Supplementary materials

Journal: *Microorganisms*

Title: BNRdb: A Manually Curated Data Resource for Biological Nitrogen Removal

Tanyaradzwa R. Ngara[1,#], Peiji Zeng[1,#] and Houjin Zhang[1,*]

[1] *Department of Biotechnology, College of Life Science and Technology, Huazhong University of Science and Technology, MOE Key Laboratory of Molecular Biophysics, Wuhan 430074, China*.

* Corresponding author: Dr. Houjin Zhang,

Address: College of Life Science and Technology, Huazhong University of Science and Technology, MOE Key Laboratory of Molecular Biophysics, Wuhan 430074, China.

Email address: hjzhang@hust.edu.cn

Telephone: +86 27 87793085

[#]These authors contributed equally.

Annotation Procedure of Nitrogen-removing Microbes

Table S1. Description of entry fields used to annotate BNR microbes based on functional genes that encode functional enzymes involved in BNR processes.

Field Name	Description
GenBank ID	Unique database identifier for a particular key gene
Microorganism	Microorganism that possess the key gene
Isolation Source	The geographical source of the organism from which the sequence was derived
Taxonomy	Identification and Classification of the microorganism
Encoding Gene	Approved name of the gene
Enzyme Name	Approved common name of the protein encoded for by the gene
DNA size(bp)	Size of the DNA of the gene
Nucleotide FASTA sequence	External link to NCBI to access the nucleotide sequence of the gene in FASTA format
UniProtKB ID	External link to UniProtKB database
Protein GenBank IB	External link to the NCBI protein sequence
Length (aa)	Number of amino acids in the canonical sequence
Protein FASTA sequence	External link to the NCBI amino acid sequence of the gene in FASTA format
Reference	Name of authors of the research article
Title	The name of the academic paper that summarizes the main ideas of the study
PMID	External link to PubMed
Research Link	External link to full article
Abstract	Quick overview of the academic paper

Table S2. Description of fields used to annotate nitrogen-removing microbial strains based on 16S rRNA gene analysis.

Field Name	Description
GenBank ID	Unique database identifier for BNR microbe identified by their ability to be involved in any of the BNR processes
Microorganism	Microorganism that has the ability to perform BNR
Isolation Source	The geographical source from which the microorganism was taken
Taxonomy	Identification and Classification of the microorganism
Electron Donor/ Energy Source	A reactant that donates electrons in an oxidation-reaction in an oxidation-reduction reaction
Final product	The intermediary and final end products of the BNR process under study
Respiration	Ability of the BNR microbe to perform BNR proccesses in anaerobic or aerobic conditions
Electron Acceptor	The nitrogenous compound/ intermediate consumed or reduced
Reference	Name of authors of the research article
Title	The name of the academic paper that summarizes the main ideas of the study
PMID	External link to PubMed
Research Link	External link to full article
Abstract	Quick overview of the academic paper

