Genomic and transcriptomic analysis of NDM-1 *Klebsiella pneumoniae* in spaceflight reveal mechanisms underlying environmental adaptability

Jia Li1,4, Fei Liu2, Qi Wang2, Pupu Ge2, Patrick C. Y. Woo3, Jinghua Yan2, Yanlin Zhao5, George F. Gao2, Cui Hua Liu2 & Changting Liu1

1Nanlou Respiratory Diseases Department, Chinese PLA General Hospital, No. 28 Fuxing Road, Haidian District, Beijing 100853, China, 2CAS key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, No. 1 West Beichen Road, Chaoyang District, Beijing 100101, China, 3State Key Laboratory of Emerging Infectious Diseases, Department of Microbiology, The University of Hong Kong, University Pathology Building, Compound Pokfulam Road, Hong Kong, China, 4School of medicine, Nankai University, 94 Weijin Road, Nankai District, Tianjin 300071, China, 5National Center for Tuberculosis Control and Prevention, Chinese Center for Disease Control and Prevention, No. 155 Changbei Road, Changping District, Beijing 102206, China.

The emergence and rapid spread of New Delhi Metallo-beta-lactamase-1 (NDM-1)-producing *Klebsiella pneumoniae* strains has caused great concern worldwide. To better understand the mechanisms underlying environmental adaptation of those highly drug-resistant *K. pneumoniae* strains, we took advantage of the China’s Shenzhou 10 spacecraft mission to conduct comparative genomic and transcriptomic analysis of a NDM-1 *K. pneumoniae* strain (ATCC BAA-2146) being cultivated under different conditions. The samples were recovered from semisolid medium placed on the ground (D strain), in simulated space condition (M strain), or in Shenzhou 10 spacecraft (T strain) for analysis. Our data revealed multiple variations underlying pathogen adaptation into different environments in terms of changes in morphology, H2O2 tolerance and biofilm formation ability, genomic stability and regulation of metabolic pathways. Additionally, we found a few non-coding RNAs to be differentially regulated. The results are helpful for better understanding the adaptive mechanisms of drug-resistant bacterial pathogens.

*Klebsiella pneumoniae* is a highly important bacterial pathogen that causes a wide range of community- and hospital-acquired infections, such as pneumonia, urinary tract infections, intra-abdominal infections and respiratory tract infections1,2. The emergence and rapid spread of carbapenem-resistant *K. pneumoniae* isolates such as the New Delhi Metallo-beta-lactamase-1 (NDM-1)-producing strains has caused great concern worldwide3–5. The NDM-1, an Ambler class B metallo-β-lactamase (MBL), is capable of hydrolyzing all β-lactams (including carbapenems) except monobactams. It was first identified in *K. pneumoniae* and *Escherichia coli* in India in 2008. To date, NDM-1 producing pathogens have been reported in more than 40 countries6–8. The successful persistence and transmission of this “superbug” remain largely unknown, but could be partly due to the ability of this pathogen to sense and react to environmental and host stress signals, which allow it to persist and disseminate in harsh conditions such as in various medical settings and inside the human host. Thus, understanding the ability of those pathogens to adapt to various stressors they encountered during environmental persistence could facilitate the comprehension of their pathobiology. In addition, identification of specific genomic variations and transcriptome patterns critical for drug resistance and environmental adaptation of the drug-resistant pathogenic bacteria is important for the development of more effective pathogen control strategies.

