nature research

Corresponding author(s): Florian Klein, M.D., M.Sc.

Last updated by author(s): Aug 29, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

No data were collected except own data (see also below). Flow cytometry data were gained by FACSAriaIII with BD FACSDiva software. ELISA plates were measured by Tecan's Sunrise absorbance microplate reader and associated software. Microscopic pictures were taken with Leica DMI microscope (fluorescence) and a Titan Krios electron microscope (FEI) operating at 300kV (cryo-EM). X-ray diffraction data were collected at the European Synchrotron Radiation Facility (ESRF) at beamline ID23-2 using a Pilatus3 2M detector. B cell receptor repertoire sequence data were generated by an unbiased template-switch-based NGS approach (Kreer et al., manuscript in preparation). Raw read pre-processing was performed with an in-house pipeline primarily based on self-written Python scripts, IgBLAST, Clustal Omega, and the pRESTO toolkit. Antibody sequence data were gained by PCR sequencing at GATC or Eurofins genomics.

Data analysis

Antibody sequence analysis: Sequencing chromatograms were filtered for a mean Phred score of 28 and a minimal length of 240 nt. Remaining sequences were annotated with IgBLAST (Ye et al., 2013) and trimmed to extract only the variable region from FWR1 to the end of the J gene. Base calls within the variable region with a Phred score below 16 were masked and sequences with more than 15 masked nucleotides, stop codons, or frameshifts were excluded from further analyses. In order to identify clonally related sequences within a single subject, we grouped all productive heavy chain sequences of that particular subject by identical VH genes, determined the pairwise Levenshtein distance for their CDRH3s, and assigned an individual clone-number to sequence groups that share the same VH gene and have a minimal CDRH3 identity of 75% (with respect to the shortest CDRH3). 100 rounds of input sequence randomization and clonal assignment were performed and the result with the lowest number of remaining unassigned (non-clonal) sequences was selected for downstream analyses.

NGS sequence processing: Raw read pre-processing was performed with an in-house pipeline primarily based on self-written Python scripts, IgBLAST (Ye et al., 2013), Clustal Omega (Sievers et al., 2011), and the pRESTO toolkit (Vander Heiden et al., 2014). Raw reads were filtered for a mean Phred quality score of 25 and a mean read length of 250 bp. Unique molecular identifiers (UMIs) were extracted and IgBLAST was used to pre-annotate reads. Each read pair was annotated with its top V gene call, its UMI, and an additional 18 nt molecular identifier (MID) starting 12 nt after the framework region (FWR) 3. UMIs were grouped and sequences (referred to as „collisions") were removed if they differed from the most abundant V gene call or had more than 1 nt difference to the remaining reads in their UMI group. We assumed that reads with the same V gene/MID and an UMI edit distance of 1 nt are more likely to represent RT,

PCR, or sequencing errors within the UMI than real different molecules. In order to rescue those erroneous UMIs, we re-grouped colliding reads and remaining single UMI reads by their MID and re-defined MID groups as the same UMI group, if they shared the same V gene and their UMI had no more than 1 nt difference. UMI groups were aligned with clustal omega and then collapsed to build consensus reads. To this end, we weighted each base call by its quality (1 - error probability) and generated sums for different base calls over each position to account for both, total abundance and the corresponding quality. The base with the highest sum was taken as the consensus base. Paired consensus reads were assembled to one sequence using pRESTOs AssemblePairs module with a minimal overlap of 6 nt. Assembled full-length sequences were annotated with IgBLAST and only productive sequences with full sequence annotation were kept for downstream analysis. V gene usage, CDR3 length and V gene germline identity distributions were determined from all final sequences without further collapsing.

Cryo-EM data analysis: The entire data processing was conducted using the cryoSPARC V2 suite (Punjani et al., 2017). Movies were motion-corrected and contrast transfer functions were fitted using CTFFIND4 (Rohou and Grigorieff, 2015). Templates for auto picking were derived by 2D classification of manually picked particles. Following template-based autopicking a total of 109,407 and 123,880 good particles were selected for reconstruction of EBOV GP/3T0331 and EBOV GP/4m0368, respectively based on iterative reference-free 2D classifications. Initial reference maps were calculated using Ab-initio reconstruction and high-resolution maps were obtained using the non-uniform 3D refinement, while imposing C3 symmetry. Working maps were locally filtered after calculating local resolution estimates.

