence information across the whole genome in all Delta infected patients (n=52). Each sample is assigned a predicted "Top", "2nd" and "3rd" amino acid based on proportion of reads at every genome position. Minimum read depth = 200. Minor genomic variants (>0.1 and <0.5; grey dashed lines) increase in frequency over time, with viral RNA from molnupiravir treated patients showing more diversity. **b,** NSP12 showed very little minor genomic variation over the five days. **c,** NSP14 also showed minor genomic stability, but had sites of low-level minor variation at 199 and 202 (indicated with black arrows) that was present in all samples

tested and may represent a persistent sub-population. **d**, Spike had two sites with an amino acid mixed population at 19 and 95 (indicated with red arrows) in all Delta samples analysed. These are known VOC sites in all the Delta sub-lineages.

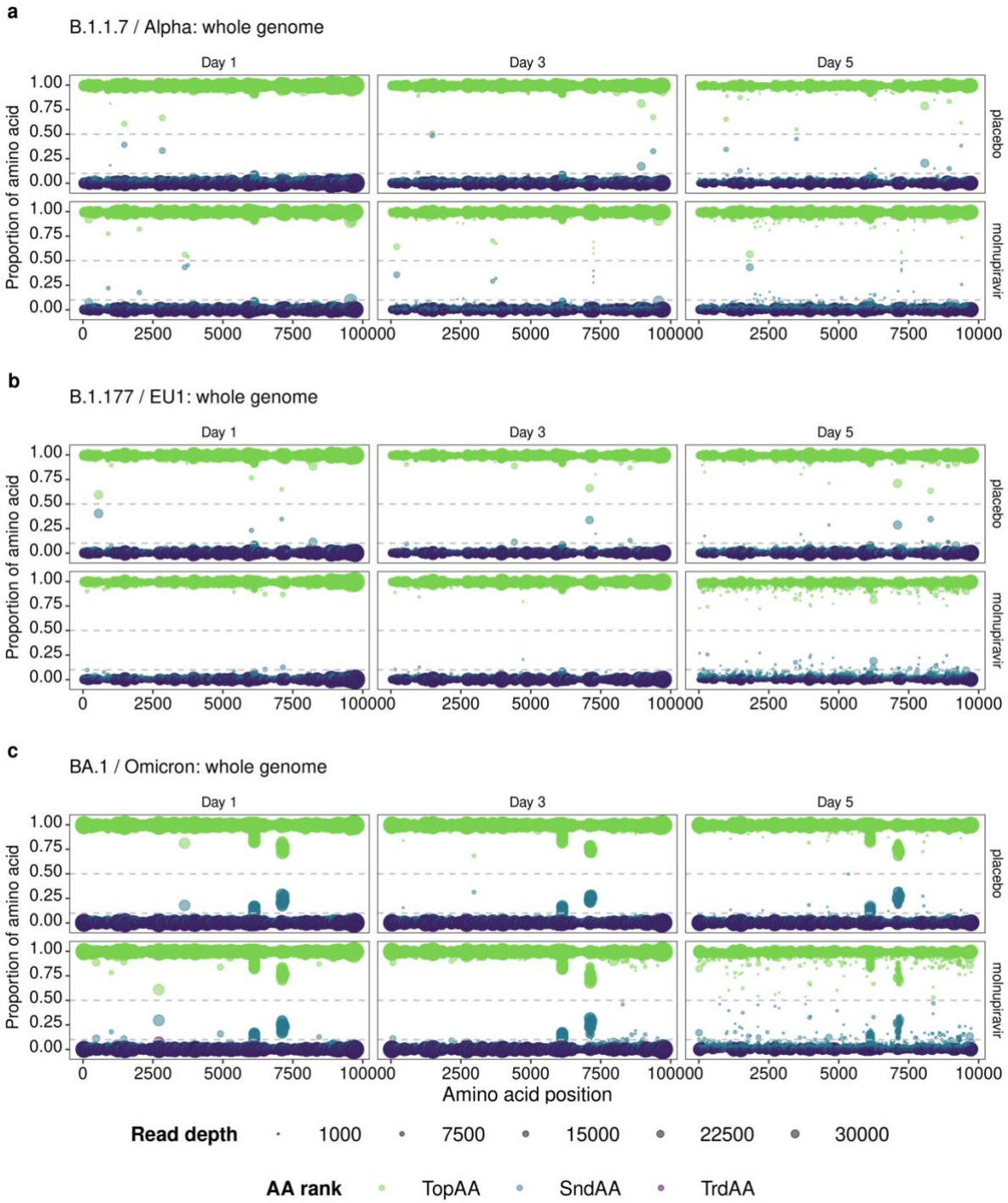
Extended Data

Lineage	***Placebo*** Total (passed)	***Molnupiravir*** Total (passed)
B.1.1.7 (Alpha)	20 (14)	17 (11)
B.1.1.1	1 (1)	0 (0)
B.1.177 (EU1)	13 (10)	15 (8)
Delta (all)	35 (24)	37 (28)
B.1.617.2	*2 (0)*	*2 (2)*
AY.120	*1 (1)*	*0 (0)*
AY.33	*0 (0)*	*1 (1)*
AY.4	*28 (21)*	*30 (22)*
AY.43	*0 (0)*	*1 (1)*
AY.4.2	*2 (1)*	*2 (2)*
AY.4.2.1	*1 (1)*	*0 (0)*
AY.98	*1 (0)*	*1 (1)*
Omicron (all)	19 (16)	20 (12)
BA.1	*12 (9)*	*15 (11)*
BA.2	*6 (6)*	*5 (1)*
XE	*1 (1)*	*0 (0)*
Failed to assign	2 (0)	1 (0)
Trial total	**90 (65)**	**90 (59)**

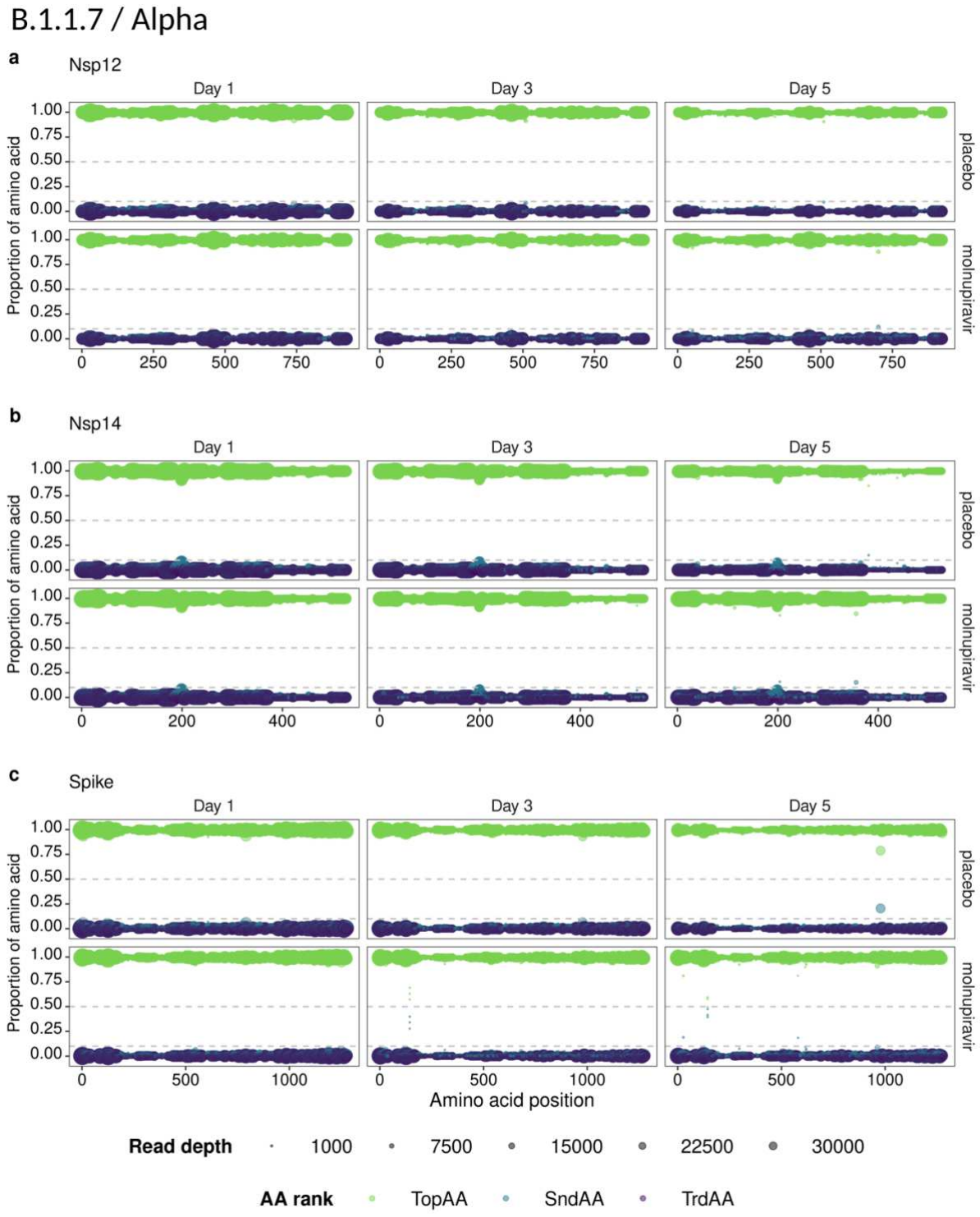
Table 1: Lineage assignment of SARS-CoV-2 from patients enrolled in the AGILE CST-2 phase IIa molnupiravir clinical trial. Viral RNA from nasopharyngeal swabs obtained from patients enrolled in the phase IIa clinical trial was sequenced as described in Methods. The consensus SARS-CoV-2 genome for each sample, assembled after mapping to the Wuhan-Hu-1 reference genome, was used to assign the lineage of SARS-CoV-2 that each patient was infected with upon entering the trial, using the software tool, Pangolin (version 4.0.6). Only patients that passed criteria of all samples (Days 1, 3 and 5) with a minimum 90% genome coverage were included in downstream analyses – numbers indicated in brackets for each (sub-)lineage. Lineages that only had one patient or an uneven balance of placebo:drug were excluded from the analysis.



