***To the Author****— Please review the editorial comments and requests below and confirm that changes have been made in the manuscript in the right-hand column.* ***This document******must be uploaded*** *as a related manuscript file.*

Please see our [final file submission checklist](https://www.nature.com/documents/commsj-file-checklist.pdf) for information about submitting your revised documents.

|  |  |
| --- | --- |
| **Files and General Policies** | |
| **Main manuscript file must be in Microsoft Word or LaTeX format.**  LaTex and Tex article source files must be accompanied by the compiled PDF for reference. If submitting a Tex/LaTex file, the bibliography must be submitted separately (as a .bib file) or contained within the .tex file. | Submitted as LaTex file. Bibliography submitted as a separate file. |
| **Each Figure must be provided as a separate file** and must be supplied whole, with all panels included in a single document. Figures should be provided at a minimum resolution of 300 dpi at final size.  Figure files must only contain images (please also leave out labels such as “Figure 1” etc). Figure captions must instead be included within the main manuscript file, grouped together at the end of the document. | Done. Figures 2 and 9 are single-column-figures and the rest are double-column. |
| All figures, tables, and supplementary items must be cited in the manuscript and **numbered in the order in which they appear**. | Done |
| **If present, Tables** in the main manuscript must be provided in an editable format and should be grouped together at the end of the main manuscript file. | Done |
| If applicable, please ensure that all **equations** are supplied in an editable format upon resubmission. If equations are numbered they must be numbered sequentially. | There are no equations. |
| Please check whether your manuscript contains **third-party images**, such as figures from the literature, stock photos, clip art or commercial satellite and map data. We strongly discourage the use or adaptation of previously published images, but if this is unavoidable, please request the necessary rights documentation to re-use such material from the relevant copyright holders and return this to us when you submit your revised manuscript.  An appropriate permissions statement must be present in the relative figure caption for any third-party images.  **Please also check whether any images in the manuscript were created using generative AI tools, such as Midjourney or DALL-E. In accordance with** [**Nature Portfolio policy**](https://www.nature.com/nature-portfolio/editorial-policies/ai)**, we cannot permit the use of these images for publication.** | No third-party images or images created with AI. |
| **Please check that you have not copied any text directly from published work** (even your own) without clear attribution, including one or more references. We run a plagiarism detection software and may need to request additional changes if we identify large blocks of identical text. | We have not. |
| If anyone is cited in ‘personal communication’, please arrange for them to e-mail [commsbio@nature.com](mailto:commsbio@nature.com), giving permission for this statement to be published. | Not relevant |
| An updated **editorial policy checklist** that verifies compliance with all required editorial policies must be completed and uploaded with the revised manuscript. All points on the policy checklist must be addressed; if needed, please revise your manuscript in response to these points.  <https://www.nature.com/documents/nr-editorial-policy-checklist.pdf>.  Please note that this form is a dynamic ‘smart pdf’ and must therefore be downloaded and completed in Adobe Reader. This file will not open in an internet browser. | Done |
| The **reporting summary** will be published alongside your manuscript therefore **it needs to accurately represent your wor**k. In this case, please take a closer look at the reporting summary and make sure things are completed correctly. If an item does not apply, for example human participants, please check the NA box next to that item. No section should be left blank.  Also, please make sure to include your name and date at the top of the document.  If you require a new Reporting Summary form, please download it here: <https://www.nature.com/documents/nr-reporting-summary.pdf>.  Please note that this form is a dynamic ‘smart pdf’ and must therefore be downloaded and completed in Adobe Reader. This file will not open in an internet browser. | Done |
| Your paper will be accompanied by a brief summary when it is published on our homepage.  In the box to the right, please provide us a with a suggested summary of **no more than 250 characters including spaces** in the style below.  *An fMRI study on awake dogs suggests that non-primary auditory dog brain regions respond more to dog- and/or infant-directed speech than to adult-directed speech, especially when spoken by females.* | Glycosyltransferases that build bacterial surface polysaccharides were classified into 17 new CAZy families. A glycan analysis reveals that the enzymes within each family transfer similar sugars even across Gram-negative and Gram-positive bacteria. |