Explore and browse interface

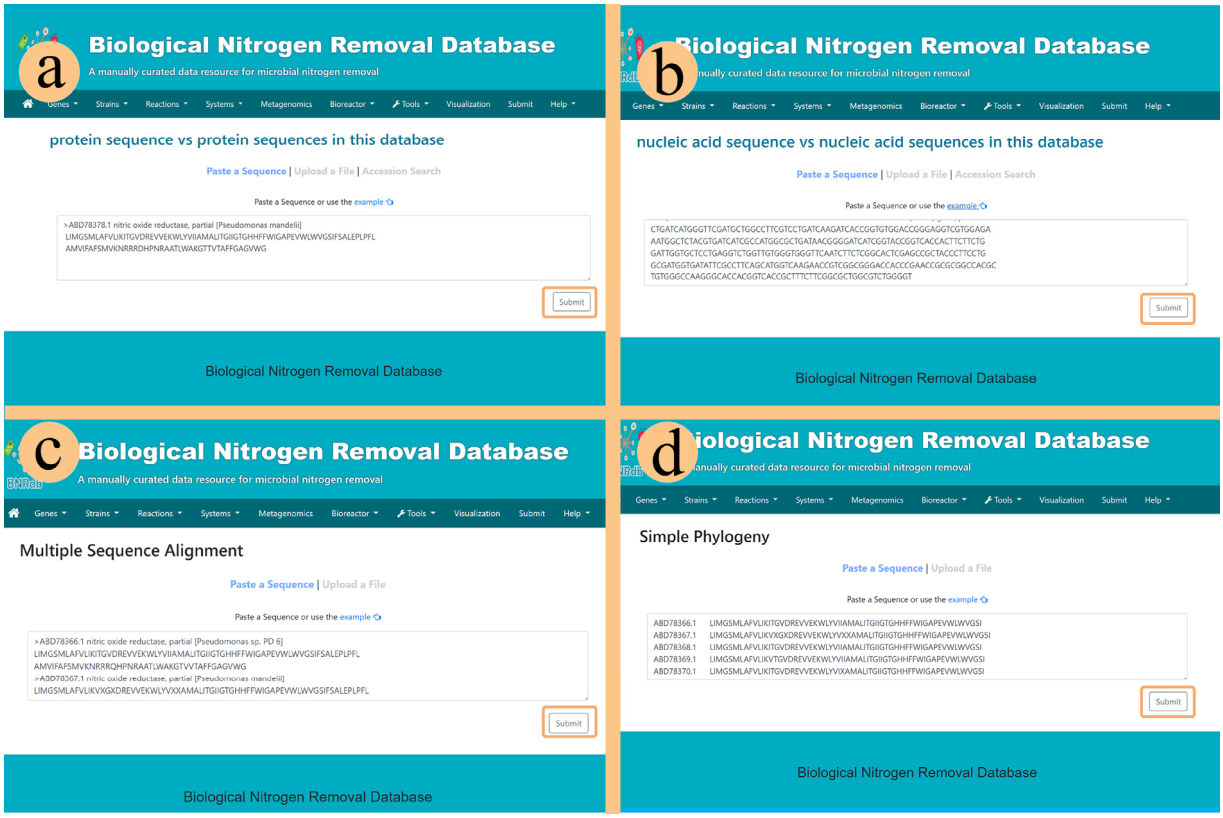

Figure S1. Overview of the BNRdb user-interface to access data. (A) BLAST search and (B) HMMER Users can perform local similarity searches between nucleotide and protein sequences, respectively. (C) Multiple sequence alignment using Clustal O. (D) User can use the simple phylogeny to infer homology and the evolutionary relationship between the sequences studies by submitting their aligned sequences either by pasting it in or uploading the file with the aligned sequences.