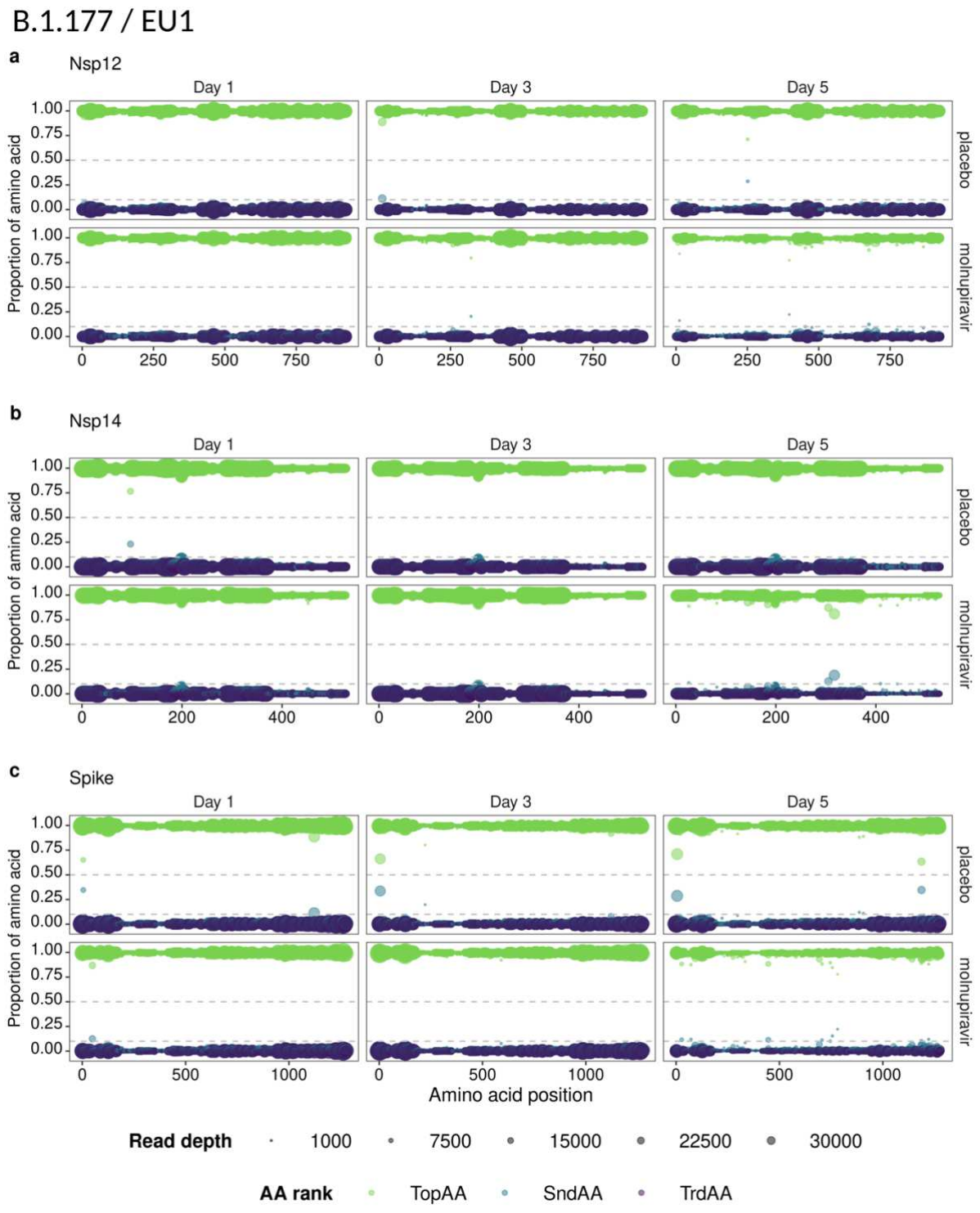
S1: All base changes over time. The mean frequency of all possible base change combinations was calculated per sample, with data grouped by treatment (placebo n = 65, green; molnupiravir n = 59, blue) and day of swab sample. Wilcoxon rank sum test was performed, to calculate the statistical significance of the mean difference in bases change frequency between treatment groups on each sample day. Of the twelve possible base changes, only 'G to A', 'C to U' and 'U to C' showed statistically different mean frequencies between groups at Days 3 and 5. ****$P \leq 0.0001$, ***$P \leq 0.001$, **$P \leq 0.01$, ns = $P > 0.05$.