Analysis of antibody protein sequences: The calculations of phylogenetic trees and sequence alignments were done with Geneious (v10). Sequence similarity of CDRH3 was also determined by Geneious software using BLOSUM (BLOcks Substitution Matrix) with residue-specific and hydrophilic gap penalties (open gap penalty of 10, extended gap penalty 0.2. For comparing the heterogeneity of sequences within the same donor with that among different donors, pairwise alignment scores among all sequences were computed by needle (with default parameters) and grouped accordingly. The difference of both distribution of alignment scores was assessed with a Mann-Whitney-U test. Sequence logo plots were generated with WebLogo3 followed by minor adjustments.

Flow cytometry, ELISA and statistical analysis: Flow cytometry analyses and quantifications were performed with FlowJo10. EC50 values, means and standard deviations of serum and antibody reactivities were calculated from technical replicates using GraphPad Prism (v7). Tests for significance were conducted using GraphPad Prism and R (R Core Team; 2018).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

In the manuscript we present the data necessary to interpret and replicate our results. All experiments were performed as described in the method section. All antibody sequences are currently being processed to be uploaded to GenBank.
EM structures have been deposited at PDB and are hold until final publication.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size of investigated subjects was chosen based on previous studies (Davis et al., Cell 2019; Corti et al., Science 2016; Klein et al., J Exp med. 2012) to sufficiently inform on proof of concept.
Data exclusions	Sequencing chromatograms were filtered for a mean Phred score of 28 and a minimal length of 240 nt. Remaining sequences were annotated with IgBLAST (Ye et al., 2013) and trimmed to extract only the variable region from FWR1 to the end of the J gene. Base calls within the variable region with a Phred score below 16 were masked and sequences with more than 15 masked nucleotides, stop codons, or frameshifts were excluded from further analyses. Exclusion criteria were not pre-established and first used in this study.
Replication	To test reproducibility experiments were repeated with the same setup at least two times. The results of single cell sort were verified by repeating the flow cytometric staining at least one time (except for EV07 due to low amount of samples). All attempts of replication were successful.
Randomization	Randomization was not applicable in this study.

Blinding	HepG2 cell assay to determine autoreactivity: Stained slides were mounted and and fluorescence intensity was evaluated by four independent persons. The investigators were blinded during data collection.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	1.) Mouse anti-human CD20-AF700 (clone 2H7), BD Bioscience, Cat#560631, RRID: AB_2687799, 2.) Mouse anti-human IgG-APC (clone G18-145), BD Bioscience, Cat#550931, RRID: AB_2738854, 3.) Goat anti-human IgG-HRP, Southern Biotech, Cat# 2040-05, 4.) Human anti-EBOV GP ADI-15758, GenBank: KU602137-8, 5.) Human anti-EBOV GP ADI-15999, GenBank: KU602605-6, 6.) Human anti-EBOV GP ADI-16037, GenBank: KU602673-4, 7.) Human anti-EBOV GP mAb100, PDB: 5FHB, 8.) Human anti-EBOV GP mAb114, PDB: 5FHA, 9.) Human anti-EBOV GP KZ52, PDB: 3INU. 10.) FITC-conjugated anti-human IgG (NOVA Lite HEp-2 ANA Kit; Inova Diagnostics)
Validation	1.) Mouse anti-human CD20-AF700 (clone 2H7): validated by Flow cytometry 2.) Mouse anti-human IgG-APC (clone G18-145): validated by Flow cytometry 3.) Goat anti-human IgG-HRP, Southern Biotech: validated by ELISA, FLISA and flow cytometry 4.) Human anti-EBOV GP ADI-15758, GenBank: KU602137-8: published in Rapid and large-scale isolation of potent neutralizing antibodies from a survivor of the 2014 Ebola virus outbreak; Bornholdt et al.; Science (2016) 5.) Human anti-EBOV GP ADI-15999, GenBank: KU602605-6, published in Rapid and large-scale isolation of potent neutralizing antibodies from a survivor of the 2014 Ebola virus outbreak; Bornholdt et al.; Science (2016) 6.) Human anti-EBOV GP ADI-16037, GenBank: KU602673-4, published in Rapid and large-scale isolation of potent neutralizing antibodies from a survivor of the 2014 Ebola virus outbreak; Bornholdt et al.; Science (2016) 7.) Human anti-EBOV GP mAb100, PDB: 5FHB, published in Structural and molecular basis for Ebola virus neutralization by protective human antibodies; Misasi, J. et al.; Science (2016) 8.) Human anti-EBOV GP mAb114, PDB: 5FHA, published in Structural and molecular basis for Ebola virus neutralization by protective human antibodies; Misasi, J. et al.; Science (2016) 9.) Human anti-EBOV GP KZ52, PDB: 3INU, published in Techniques and tactics used in determining the structure of the trimeric ebolavirus glycoprotein; Lee, J.E. et al.; Acta Crystallogr. (2009) 10.) FITC-conjugated anti-human IgG (NOVA Lite HEp-2 ANA Kit; Inova Diagnostics) validated by Assay quality control