| **ORCID**  *Communications Biology* is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as ‘corresponding author’ create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS) prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. For more information please visit <http://www.springernature.com/orcid>.  For all corresponding authors listed on the manuscript, please follow the instructions in the link below to link your ORCID to your account on our MTS before submitting the final version of the manuscript. If you do not yet have an ORCID you will be able to create one in minutes. **We are, however, unable to do this on your behalf.**  <https://www.springernature.com/gp/researchers/orcid/orcid-for-nature-research>  IMPORTANT: All authors identified as ‘corresponding author’ on the manuscript must follow these instructions. Non-corresponding authors do not have to link their ORCIDs but are encouraged to do so. Please note that it will not be possible to add/modify ORCIDs at proof. Thus, if they wish to have their ORCID added to the paper they must also follow the above procedure prior to acceptance.  To support ORCID's aims, we only allow a single ORCID identifier to be attached to one account. If you have any issues attaching an ORCID identifier to your MTS account, please contact the Platform Support Helpdesk at <http://platformsupport.nature.com/> | Done |
| We regularly highlight papers published in *Communications Biology* on the journal’s **Twitter** account (@CommsBio). If you would like us to mention authors, institutions, or lab groups in these tweets, please provide the relevant twitter handles in the right-hand column. |  |
| We would welcome the submission of material for the **‘Featured Image’** section on the Communications Biology home page. Images should relate to the content of your manuscript but need not be contained within the paper. **Photographs and aesthetically interesting images are preferred; diagrams are generally not used. Per** [**Nature Portfolio policy**](https://www.nature.com/nature-portfolio/editorial-policies/ai)**, we cannot use any images that were created using a generative AI image tool like Midjourney or DALL-E (among others).**  Suggestions should be uploaded as a Related Manuscript file. Please provide 1200x675-pixel RGB images. You will also need to submit a completed [Image License to Publish](https://www.nature.com/documents/snl-image-ltp.docx).  Unfortunately, we cannot promise that your suggestions will be used. | We have no suggestions |
| **Supplementary information** | |
| **Supplementary Information Format and referencing**   * Supplementary Figures, small Tables, and any supplementary text must be provided **in a single PDF**. Figures and their captions should be presented together. We noted that some of your supplemental Figures span several pages. You might want to introduce something like “Fig SX – continued” for Fig S4 and 6.   + If you include a title page, please check that the title and author list matches the main manuscript. * **All Supplementary items must be referred to in the manuscript**, and items must be mentioned in numerical order. Please do not include general references to “Supplementary Material”; instead refer to specific items. * Additional files can be provided as **Supplementary Data** (Excel files, text files, .zip folders), **Supplementary Movies**, **Supplementary Audio**, or **Supplementary Software** (.zip folder)   Supplementary Information files will be uploaded with the published article as they are submitted with the final version of your manuscript. Any highlighting or tracked changes should be removed from the file. | We have added “Supplementary Figure 4 – continued.”. We have included a title page. All items are cited, and the tables are submitted as separate excel files. |
| **Supplementary items must be cited in a consistent format**. Names of items in the Supplementary file(s) must match those used in the main manuscript.  **We recommend using the following naming formats:** Supplementary Figure 1, Supplementary Table 1, Supplementary Data 1, Supplementary Note 1, and Supplementary References. | Done |
| **IF YOUR Supplementary Information contains large tables and other data types:**  **Large or data-rich tables and other data types:** We strongly recommend depositing these to suitable repositories (such as Figshare, Dryad, or a data type-specific repository if one exists).  Otherwise, these must be supplied as **Supplementary Data** files separate from the Supplementary Information pdf. Each file must be labelled as Supplementary Data 1, etc. | The two supplementary tables are uploaded as excel files. |