S2: Predicted amino acid variations derived from SARS-CoV-2 RNA in the whole genome of alpha, B.1.177/EU1 and BA.1/Omicron lineages. Predicted amino acid variation derived from RNA sequence information across the whole genome in all **a,** alpha (placebo n=14, molnupiravir n=11); **b,** B.1.177/EU1 (placebo n=10, molnupiravir=8); and **c,** BA.1/Omicron (placebo n=9, molnupiravir=11) infected patients. Each sample is assigned a predicted "Top", "2nd" and "3rd" amino acid based on proportion of reads at every genome position. Minimum read depth = 200. Minor genomic variants (>0.1 and <0.5; grey dashed lines) increase in frequency over time, with viral RNA from molnupiravir treated patients showing more diversity.



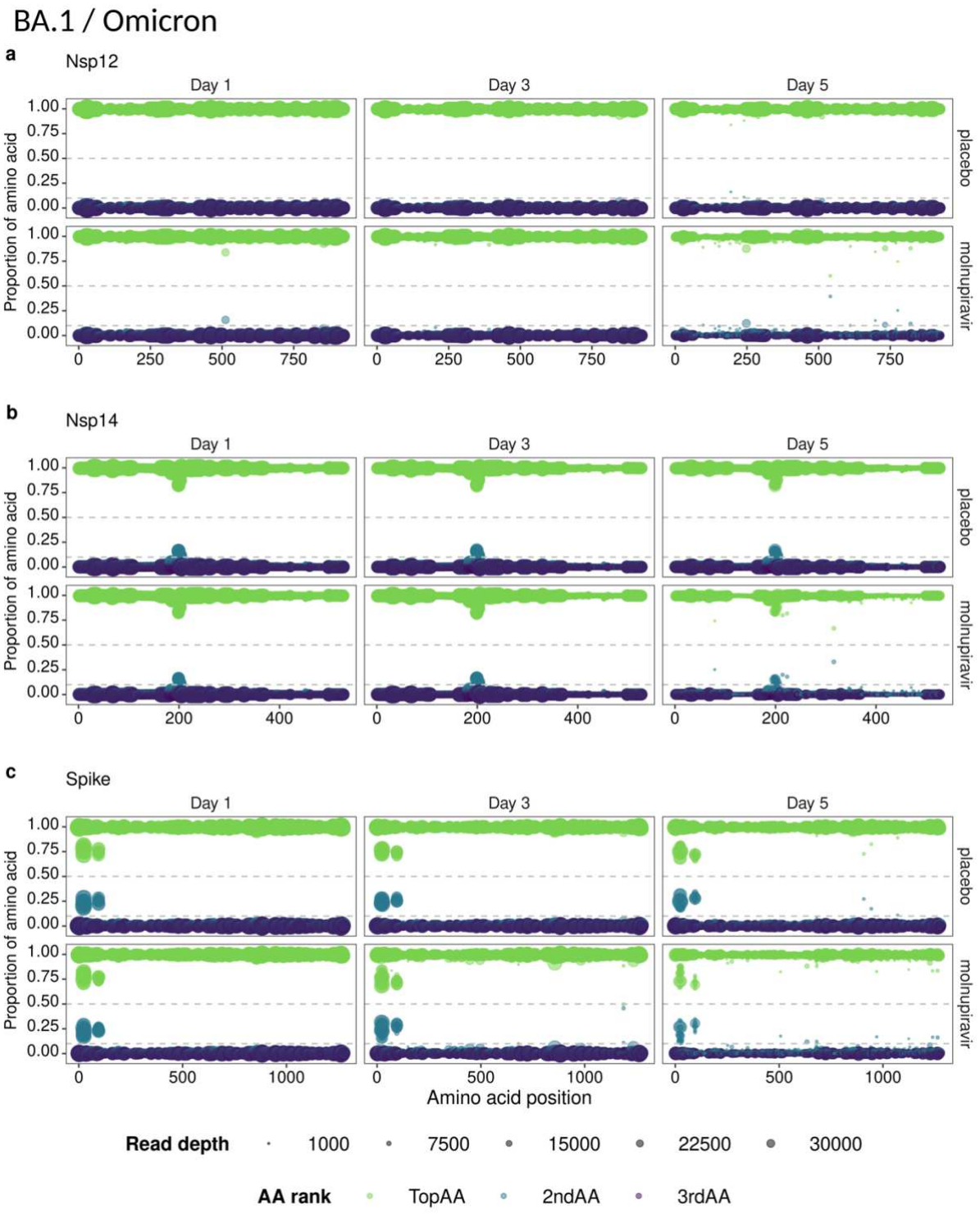
S3. Alpha - predicted amino acid variations in Nsp12, Nsp14 and Spike proteins.