During spaceflight, a variety of physiological stressors associated with the space environment and spacecraft conditions could potentially contribute to detrimental alterations in the human immune system. At the same time, the bacteria introduced into the extraterrestrial environments by the space crew members inhabiting the stressful aerospace environments might develop resistance traits which could be important to spacelighit missions and the general public medicine. Indeed, the recurrent isolation of various extremotolerant bacteria from space-
microscopy. Field emission scanning electron microscopy (SEM) was further used to monitor single cell morphology of the strains and the results showed that the bacterial cell walls of all strains were intact. But interestingly, some of the cells of the T strain turned into elongated forms and adhered to each other under SEM (Fig. 1). In the H₂O₂ sensitivity assay, the 1 mM and 5 mM H₂O₂ exposure resulted in slightly reduced survival for the T strain in comparison with the BAA-2146 control strain (P < 0.001), whereas, the D and M strains showed no obvious differences in sensitivity to H₂O₂ when compared to the BAA-2146 strain (Fig. 2a). The ability of K. pneumoniae to form biofilms is thought to be an important phenotype with respect to adaptability and virulence traits such as host colonization, antibiotic resistance, and environmental persistence, etc. We therefore performed biofilm assays to determine the biofilm formation abilities of the strains, and the results showed that the T strain showed enhanced biofilm-forming capacity compared to the BAA-2146 strain (P < 0.001) while the D strain and the M strain showed similar biofilm-forming capacity with the reference strain BAA-2146 (Fig. 2b). The fitness of the K. pneumoniae strains was further assessed by determining the growth rates of the strains in a nutrient LB medium as well as in chemically defined media (CDM), as this parameter also could contribute to environmental adaptability. All strains showed similar growth curves in LB medium (Fig. 3a) and in CDM (Fig. 3b). Furthermore, the growth curves were also similar when the bacteria were cultivated in LB medium at different temperatures including 30°C, 37°C and 40°C (Fig. 3c, 3a and 3d). To better understand the phenotypic changes in the strains, we also performed carbon source utilization assays and chemical sensitivity assays using the 96-well Biolog GEN III MicroPlate. Among the 71 carbon source utilization assays, the α-D-Lactose utilization ability was found to be defective in all three strains (including the D, M and T strains) as compared to the reference strain BAA-2146. Interestingly, the M strain was shown to gain the ability to use D-Mannose. Among the 23 chemical sensitivity assays, no obvious changes were observed for all three strains compared to the reference strain (Supplementary Table S2).

Whole genome sequencing statistics. The basic whole genome sequencing statistics were shown in Table 1. The sequencing depth ranged between 127× and 128× and read mapping results reported nearly complete genome coverage (~99.99%) for all three strains. Fig. 4 showed the circular representation of the genome features including the COG annotated coding sequences, KEGG enzymes, RNA genes, GC content, GC skew, strain-specific SNPs, etc. As shown in Supplementary Fig. S1, all three samples had similar mapping coverage across the reference genome and the coverage was higher for the regions closer to the origin of replication.
Figure 2 | Oxidative stress tolerance and biofilm formation ability of *K. pneumoniae* strains. (a) Survival of *K. pneumoniae* strains in oxidative stressed environment. The ATCC BAA-2146, D, M and T strains were incubated with different concentrations of H$_2$O$_2$ in sodium phosphate buffer, and then tested for their survival on LB plates. Growth is expressed as relative survival after 2 h incubation with H$_2$O$_2$. (b) Biofilm assays of *K. pneumoniae* strains. OD$_{595}$ readings were measured for each strain to determine the amount of biofilm formed. Data represent one of three independent experiments.

Figure 3 | Growth curves of the *K. pneumoniae* strains in LB medium (a, c and d) or CDM (b). Overnight-grown bacteria were diluted 100 folds in fresh LB medium or CDM and were cultivated at designated temperature with shaking (180 rpm). OD$_{600}$ readings were measured at designated time points for each strain. The data represent one of two independent experiments performed in triplicate, with s.d. indicated by error bars.
probably because of a higher speed of the DNA replication than that of cellular doubling during exponential growth.

Identification of the mutations in D, M and T strains compared to reference strain ATCC BAA-2146. A total of 34 mutations (excluding the synonymous mutations) were identified in all strains (Supplementary Tables S3), among which, 25 mutations were shared by all three strains (Supplementary Tables S4). No strain-specific mutation was identified in the D strain, while 3 and 2 strain-specific mutations were identified in the T and M strains, respectively. In addition, a few mutations were shared by T and M, T and D, or M and D strains (Table 2). PCR and sequencing analysis confirmed that these mutations were not sequencing or assembly errors.

Drug resistance determinants of K. pneumoniae strains. Electronic PCR was conducted for ATCC BAA-2146, D, M and T strains to analyze drug resistance determinants conferring resistance to carbapenems, folate pathway inhibitors, fluoroquinolones, aminoglycosides, etc. The drug resistance-associated genes and gene mutations detected in the D, M and T strains were identical to those of the reference strain ATCC BAA-2146 (Supplementary Table S5).