|  |  |
| --- | --- |
| **It’s mandatory to provide access to the numerical source data for graphs and charts:** We strongly recommend depositing these to suitable repositories (such as Figshare, Dryad, or a data type-specific repository if one exists).  Otherwise, all source data underlying the graphs and charts presented in the main figures must be uploaded as **Supplementary Data** (in Excel or text format). If you have more than 5 such excel files, please consolidate them as separate tabs in one Excel file. |  |

|  |  |
| --- | --- |
| For any Supplementary Files such as those mentioned above that are not included your combined PDF (e.g. Supplementary Data, Movies, Audio, Software), please provide a title and description for each file **here in the column to the right**.  For example:  File name: Supplementary Data 1  Description: The source data behind the graphs in the paper | **Supplementary Table 1.** O-Lig seeds. List of seed sequences used for creation of the O-Lig family.  **Supplementary Table 2.** BP-Pol seed. List of BP-Pol sequences that were considered in this work. |
| **Title Page** | |
| Please ensure that the author list provided in our manuscript tracking system matches the author list in the main manuscript. | Done |
| **Please check that your author list and affiliations (not present currently) comply with the following:**   * We recommend that author first names be written out in full rather than provided as initials. * Titles (Professor, Dr, etc.) must not be included. * Affiliations must be numbered in the order of their first appearance in the author list. * Each affiliation must contain only one address. * Each affiliation must include the institution, city and country. The name of the city must be provided separately from the institution even if it is a part of the institution name; e.g. ‘University of Science and Technology Beijing, Beijing, China’ * Author tagging statements are limited to the following two options: "These authors contributed equally" and "These authors jointly supervised this work", with no more than one of each tag permitted. * Where relevant, “present address” must be provided separately as the final affiliation. * At least one corresponding author must be designated, and an e-mail address must be provided for each corresponding author (with a limit of one e-mail address per author). | Done |
| **Manuscript title**  Please ensure the title clearly describes the central finding of the paper. We recommend writing the title as a declarative statement of approximately 15 words or fewer.  Be sure to include any key species, protein names, or gene names to ensure optimal retrieval of the paper in database searches. | Done |
| **Abstract**  The abstract should be **accessible to non-specialists** and avoid jargon and abbreviations. Please write the abstract in **the present tense**.  We recommend structuring the abstract as follows:  Opening statement explaining why this area of research is important. A sentence explaining the gap in knowledge that your research will address. **Here we show** (or an equivalent phrase), and then the major results and conclusions of the paper. Final sentence indicating any broader impacts and how this research will be used in the future. | Done |
| **Main text** | |
| **Format of the main text**  Please ensure your manuscript includes the following sections, presented in this order:   1. “**Introduction”**: The background and rationale for the work. The final paragraph should be a brief summary of the major results and conclusions. The results of the current study must only be discussed in this final paragraph. The Introduction should contain no references to figures or tables. Do not include subheadings. 2. “**Results” or “Results and Discussion”**. This should be split into subheaded sections; we recommend 1 subheading per main figure or table. Figures should not be embedded in the text but submitted separately.    1. Do not use more than 1 layer of subheadings.    2. A “Conclusions” paragraph can be included **only if the results and discussion are combined into a single section.** 3. “**Discussion” (optional)**, without subheadings. 4. **Methods**, which should be split into subheaded sections. Do not use more than 1 layer of subheadings.   **To improve readability**, we recommend that the main text (Introduction, Results and Discussion) be limited to approximately 5000 words or fewer. | Done |
| **Please include exact p-values where possible.** We ask that you also include the name of the statistical test and the estimated effect size. If applicable, please also include the confidence interval. | We have no statistical tests |
| Avoid the use of the word “**significant**” unless referring the results of a statistical test. | Done |
| **Please check that all gene and mRNA names are in italics**. Protein names should not be in italics. Please confirm that only official gene/protein symbols are used and that species names are in italics. | Done |
