water

MDPI

Article

Effects of Selenium in Different Valences on the Community Structure and Microbial Functions of Biofilms

Wenbin Hao [1], Chunmei Tao [2], Tanveer M. Adyel [3], Junjie Zhao[4,*], Jun Hou [5], Lingzhan Miao [5] and Yuan Zeng [6,*]

Supplementary Material

Text S1: High-throughput sequencing

The biofilms were sequenced with high throughput (Illumina HiSeq 2500 platform, Guangdong Magigene Biotechnology Co. Ltd., China) to explore the composition and function of microbial communities. Approximately 1 g of each biofilm sample were collected for bioinformatics analysis and DNA extraction using a DNA extraction kit (Omega Bio-tek, Norcross, GA, US). 16S rRNA genes of V4 region were amplified using primer (806R/515F) to identify the diversity of bacteria. DNA integrity and purity were monitored on 1% agarose gels and then measured using the NanoDrop One (Thermo Fisher Scientific, MA, USA) at the same time. Primers were synthesized by Invitrogen (Invitrogen, Carlsbad, CA, USA). PCR reactions, containing 25 μl 2x Premix Taq (Takara Biotechnology, Dalian Co. Ltd., China), 1 μl each primer(10 mM) and 3 μl DNA (20 ng/μl) template in a volume of 50 μl, were amplified by thermocycling: 5 min at 94°C for initialization; 30 cycles of 30 s denaturation at 94°C, 30 s annealing at 52°C, and 30 s extension at 72°C; followed by 10 min final elongation at 72°C. 3 replicates per sample and each PCR products of the same sample were mixed, the PCR instrument was BioRad S1000 (Bio-Rad Laboratory, CA, USA). Then, the PCR products were detected by 1% agarose gel electrophoresis and purified. Sequencing libraries were generated using NEBNext® Ultra™ DNA Library Prep Kit for Illumina® (New England Biolabs, MA, USA) following manufacturer's recommendations and index codes were added. The library quality was assessed on the Qubit@ 2.0 Fluorometer (Thermo Fisher Scientific, MA, USA) and Agilent Bioanalyzer 2100 system (Agilent Technologies, Waldbron, Germany). At last, the library was sequenced on an IlluminaHiseq2500 platform and 250 bp paired-end reads were generated.

Quality filtering on the paired-end raw reads were performed under specific filtering conditions to obtain the high-quality clean reads according to the Trimmomatic (V0.33) quality-controlled process. Then, paired-end clean reads were merged using FLASH (V1.2.11) according to the relationship of the overlap between the paired-end reads, when at least 10 of the reads overlap the read generated from the opposite end of the same DNA fragment, the maximum allowable error ratio of the overlap region of 0.1, and the spliced sequences were called Raw Tags. Then, sequences were assigned to each sample based on their unique barcode and primer using Mothur software (V1.35.1, http://www.mothur.org, accessed on 21 July 2021), after which the barcodes and primers were removed and got the effective Clean Tags.

Sequence's analysis were performed by usearch software (V8.0.1517). Sequences with ≥ 97% similarity were assigned to the same OTU. An OTU is thought to possibly represent a species. The most frequently occurring sequence was extracted as representative sequence for each OTU and was screened for further annotation. During the clustering, usearch can remove the chimera sequence and singleton OTU at the same time. For each representative sequence, the silva (for 16S) database was used to annotate taxonomic information (set the confidence threshold to default to ≥ 0.5). Removes the OTU and its Tags, which are annotated as chloroplasts or mitochondria (16S amplicons) and can't be annotated to the kingdom level. Then get the number of effective tags (No. of seqs) and OTU taxonomy synthesis information table (otu_table) for the final analysis. In order to study

the difference of the dominant species in different groups, the OTU representative sequence with the relative abundance in the first 50 and annotated to the level of the genus was selected, multiple sequence alignment was conducted using the FastTree software, and the relative abundance of each OTU and the species annotation information of the representative sequence were combined with the ggtree software package for visual display. After that, OTU abundance information were normalized using a standard of sequence number corresponding to samples. Subsequent analysis of alpha diversity and beta diversity were all performed basing on this output normalized data.