BLAST search results

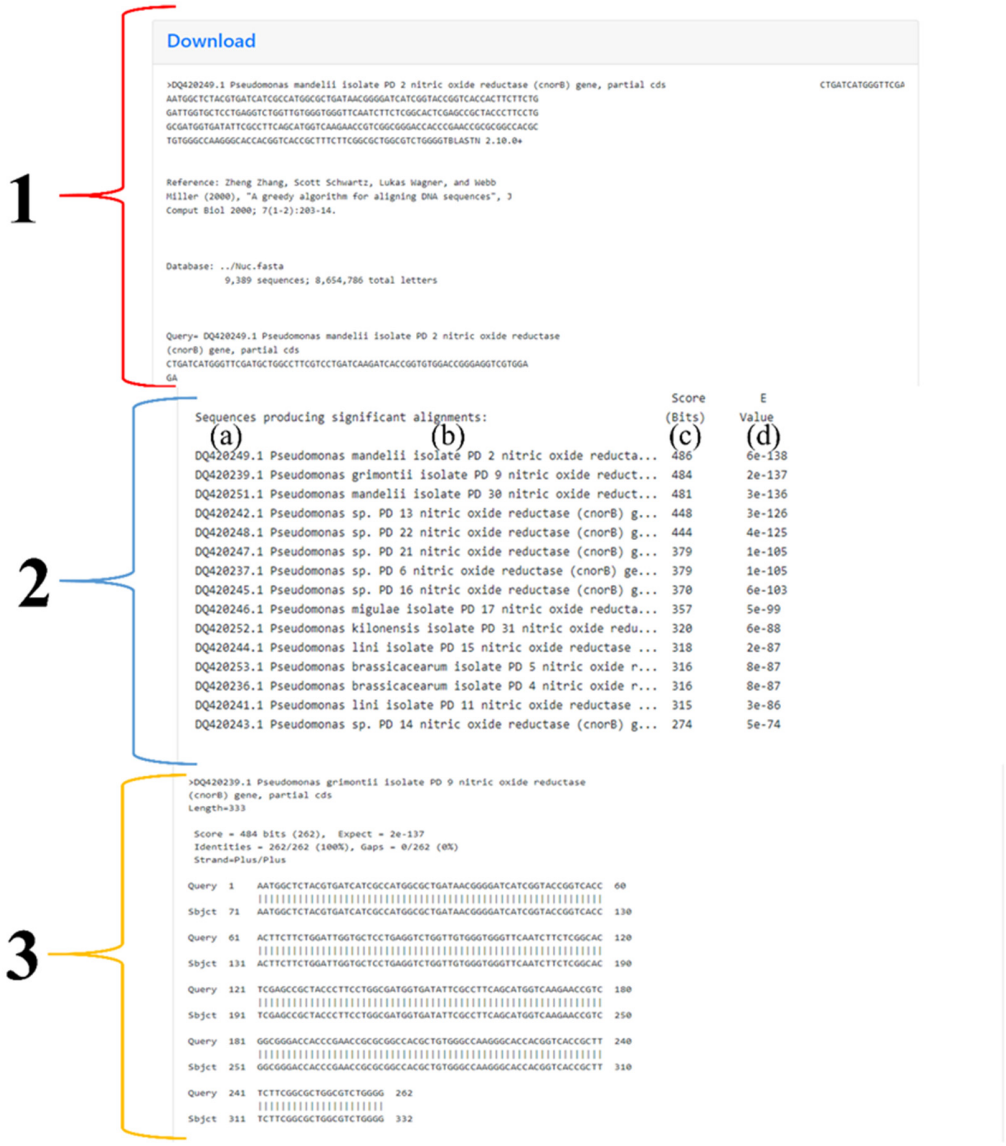

Figure S2. Description of fields of the BLAST hits on the query sequence. (1) the header of the BLAST report: the top line shows information about the query sequence. The scientific article that describes BLAST is then cited. Followed by the summary of the database searched. The length of the query sequence is also shown. (2) the one-line description in the BLAST report: Each line has four fields (a) accession number; (b) description of sequence; (c) the score of the alignment in bits. Hits found at the top of the list have a higher score; (d) with the E-value providing an estimate of statistical significance. (3) Pairwise sequence alignment from the BLAST report. The alignment is

preceded by the accession number, description line and the length of the matched sequence in nucleotides. This is followed by the bit score (with the raw score shown in parenthesis) and the E-value. Next comes the line which shows information concerning the number of identical residues in the pairwise alignment (identities) and the number of gaps the alignment has if any. Lastly, the actual alignment is displayed, with the query sequence on top and the database sequence match below (which labelled Sbjct). The numbers on the left and right help show the position in the nucleotide sequence. Note that (-) indicates deletions and insertions within the sequence. Similarities between the sequences are indicated by a line between the two sequences. If the query and subject sequence do not have similar nucleotide bases at a given point the area between the two sequence is presented as blank.

The phylogenetic tree obtained in Phylotree.js

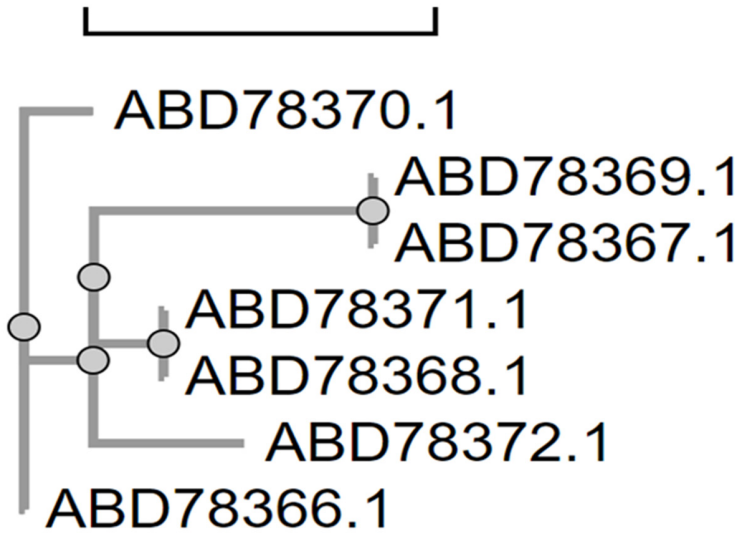

Figure S3. The phylogenetic tree obtained in PHYLIP for microbial denitrifiers that possess the nitric oxide reductase gene using the example sequences provided as references to use in BNRdb.