| Language such as “new”, “novel”, “for the first time”, “unprecedented”, etc, should be avoided, or qualified with “**to the best of our knowledge**” or similar, because it often leads to unproductive controversy. **Novelty should be made clear from the context.** | Done |
| All data that support the conclusions drawn must be presented in the manuscript unless they are published elsewhere. **We do not allow statements of “data not shown”.** | Done |
| **Style checks**  -Please avoid abbreviating terms unless they are used five or more times. We ask that you avoid all non-standard 2 letter abbreviations.  -Our style does not allow for the use of bold or italics to convey emphasis. If applicable, please remove such usage from your manuscript.  -Use of speech marks around words or phrases should be avoided; if a phrase is non-standard, please explain the meaning instead; otherwise they are usually unnecessary. | Done |
| **Display items** | |
| **Figure captions/legends**  Figures must have a title that will appear above the Figure **and** a legend that will appear below the Figure (see e.g. <https://www.nature.com/articles/s42003-020-1059-1/figures/1>)  The Figure title must describe the Figure as a whole and must not contain reference to specific figure panels.  We noticed that Fig 5 and Fig 6 might need a color code legend within the figure.  The Figure legend must refer to and describe **all panels**. Abbreviations, symbols, colors, and shading present in the Figure must be defined. Please write out the symbols/colors in words (blue circles, red dashed line, etc.) within these definitions.  **All figure panels must be labelled using lower case letters and cited as such in the text. Please refrain from referring to sections of figures as top/bottom/left/right/, etc.** | We have added color code legends in figures 5 and 6. |
| **Axis and panel labels will be published as received.** We recommend using a sans-serif font such as Arial or Helvetica. | Done |
| **Tables in the main text**  Please check that any Tables comply with the following (does not apply to Supplementary Tables):   * Do not include shading or colors. All Tables must contain black and white text only. * Any bold/italic formatting must be either removed or defined clearly in a Table footnote. * Where Tables contain images, each image should appear in its own cell in the absence of any text. * All Tables must have a brief title. | Done |
| Please pay close attention to our [Digital Image Integrity Guidelines](https://www.nature.com/nature-research/editorial-policies/image-integrity). Also ensure that you retain unprocessed data and metadata files after publication, ideally archiving data in perpetuity, as these may be requested during the peer review and production process or after publication if any issues arise. | Done |
| **Methods** | |
| Please ensure that all information present in the Reporting Summary is also in the manuscript. This information is usually most appropriate in the Methods section. | Done |
| **We allow unlimited space for Methods.** The Methods must contain sufficient detail such that the work could be repeated. It is preferable that all key methods be included in the main manuscript, rather than in the Supplementary Information.  **Please avoid use of “as described previously”** or similar, and instead detail the specific methods used with appropriate attribution. | Done. |
| **Data Policies** | |
| Please add a **Data Availability statement**.  The **Data Availability** **statement** must include:   * Access details for deposited data, including repository name and unique data ID. * How source data can be obtained. * A statement that all other data are available from the corresponding author (or other sources, as applicable) on reasonable request. **Note that ‘available upon request’ is only appropriate if immediate data access has not been mandated by our policies or by the editors.**   See here for more information about formatting your Data Availability Statement: <http://www.springernature.com/gp/authors/research-data-policy/data-availability-statements/12330880> | Done |
| *Communications Biology* has a strong preference for all data to be deposited in an approved repository. In some cases, data deposition may be required by the editor.  **We recommend the following data repositories:**   * [GenBank](https://www.ncbi.nlm.nih.gov/genbank/) (all DNA sequence data) * [NHGRI-EBI GWAS Catalog](https://www.ebi.ac.uk/gwas/) (GWAS summary statistics) * [PGS Catalog](https://www.pgscatalog.org/) (polygenic risk scores) * [Gene Expression Omnibus](https://www.ncbi.nlm.nih.gov/geo/) (Microarray or RNA sequencing data) * [Sequence Read Archive](https://www.ncbi.nlm.nih.gov/sra) (WGS or WES data) * [Protein Data Bank](https://www.wwpdb.org/) (protein structural data) * [OSF](https://osf.io/) (neuroimaging raw data and EEG/EMG/MEG raw data) * [Neurovault](https://neurovault.org/) (unthresholded statistical maps, parcellations, and atlases produced by MRI and PET studies) * [Image Data Resource](https://idr.openmicroscopy.org/about/) (microscopy data) * [PRIDE](https://www.ebi.ac.uk/pride/) (proteomics data)   Data types without a specific repository can be deposited in a generalist repository, such as [figshare](https://figshare.com/) or [Dryad](https://datadryad.org/stash).  