RNA-Seq mapping statistics. Approximately 96% of the sequencing reads could be mapped to ATCC BAA-2146 (CP006659) reference genome (Table 3). The uniquely mapped reads for D, M, and T were 96.0%, 95.9% and 96.3% respectively. The number of reads mapped each gene ranged from 1 to 449,669 with a median of 206, 209 and 214 for D, M and T, respectively. In addition, the RPKM for D, M and T was 48.8, 49.5 and 52.3 for D, M and T respectively. The distribution of the number of mapped reads and RPKM values across all three samples was displayed in Supplementary Fig. S2. The expression of total genes and the differentially expressed genes identified among D, M and T strains (fold changes >2) were shown in Fig. 5.

Comparative transcriptomic analysis. By using the KEGG orthology based annotation system to identify metabolic pathways, we identified extensive changes in the transcriptomes of the M and T strains in comparison to the D strain. Expression of the differentially expressed genes identified among D, M and T strains (including: T strain vs. D strain, M strain vs. D strain, and T strain vs. M strain) were shown in Supplementary Tables S6–S8. Compared to the D strain transcriptome, the T strain transcriptome was characterized by regulation of a number of genes involved in fatty acid degradation ($P=0.0464189$), microbial metabolism in diverse environments ($P=0.0464189$), and phenylalanine metabolism ($P=0.0464189$). Among those three categories, the majority of the genes involved in fatty acid degradation and phenylalanine metabolism were down-regulated, while the genes involved in microbial metabolism

| Table 1 | Statistics of whole genome sequencing |
|---|---|---|
| Sample | D strain | M strain | T strain |
| **Raw read statistics** | | | |
| Total reads | 4011112 | 4038889 | 4038889 |
| Total base pairs | 727 | 727 | 727 |
| Sequencing depth (X) | 127 | 128 | 128 |
| Genome coverage (%) | 99.99 | 99.99 | 99.99 |
| **Assembly statistics** | | | |
| Chromosome size | 5,674,681 | 5,671,078 | 5,671,920 |
| No. of scaffolds | 72 | 73 | 66 |
| Largest scaffold length | 840,095 | 840,520 | 839,882 |
| N50 scaffold length | 236,010 | 236,344 | 288,168 |
| G+C content (%) | 57.0 | 57.0 | 57.0 |

Figure 4 | Circular representation of the genome features. Genome sequences (ring 1), COG Annotated coding sequences (rings 2 + 3), KEGG enzyme (ring 4), RNA genes (ring 5: red, rRNA; blue, tRNA), GC content (rings 6), GC skew (ring 7), strain-specific SNPs (ring 8) are shown. All strain-specific SNPs from D (yellow), M (red), and T (blue) are shown on ring 8. Very short features were enlarged to enhance visibility. Clustered genes and SNPs, such as several rRNA genes, may appear as one line due to space limitations. The image was created by using the software Circos.
in diverse environments include many up-regulated and many down-regulated genes (Supplementary Table S6). Compared to the D strain transcriptome, the M strain transcriptome showed that many genes associated with tyrosine metabolism (P<0.001), degradation of aromatic compounds (P=0.003804639), and microbial metabolism in diverse environments (P=0.005292205) were regulated (Supplementary Table S7), and the majority of those genes were up-regulated. Genes altered in their expression common to both M strain and T strain transcriptomes included a variety of genes involved in microbial metabolism in diverse environments. In comparison to the M strain, the T strain transcriptome was further characterized by regulation of a number of genes involved in citrate cycle (P=0.0179529) and phenylalanine metabolism (P=0.0404427). In addition, many more genes involved in microbial metabolism in diverse environments (P=0.0404427) were regulated in the T strain than in the M strain. Additionally, we found a few ncRNAs to be differentially regulated in the T and M strains as compared to the D strain. Some of those ncRNAs (including GlmZ RNA activator of glmS mRNA and RT RNA) were similarly regulated in both M and T strains. Some of those ncRNAs (including GlmZ RNA activator of glmS mRNA and RT RNA) were similarly regulated in both M and T strains. (M/D) belonged to the following three COG categories: Cell motility, amino acid transport and metabolism (P=0.04659). In addition, many more genes involved in microbial metabolism in diverse environments (P=0.004304), transport and catabolism (P=0.007086), and Inorganic ion transport and metabolism (P=0.0346).