BNR Treatment Systems page

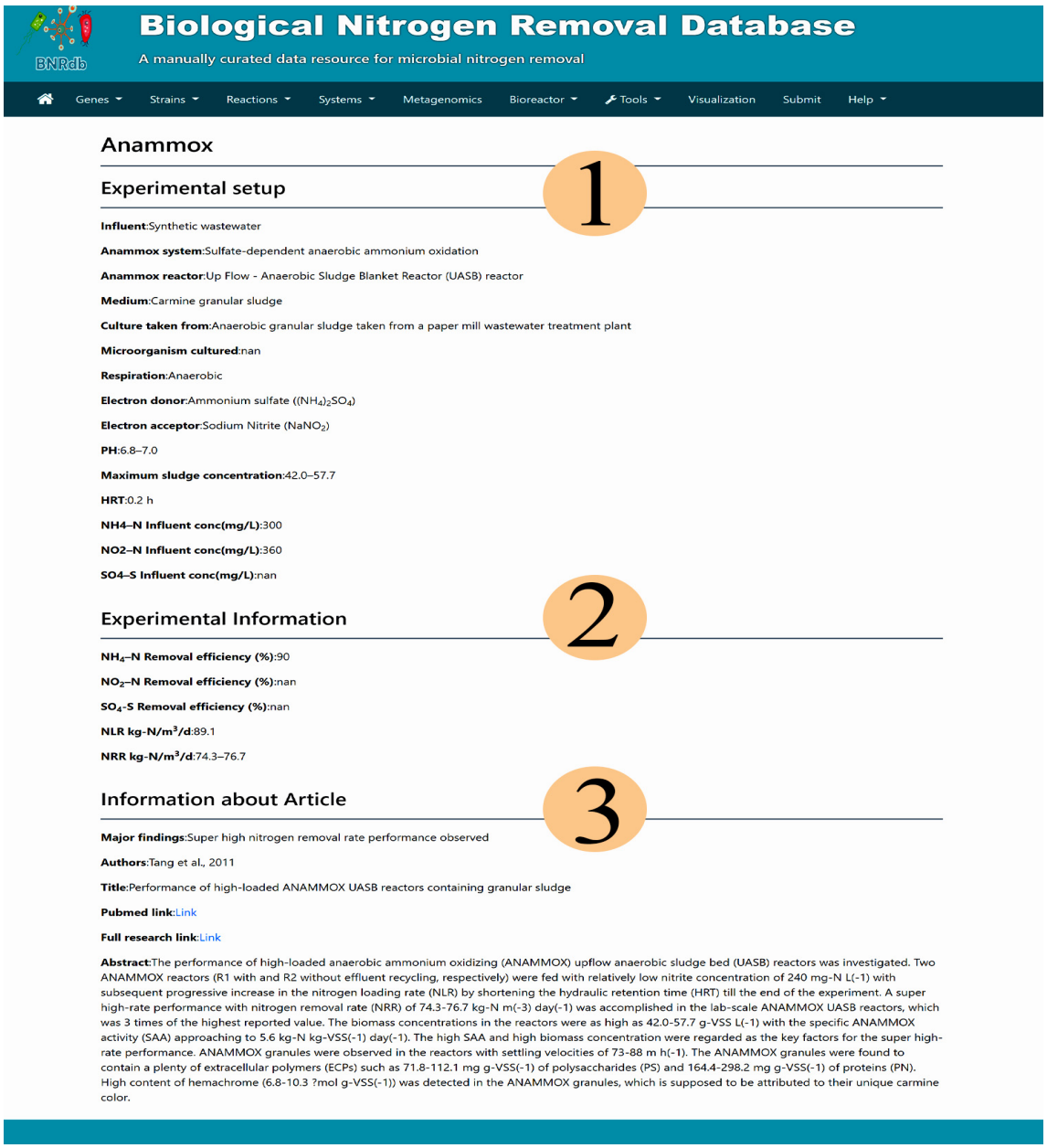

Figure S4. Screenshot of the BNR treatment systems experimental methods and conditions. (1) Experimental setup. (2) Experimental results in nitrogen removal from the system. (3) Literature information is provided with links to external databases.