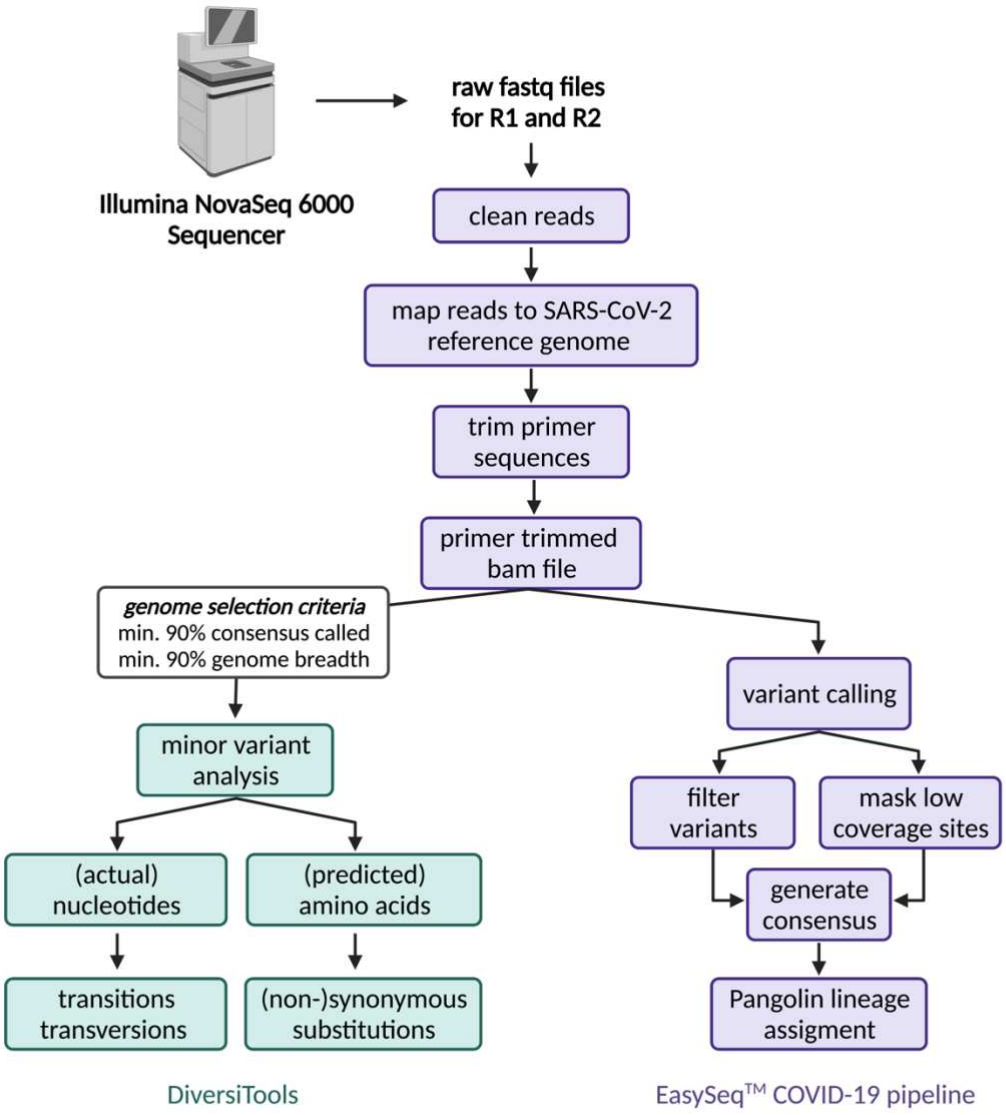
240

241 **S4. B.1.177/EU1 - predicted amino acid variations in Nsp12, Nsp14 and Spike proteins.**

242



S5. BA.1/Omicron - predicted amino acid variations in Nsp12, Nsp14 and Spike proteins.



S6: Computational workflow used to generate SARS-CoV-2 genomic data, assign PANGO lineage and analyse minor genomic variants.

Methods

Sample Collection

AGILE is a randomised multi-arm, multi-dose, phase I/IIa platform in the UK using a seamless Bayesian adaptive design to determine the safety, activity, and optimal dose of multiple SARS-CoV-2 candidate therapeutics[6]. This trial evaluated molnupiravir (EIDD-2801/MK-4482), for the treatment of COVID-19 in a seamless phase I/II trial (clinicaltrials.gov registration number NCT04746183). Eligible participants were men and women aged ≥18 years with PCR-confirmed SARS-CoV-2 infection who were within five days of symptom onset, free of uncontrolled chronic conditions, and ambulant in the community with mild or moderate disease. Nasopharyngeal swabs were obtained from patients on days 1, 3, 5, 8, 11, 15, 22 and 29.

RNA Extraction, Amplicon Library Preparation, and Illumina Sequencing

RNA was extracted from the nasopharyngeal swabs by the GCP Laboratory Facility at the University of Liverpool using a Maxwell® RSC instrument, an automated nucleic acid extraction instrument (Promega, USA). Aliquots of surplus RNA were provided for sequencing analysis. Briefly, library preparation consisted of converting RNA to cDNA using LunaScript™ (Thermofisher, Waltham, Massachusetts), then amplified by reverse complement (RC)-PCR amplification (EasySeq™ SARS-CoV-2 Whole Genome Sequencing kit, NimaGen, Nijmegen, The Netherlands)[9]. This kit barcodes and ligates Illumina adapters in a single PCR reaction, with two separate pools of primers (pools 1 and 2). After amplification, primer pools 1 and 2 for each amplified sample were mixed 1:1 before being cleaned with Beckman Coulter™ Agencourt AmpureXP beads (Fisher Scientific, Hampton, New Hampshire), quantified and the library quality assessed on an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, California). All purified samples were then pooled together and denatured. Finally, the denatured amplicon

library was loaded into the NovaSeq cartridge (2 × 150 bp run) before loading on the NovaSeq 6000 machine. The sequencing was conducted in two separate sequencing runs, one for the first 120 patients' swab samples, and a second for the final 60 patients' swab samples.