Table 2 | Comparative analysis of SNPs and Indels identified in D, M, and T strains. Synonymous SNPs were excluded

| Mutation position | Gene name | Product | Base mutation | AA mutation | Type of mutation |
|-------------------|-----------|---------|---------------|-------------|-----------------|
| **SNPs and Indels specific to T strain** |
| G1539668C | Kpn2146_1513 | D-3-phosphoglycerate dehydrogenase | G20C | R7P | Non-synonymous |
| T2546875G | Kpn2146_2552/Kpn2146_2553 | 2-carboxyglycerate decarboxylase/Hypothetical protein | T-327G | | Intergenic regions |
| G4946124A | Kpn2146_5015 | Small subunit ribosomal RNA | G1400A | | |
| **SNPs and Indels specific to M strain** |
| T3252757A | Kpn2146_3273/Kpn2146_3274 | RiT RNA/riRNA-Tyr(GTA) | T-39A | | Intergenic regions |
| C3252759T | Kpn2146_3273/Kpn2146_3274 | RiT RNA/riRNA-Tyr(GTA) | C-37T | | Intergenic regions |
| **SNPs and Indels specific to D strain** |
| None |
| **Non-synonymous SNPs and Indels shared by T strain and M strain and not present in D strain** |
| G2613880A | Kpn2146_2620/Kpn2146_2621 | Ambler Class A beta lactamase | G274A | | Intergenic regions |
| **SNPs and Indels shared by T strain and D strain and not present in M strain** |
| C5323238G | Kpn2146_5387/Kpn2146_5388 | Transcriptional regulator RuR of Tn10 family | C289G | | Intergenic regions |
| **SNPs and Indels shared by M strain and D strain and not present in T strain** |
| G2610928T | Kpn2146_2617 | Hypothetical protein | G112T | L38I | Non-synonymous |
| G2612995C | Kpn2146_2619/Kpn2146_2620 | DeoR family transcriptional regulator probably involved in glycerate glycolaldehyde metabolism/Ambler Class A beta lactamase SHV-11 | G-1C | | Intergenic regions |

Table 3 | Summary of RNA sequencing data

| Sample name | D reads number | Percentage | M reads number | Percentage | T reads number | Percentage |
|-------------|----------------|------------|---------------|------------|---------------|------------|
| Total reads | 13,367,754 | 100.0% | 13,887,820 | 100.0% | 13,284,972 | 100.0% |
| Total base pairs | 1,203,097,860 | 100.0% | 1,249,903,800 | 100.0% | 1,195,647,480 | 100.0% |
| Total mapped reads | 12,831,916 | 96.0% | 13,321,176 | 95.9% | 12,796,903 | 96.3% |
| Perfect match | 10,367,952 | 77.6% | 10,822,875 | 77.9% | 10,284,372 | 77.4% |
| ≤3bp mismatch | 2,463,964 | 18.4% | 2,498,301 | 18.0% | 2,512,531 | 18.9% |
| Unique match | 12,829,558 | 96.0% | 13,317,899 | 95.9% | 12,793,482 | 96.3% |
| Multi-position match | 2,358 | 0.0% | 3,277 | 0.0% | 3,421 | 0.0% |
| Total unmapped reads | 535,838 | 4.0% | 566,644 | 4.1% | 488,069 | 3.7% |
| Reads that aligned to rRNA | 10,366 | 0.2% | 11,075 | 0.2% | 4,232 | 0.1% |
Discussion

The explanation for the epidemiological success of highly drug-resistant bacterial pathogens, such as the NDM-1-producing *K. pneumoniae* strains, is complex. In general, the antibiotic resistance phenotype of bacteria is attributed to their capacity to acquire and express genes associated with a wide range of antibiotic resistance functions, as well as their intrinsic properties such as porins and gene mutations, etc. In this study, the reference strain ATCC BAA-2146 was shown to possess a large number of previously reported genes associated with resistance to β-lactams, fluoroquinolones, aminoglycosides and folate pathway inhibitors. Among the more recently reported carbapenemases genes (such as *bla*<sub>IMP</sub>, *bla*<sub>VIM</sub>, *bla*<sub>NDM-1</sub>, *bla*<sub>KPC</sub> and *bla*<sub>OXA-48</sub>, etc.), only NDM-1 was detected in ATCC BAA-2146. We also identified a *gyrA* mutation T247A (Ser83Ile) which was shown to be associated with fluoroquinolone resistance.