Bioreactor Page

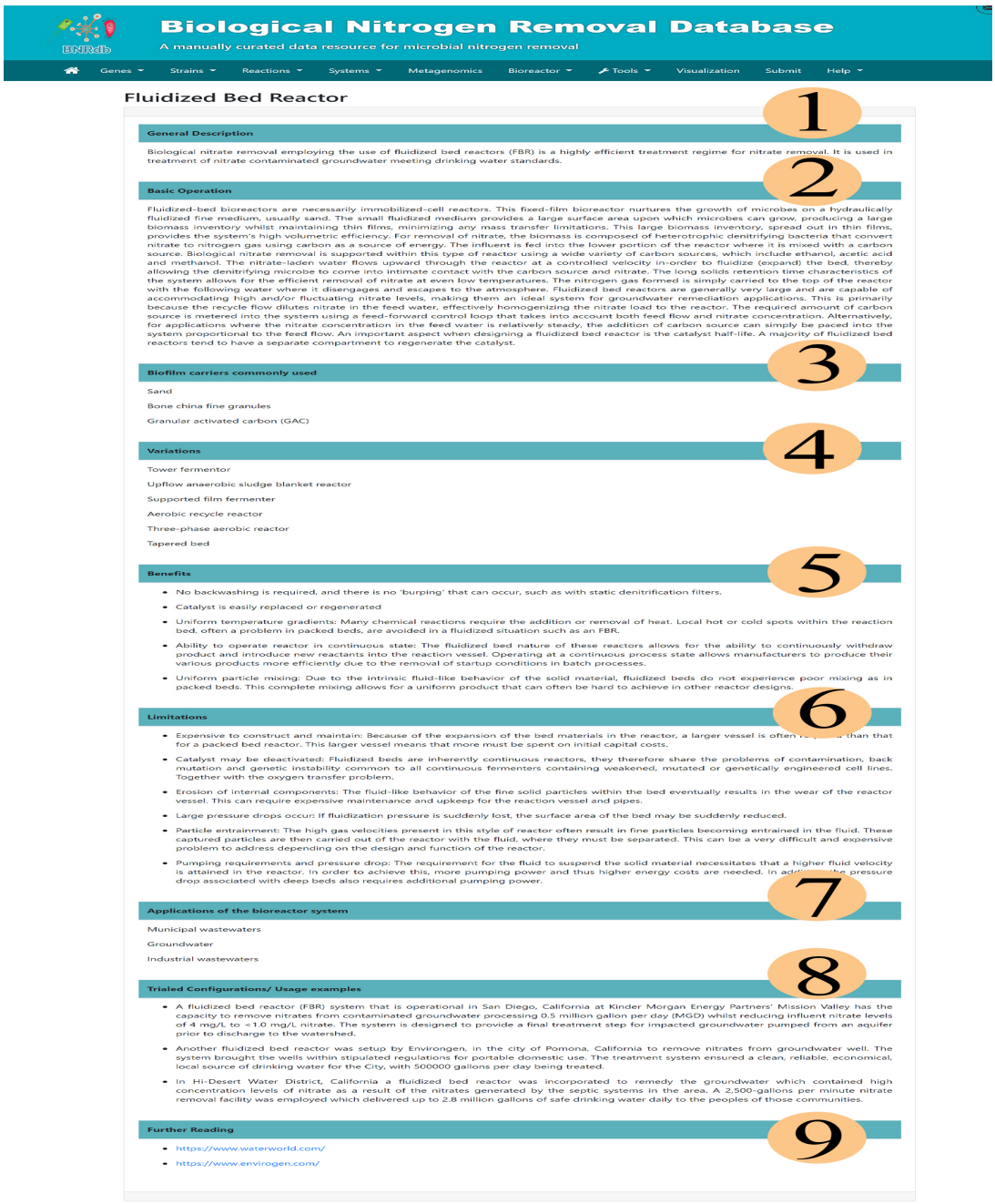

Figure S5. Screenshot display of the bioreactor display section using Fluidized Bed Reactor as an example. (1) General description. (2) Basic operation. (3) Carriers commonly used. (4) Variations. (5) Benefits of the reactor system (6), and Limitation associated with its use. (7) Possible applications. (8) Trailed configurations. (9) Further reading.