For an up-to-date list of approved repositories, please visit <https://www.springernature.com/gp/authors/research-data-policy/repositories/12327124>. | The content of the CAZy families is available at the CAZy website. The lists of seed proteins are in the supplementary tables. |
| **Data citation**  Please cite datasets stored in external repositories that provide a doi (such as Figshare, Zenodo etc) **in the main reference list**, example below:  110. Alam, A., Wang, T. & Chiosis, G. Cytoscape files—Systems-level analyses of protein-protein interaction network dysfunctions via epichaperomics identify cancer-specific mechanisms of stress adaptation [Data set]. *Zenodo* <https://doi.org/10.5281/zenodo.7433980> (2022).  For previously published datasets, we ask authors to cite both the related research articles and the datasets themselves.  For more information on how to cite datasets in submitted manuscripts, please see our [data availability statements and data citations policy](https://www.nature.com/documents/nr-data-availability-statements-data-citations.pdf). | We do not use any datasets stored in external repositories. |
| **Code availability**  Please include a Code Availability statement if your manuscript contains custom code, indicating whether and how the code can be accessed, including any restrictions to access. **In some cases, the editor may require that code be made immediately available.**  This section should also include information on the versions of any software used, if relevant, and any specific variables or parameters used to generate, test, or process the current dataset. **The Code Availability statement must be provided as a separate section after the Data Availability section.**   Please see our policy on code availability for more information. <http://www.nature.com/sdata/for-authors/editorial-and-publishing-policies#code-avail>  In addition to making the custom code available, please ensure that the version of the code/software described in the paper is **deposited in a DOI-minting repository** (eg, Zenodo) and that this DOI is also cited in the main Reference list. | We have added a Code Availability statement with reference to the GitHub repository for Aclust. |
| **End Notes** | |
| Please check that your bibliography complies with the following:   * Your bibliography should start with the heading “References”. The references must be numbered in the order of appearance in the text, then tables, then figures. * Any in-text citations to references (e.g. "Gupta et al. show...") should be followed by their corresponding reference citation number from the reference list. * Manuscript citations must include journal title, article title, volume number, page or article number or DOI, and year of publication. *Communications Biology* uses standard *Nature* referencing style. See [here](https://www.nature.com/commsbio/submit/submission-guidelines#references) for full details * No publication can be present more than once in the reference list. * No footnotes are permitted in the references or elsewhere. Text should be incorporated into the main text, the Methods section, or the Supplementary Information instead. * Websites should only be listed in the references if they are in common use or curated. * Where possible, preprints in the reference list should be updated with details of the published, peer-reviewed paper. * Citations should be formatted in the text using **superscript numbers**. | Done |
| Please ensure to provide a '**Competing interests**' statement using one of the following standard sentences:   * The authors declare the following competing interests: [specify competing interests] * The authors declare no competing interests.   See our competing interests policy for further information: <https://www.nature.com/nature-research/editorial-policies/competing-interests> | Done |
| Please ensure to provide an **'Author Contributions**' section **that individually lists the specific contribution of each author to the work**. Each author must be referred to by name or initials. Where multiple authors possess identical initials, they must be clearly disambiguated from one another.  See our author contributions policy for further information: <https://www.nature.com/nature-research/editorial-policies/authorship#author-contribution-statements> | Done |
| No separate funding section is permitted. If applicable, please include your funding information in Acknowledgements instead. | Done |