Supplementary information

Materials and methods

Sample collection

Four Chinese crewmembers (3 men and 1 woman，two was the regular non-astronaut volunteers and two was real astronauts) entered the controlled ecological life support system (CELSS) for 180 days for integrated experiments which is significant for maintaining human health during Mars travel (figure 1). The CELSS was composed of 8 compartments with 4 categories: the crew capsule, biological capsule, life protection capsule and resource capsule, with an area of 370 square metres and a volume of 1340 cubic metres. The CELLS has efficient recycling of air, water and food in an enclosed environment, and the establishment of a life and health guarantee system suitable for long-term human residence was designed to significantly reduce the need for ground supplies. With the goal of achieving the enclosure of oxygen, water, food and solid waste cycles of 100%, 90%, 70% and 60%, respectively, the experiment was carried out with controlled ecological life and protection technology verification through several key technologies, including plant cultivation and management, waste treatment and reuse, water circulation management and atmospheric regeneration and regulation. In addition, a team of nutritionists systematically planned the diet of volunteers to ensure the stability of nutrition components before, in and after of the CELSS, the proportion of all ingredients in food remains constant. Some food came from scientific research personnel cultivated 5 classes containing 25 kinds of plants, including crops, vegetables, fruits and traditional Chinese medicines. Moreover, the exercise plans were also made by the relevant experts and strictly implemented according to physical ability, psychological condition and other factors.

The experiment was divided into five stages: before going into the space capsule (Before), pre-Mars landing (M1), Mars (M2), leaving Mars and returning to Earth (M3), landing on Earth and exiting the space capsule (After) (figure 1). On Mars, the day and night lighting followed the solar cycle of Mars (24 h 39 min 35.24409 s), so from 22:30 on the 71st day to 22:30 on the 108th day, the Mars solar day was established; that is,

the daily duration was changed to 24 hours 40 minutes (figure 1). Microbiological faeces and saliva samples and metabolic faeces, urine, and plasma samples were obtained every 15 days for 180 days (figure 1). Two hours after a meal, no stimulated saliva was obtained by rinsing the mouth 3 times with sterile double-distilled water resting for at least 10 minutes sitting upright before saliva collection and allowing saliva to naturally flow into an Eppendorf tube. Faecal samples were collected with sterile manure collectors. All salivary and faecal samples were stored at -20 °C during space travel and transferred to storage at -80 °C. The entire process was approved by the Ethics Committee of the China Astronaut Research and Training Centre (ACC). All crewmembers gave their written informed consent prior to their inclusion in the study.

Salivary and intestinal microbiota analysis and data generation

DNA was extracted from samples with the CTAB/SDS method. After extracting the genomic DNA of all the samples, PCR amplifications were carried out. The V3-V4 hypervariable regions of the 16S rRNA gene were amplified using specific primers (338F: ACTCCTACGGGAGGCAGCA, 806R: GGACTACHVGGGTWTCTAAT). All PCRs were carried out in 30-μL reactions with 15 μL of Phusion® High-Fidelity PCR Master Mix (Thermo Scientific, Waltham, MA, USA). PCR products were purified with a GeneJET Gel Extraction Kit (Thermo Scientific, Waltham, MA, USA) and analysed by electrophoresis on 2% agarose gels. Samples with strong, clear bands were chosen for further analysis. Finally, the library was sequenced on an Illumina HiSeq 2500 platform. Paired-end HiSeq 2500 sequencing reads were merged with sequence tags according to the overlap relationship between the reads and passed the QC test with UPARSE[1]. Poor-quality reads and chimaeras were filtered by USEARCH software[2]. After dereplication, clean reads were clustered into different operational taxonomy units (OTUs) with similarities higher than 97%. The representative reads of OTUs were aligned to the Ribosomal Database Project (RDP, release 11.5)[3] to obtain a clear taxonomy of the microbiome.

Ultra-performance liquid chromatography quadrupole coupled to Q Exactive Orbitrap mass spectrometry (UPLC-QE/MS) metabolomics profiling of faecal, plasma and urinary samples

The metabolic samples were mixed with methanol, and an internal standard (L-2-chlorophenylalanine, 1 mg/ml stock in dH_2O) was added. All samples were subjected to ultrasonic extraction in ice water for 10 min and centrifuged at 12,000 rpm at 4°C for 15 minutes, and the supernatant was harvested. Untargeted metabolite screening was performed on Q Exactive Orbitrap mass spectrometer (Thermo Scientific, Waltham, MA, USA) after calibration following the manufacturer's guidelines. Fullms $ddMS^2$ method was used to select QC samples, and multiple scanning segments were set to obtain more secondary data for identification. The samples were operated in both the positive and negative ion modes. Standardized samples were prepared for quality control.

Statistical analyses

The alpha index (PD_whole_tree) and distance matrix (Weighted UniFrac) were carried out by QIIME software (1.9.1). Other analyses and pictures were performed with Rstudio software (R version i386 3.6.2). Alpha diversity determination was applied to analyse the complexity of the species diversity in each sample. Differences in alpha diversity were tested by Wilcoxon's rank-sum tests. Principal coordinate analysis (PCoA) was used to examine the differences in the complexity of the samples by the vegan (2.5-6) and ggplot2 (3.3.0) packages. The differential analysis of microorganisms in phyla or species was performed by the Wilcoxon test (p value < 0.05), and heatmaps were generated by the pheatmap package (1.0.12). The Pearson correlation coefficient was calculated by psych (1.9.12.31), and the network was visualized by Cytoscape (version: 3.6.1) (absolute Pearson ≥ 0.6, p value < 0.05).

Mass spectrometry data were converted to the mzML format using MS converter software. Then, XCMS (version 1.50) software was used to process data, including a peak search and peak alignment. Finally, OSI-SMMS software (version 2.0.0, Dalian ChemDataSolution Information Technology Co. Ltd) was used for substance identification. The positive and negative ion data were first calculated as the relative abundance, and those with an ms2.score >0.5 were retained, combined and finally used for data analysis. The different metabolites were analysed by Wilcoxon test analysis with R studio software (version 3.6.2) (p < 0.05). Pathway enrichment analysis of the

differential metabolites was performed with the KEGG database (https://www.metaboanalyst.ca/home.xhtml). Pathways were considered significantly altered when $p < 0.05$. A network between different species and all or different metabolites was generated by R studio software and visualized by Cytoscape (version: 3.6.1) (correlation>|0.6|, p<0.05).

REFERENCES

1 Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 2013;10(10):996-8.

2 Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;26(19):2460-1.

3 Cole JR, Wang Q, Fish JA, et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 2014;42(Database issue):D633-42.