In silico processing

The raw sequencing data was processed using two different pipelines (summarised in Supplemental Figure 6). The first method, EasySeq_covid19 (version 0.9, code available at https://github.com/JordyCoolen/easyseq_covid19), performs quality control steps, maps to the reference genome (Wuhan-Hu-1; NC045512.2), variant calls and generates a consensus genome for each sample[9]. Pangolin (version 4.0.6) was used to assign virus lineage[10]. The second method, DiversiTools (code available at https://github.com/josephhughes/DiversiTools), uses the alignment file (produced in the EasySeq pipeline) to analyse the minor genomic variation and predicts the amino acid sequence based on the genomic data. DiversiTools allows an in-depth analysis of viral diversity in each sample, rather than just the consensus/dominant genomic information, as previously described[11]. Data visualisation was conducted in R (version 4.0.2). Wilcoxon rank sum tests were used to determine differences between treatment groups at each time point, using the Rstatix package (version 0.7.0). Schematic figures 1a, 1b and S6 made using Biorender.com.

Data availability

All raw data used in the analysis have been deposited to the National Center for Biotechnology Information (NCBI) Short Read Archive (SRA) (Project Accession Number PRJNA854613) and will be made publicly available upon publication.

Code availability

All custom code used in this study will be made available in a public repository upon publication.

Funding

This work was funded by a US Food and Drug Administration (FDA) Medical Countermeasures Initiative contract (75F40120C00085) awarded to JAH. The article reflects the views of the authors and does not represent the views or policies of the FDA. This work was also supported by the MRC (MR/W005611/1) G2P-UK: A national virology consortium to address phenotypic consequences of SARS-CoV-2 genomic variation (JAH as a co-I). The AGILE platform infrastructure is supported by the Medical Research Council (grant number MR/V028391/1) and the Wellcome Trust (grant number 221590/Z/20/Z). ID-B is supported by PhD funding from the National Institute for Health and Care Research (NIHR) Health Protection Research Unit (HPRU) in Emerging and Zoonotic Infections at University of Liverpool in partnership with Public Health England (PHE) (now UKHSA), in collaboration with Liverpool School of Tropical Medicine and the University of Oxford (award 200907). The Southampton Clinical Trials Unit is a UKCRC registered Clinical Trials Unit (CTU) with NIHR CTU support funding with the collection of samples undertaken from across the NIHR UK Clinical Research Facility Network, including facilities in Liverpool, Southampton, Manchester, Lancashire, and Kings College Hospital. The GCPLab facility is supported by grant funding from Cancer Research UK and the NIHR through the Liverpool Experimental Cancer Medicine Centre (C18616/A25153). The views expressed are those of the authors and not necessarily those of the Department of Health and Social Care or NIHR.

Ethics declarations

All participants provided written informed consent before enrolment. The study protocol was reviewed and approved by the UK Medicines and Healthcare product Regulatory Agency (MHRA) (EudraCT 2020-001860-27) and West Midlands Edgbaston Research Ethics Committee (20/WM/0136).

Author information

Contributions:

Conceptualization: TF, SK, MWC, JAH.

Methodology: ID-B, HG, RP-R, SK, GG, TF.

Software: ID-B, RP-R, XD.

Validation: ID-B, HG, RP-R, JP, AMR, KB, GS, JN, JAH.

Formal analysis: ID-B.

Investigation: ID-B, HG, RP-R, AMR, JP, LEW, RF, CH, DH, SA, JNP, RL, MT, LLL, TE, KB, CJE, SK, JAH, TF.

Resources: GG, TF, SK.

Data Curation: ID-B, YR, GG, GS, JN, EK, CM.

Writing - Original Draft: ID-B, TF, JAH.

Writing Review and Editing: ID-B, HG, RP-R, XD, YR, WG, TF, GG, SK, JAH.

Visualization: ID-B, HG, RP-R, XD.

Supervision: GG, DGL, ACD, SK, MWC, TF, JAH.

Project administration: ID-B, VS, LLL, EK, HR, MT, CW, GG, SK, ACD, TF, JAH.

Funding acquisition: WG, GG, SK, TF, JAH.

Corresponding authors:

Tom Fletcher. E-mail: Tom.Fletcher@lstmed.ac.uk

Julian A. Hiscox. E-mail: julian.hiscox@liverpool.ac.uk