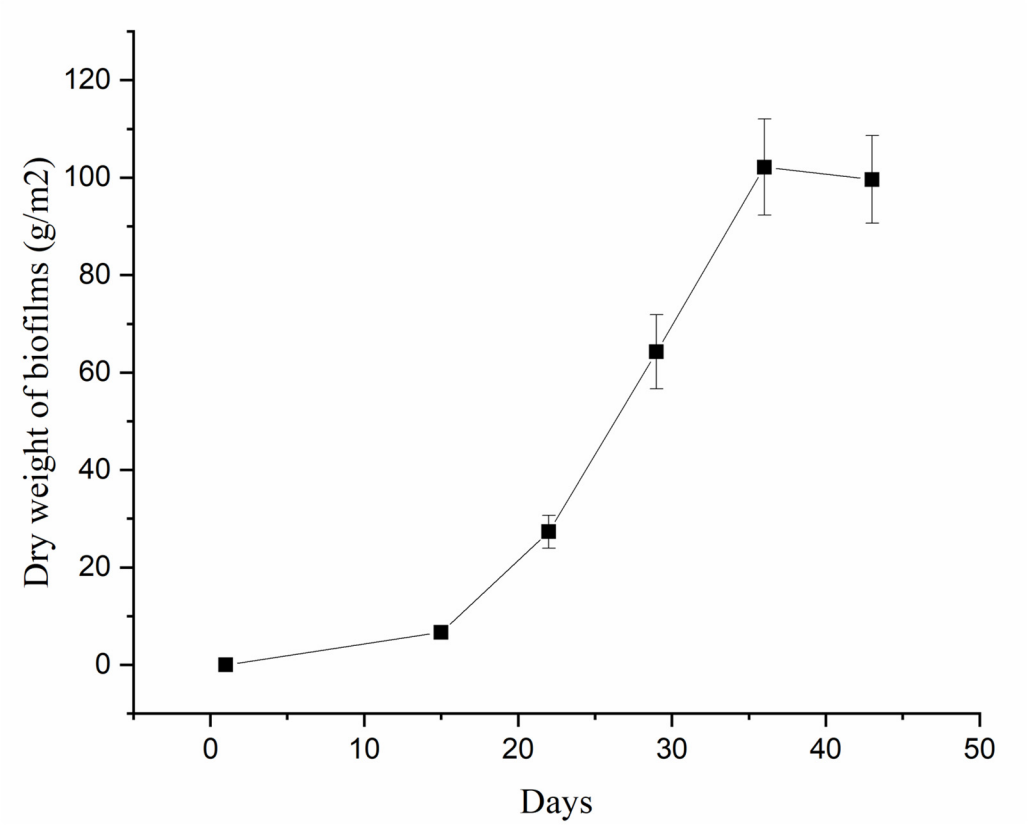


Figure S1. Dry weight of biofilms (g/m2) after colonization (Day 0). The dry weight became stable since Day 36. Mature and stabilized biofilms were obtained on Day 42 to conduct exposure experiments.