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T (ATCC); HEK293E (ATCC); HEK293F (National Research Council Canada), FreeStyle™ 293-F Cells (Thermo Fisher Scientific); Vero C1008 (ATCC); HepG2 cells (NOVA Lite HEp-2 ANA Kit; Inova Diagnostics)
Authentication	Cell lines were not authentificated.
Mycoplasma contamination	Cell lines were tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	In this study no commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	NOD.Cg-Rag1tm1Mom Il2rgtm1Wjl/SzJ (The Jackson Laboratory). Each group contained two female and one male with age between 22-36 weeks.
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve samples collected from the field.
Ethics oversight	LANUV

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	All individuals participated in the Phase I Trial to Assess the Safety, Tolerability and Immunogenicity of an Ebola Virus Vaccine (rVSVΔG-ZEBOV-GP; NCT02283099) and were vaccinated with either 3 x 10^5 (EV01, EV02, EV04 and EV06) or 3 x 10 6 pfu (EV03 and EV05). The age of the selected individuals was 55, 26, 42, 51, 44 and 45 years (EV01-EV06). Two of the individuals were female (EV01 and EV02). Samples were taken after 18.5, 18.5, 24, 20.5, 25 and 26 month after vaccination (EV01-EV06). EV07 is a health care worker (age 47) that was vaccinated with 2 x 10^7 in December 2018.
Recruitment	Individuals studied were previously enrolled in the Phase I Trial to Assess the Safety, Tolerability and Immunogenicity of an Ebola Virus Vaccine (rVSVΔG-ZEBOV-GP; NCT02283099) and were vaccinated with either 3 x 10 5 or 3 x 10 6 pfu. From this cohort, six individuals were enrolled in an observational study (INA; 16-054) at the University Hospital of Cologne to collect serum and PBMC samples. Recruitment was open for all participants in trial NCT02283099 that were enrolled at the site in Hamburg. All participants interested in our study were enrolled. No potential self-selection bias or other bias are known.
Ethics oversight	The protocol was approved by the Institutional Review Board (IRB) of the University of Cologne, Germany.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	https://www.viomedo.de/klinische-studien/6383/identifikation-neutralisierender-antikoerper-chronische-schwerwiegende-infektionen
Study protocol	https://www.viomedo.de/klinische-studien/6383/identifikation-neutralisierender-antikoerper-chronische-schwerwiegende-infektionen
Data collection	Data collection was performed at enrollement after informed consent.
Outcomes	Non-Interventional study. No outcome paramenters were defined except number of collected samples.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Peripheral blood mononuclear cells (PBMCs) were enriched for B lymphocytes by magnetic cell separation (MACS) with CD19-microbeads (Miltenyi Biotec) according to manufacturer's instruction. MACS LS separation columns (Miltenyi Biotec) were used to enrich labeled cells. B cells were spun and suspended in PBS (Gibco) with 2% FBS (Merck) and 2 mM EDTA (Thermo Fisher Scientific) and subjected to fluorescence staining with DAPI (Thermo Fisher Scientific), anti-huCD20-Alexa Fluor 700 (BD), anti-huIgG-APC (BD), and EBOV GPΔTM-DyLight 488 (Microscale Antibody Kit, Thermo Fisher Scientific). Cells were stained for 20 min at 4°C and live CD20+ IgG+ EBOV GP+ cells were sorted.

Instrument	FACSAriaIII (Becton Dickinson) in a single cell manner into 96-well plates.
Software	BD FACSDIVA™ SOFTWARE
Cell population abundance	Described in Figure 1.
Gating strategy	Gating strategies are shown in Fig 1 and Supplementary Fig. 2.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.