Submissions Page

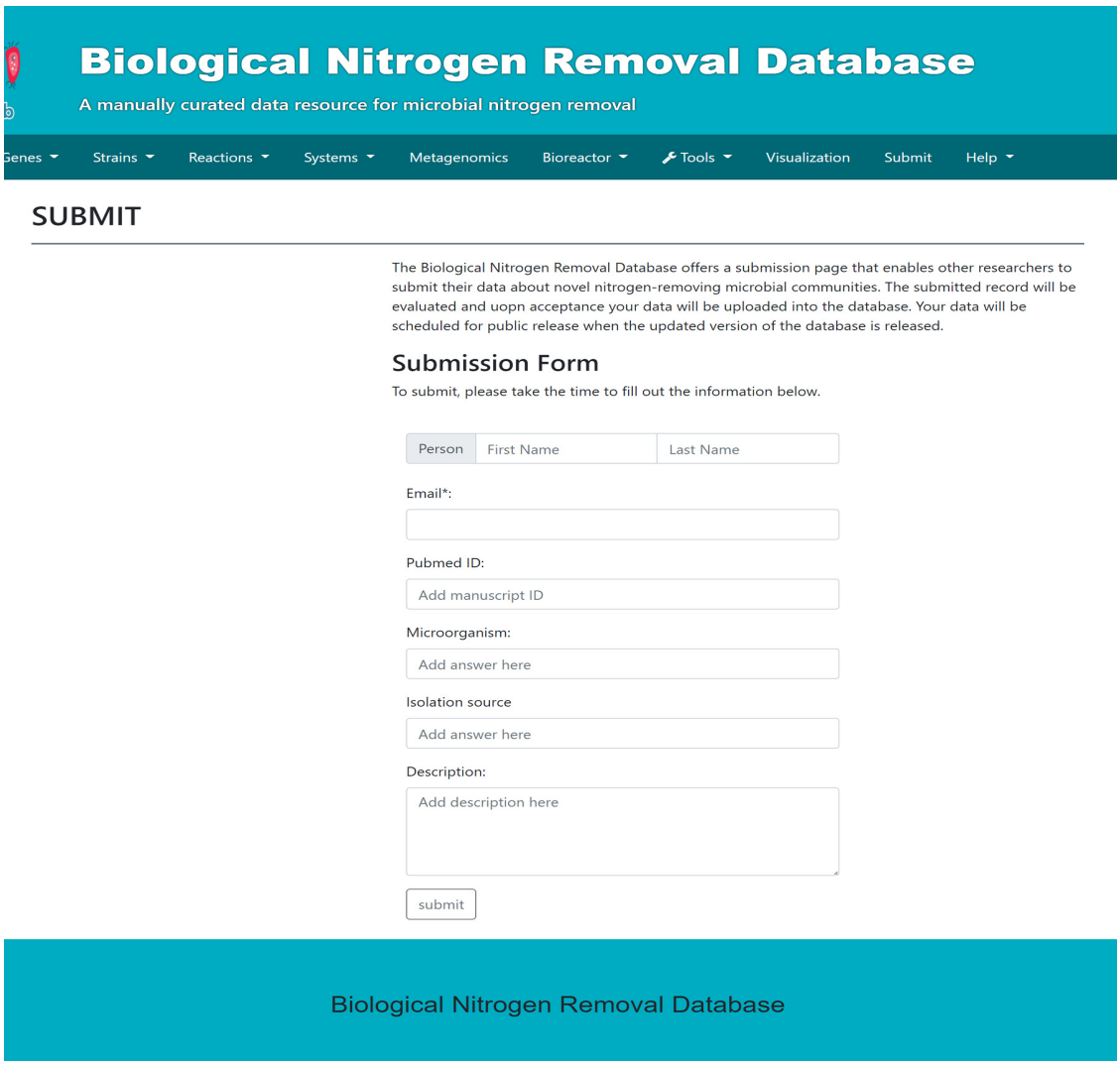

Figure S6. Screenshot of the submission page. An interactive data submission feature is incorporated into the BNRdb, to generate clean and uniform data sets that can be useful to other users. Contributors who wish to submit new data sets related to BNR, which is not already included in the database, can use this feature following the submission guidelines that have been devised. The administrator/curator may then contact the authors who submitted a new dataset to be the editor of that particular dataset and engage in reviewing any inaccurate or potentially missing data during the validation stage. The curator of the database will then upload the new dataset after its systematic validation.