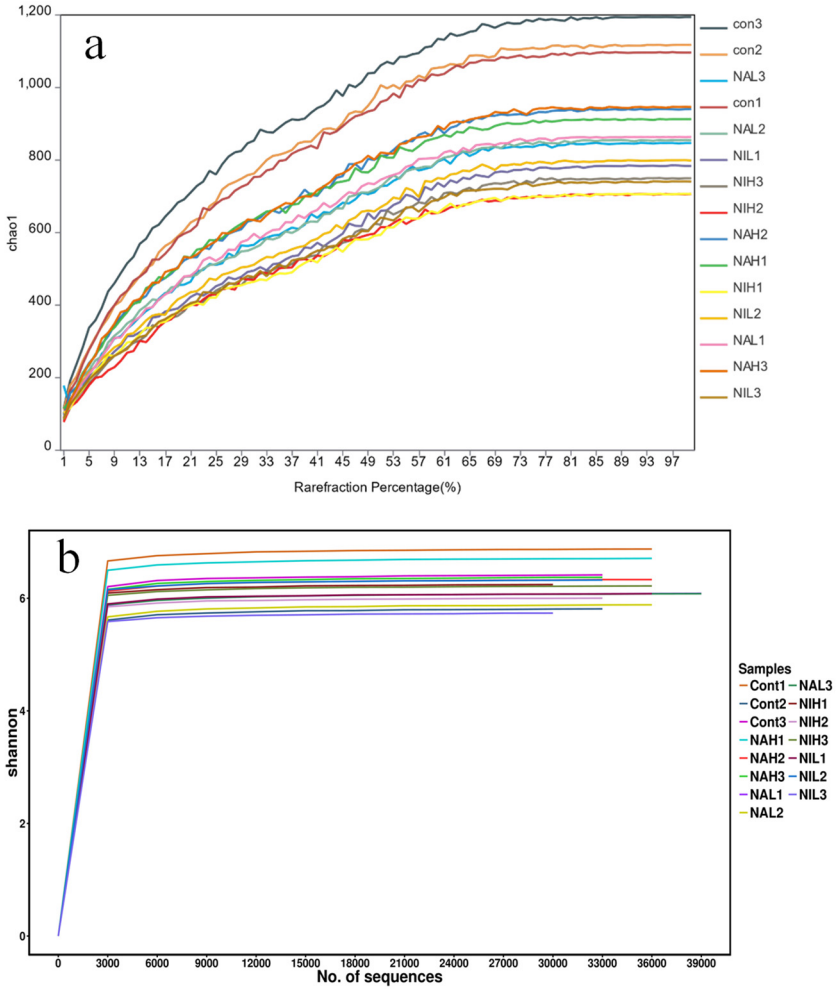


Figure S2. Dilution curve, the sequence number is in a stable stage of the curve, and the species coverage can meet the analysis requirements, (**a**) for Chao1; (**b**) for shannon.

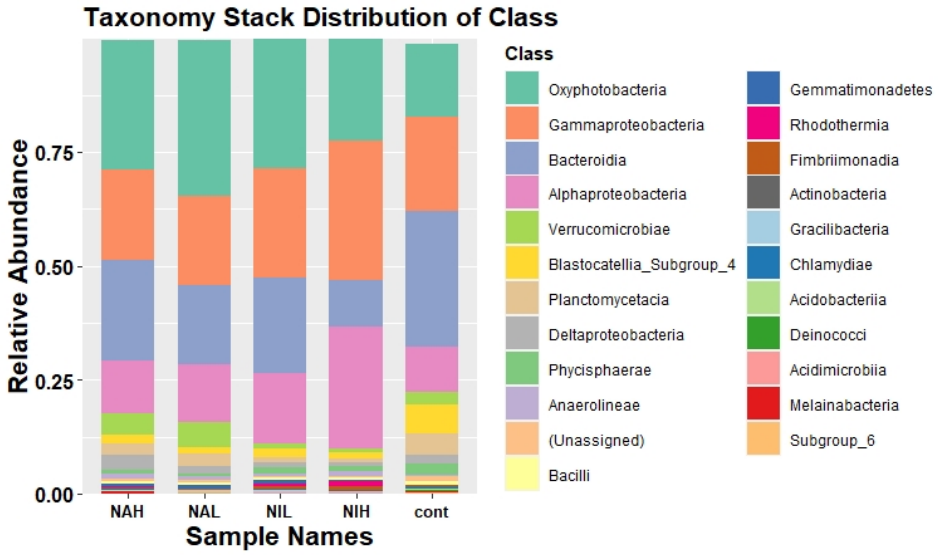


Figure S3. Community composition at class level, select species with relative abundance greater than 1%.

Table S1. Clusters of Orthologous Groups of proteins (COG) database at level 1.

Taxon	Cont_ Average	NAH_ Average	NAL_ Average	NIH_ Average	NIL_ Average	*p*-value
Metabolism	0.36294	0.36787	0.36558	0.37851	0.37222	0.01920
Cellular Processes And Signaling	0.25697	0.25628	0.26098	0.24964	0.25140	0.01920
Poorly Characterized	0.20649	0.20393	0.20346	0.20302	0.20453	0.09281
Information Storage And Processing	0.17360	0.17192	0.16999	0.16883	0.17185	0.01920

Table S2. COG database at level 2.

Taxon	Cont_ Average	NAH_ Average	NAL_ Average	NIH_ Average	NIL_ Average	*p*-value
[R] General function prediction only	0.12368	0.12099	0.12007	0.11721	0.11987	0.04083
[S] Function unknown	0.08281	0.08294	0.08339	0.08582	0.08467	0.04083
[E] Amino acid transport and metabolism	0.07292	0.07624	0.07592	0.07968	0.07743	0.04083
[K] Transcription	0.06725	0.07083	0.06921	0.06702	0.06741	0.04083
[M] Cell wall/membrane/envelope biogenesis	0.07011	0.06682	0.06538	0.06299	0.06692	0.04083
[T] Signal transduction mechanisms	0.06652	0.06699	0.06855	0.06479	0.06504	0.04552
[C] Energy production and conversion	0.05884	0.05845	0.05809	0.06149	0.05914	0.04083
[G] Carbohydrate transport and metabolism	0.05600	0.05811	0.05806	0.05211	0.05522	0.04083
[L] Replication, recombination and repair	0.05598	0.05158	0.05134	0.05058	0.05266	0.04083
[P] Inorganic ion transport and metabolism	0.05167	0.05054	0.05008	0.05240	0.05140	0.04083