In this study, the D, M and T strains all remained their drug resistance phenotype and maintained all those drug resistance-associated genes detected in the reference strain ATCC BAA-2146, suggesting that this highly drug-resistant *K. pneumoniae* strain did not lose resistance to those drugs while grown in extreme conditions. Based on MLST genotyping analysis, ATCC BAA-2146 belonged to ST11, which was reported to be the most dominant clone among the carbapenem-resistant *K. pneumoniae* isolates in China. Thus this drug-resistant *K. pneumoniae* strain could be intrinsically highly transmissible.

Morphological analysis revealed that some cells of the T strain were elongated and adhered to each other under SEM, suggesting that certain genes involved in maintaining bacterial morphology and adherence ability could be affected or regulated. Normally, the function and morphology of the bacteria are maintained by well-coordinated equilibrium of their fusion and fission activities. Therefore, the explanations for the mechanism underlying the formation of elongated bacteria could be an enhanced fusion process, a blocked fission process, or a combined action of the two. Since we did not identify mutations in genes directly involved in bacterial fission or fusion function in the T strain, we thus hypothesize that certain mutations which occurred in the intergenic regions or ncRNAs could play regulatory roles on genes associated with these morphological changes in the T strain. We also predict that the increased bacterial fusion and adherence capabilities might be beneficial for the pathogen to form biofilm under stress conditions in order to promote bacterial survival.

The elongated forms of the bacteria have previously been observed in bacteria being treated by certain antibiotics. For example, in one study, the scanning electron microscopy analysis demonstrated that *Pseudomonas aeruginosa* became elongated after exposure to ciprofloxacin. Another study also showed that most of the *P. aeruginosa* cells turned into elongated forms and adhere to each other when exposed to nalidixic acid. Thus, this morphological change could be a typical response of pathogenic bacteria towards various stressors.

Bacterial cells are constantly challenged by various environmental stressors from their natural habitats. Similar to many pathogens, *K. pneumoniae* faces several challenges during infection and colonization of the human body. Among these, the ability of the pathogen to tolerate oxidative stress is critical for their survival since the transition from aerobic to microaerophilic conditions or the transition from a microaerophilic to oxidative stress environment is frequently encountered by the pathogen during infection. Oxidative stress on aerobic bacteria is mainly mediated by partially reduced oxygen species, or reactive oxygen species, most notably superoxide and H<sub>2</sub>O<sub>2</sub>, which are by-products of aerobic metabolism. These reactive oxygen species can cause damage to DNA, proteins and membranes. As a result, all aerobic bacteria possess various mechanisms to scav-
enough superoxide and $H_2O_2$, thus protecting the cells from being damaged by these reactive oxygen species. Thus, the response of drug-resistant bacteria under oxidative stress conditions is of particular interest. Our study revealed that the T strain exhibited a slightly reduced tolerance toward $H_2O_2$, which could be partly explained by the decreased expression of the universal stress response gene as well as a few ncRNAs such as $sroC$ RNA, $6S/SsrS$ RNA, and the bacterial small signal recognition particle RNA in this strain, since those genes and ncRNAs were shown to be involved in stress tolerance or bacterial fitness\textsuperscript{23–27}. Interestingly, it was demonstrated that $S. Typhimurium$ inside macrophages led to repression of the expression of the $sroC$ and $6S/SsrS$ RNA genes, suggesting that those ncRNAs may also direct pathogen adaptation to a non-proliferative state inside the host cell\textsuperscript{30–34}. Taken together, we hypothesize that those ncRNAs could also play critical roles during environmental adaptation of $K. pneumoniae$.

The bacterial biofilm consists of an aggregate of cells contained within a matrix of surface polysaccharides, proteins and DNA. The ability to produce a biofilm results in enhanced resistance to host defense factors, antimicrobials, as well as various environmental stressors, thus it has been increasingly recognized as an important virulence property. An observation by Wu et al. showed that certain pathogens produced more biofilm than 

| Table 4 | Non-coding RNAs differentially expressed in different $K. pneumoniae$ strains |
|----------|-----------------|-----------------|-----------------|-----------------|
| up/down  | LogFC           | Gene name       | Length          | Product                     |
| ncRNAs in T strain vs. D strain |                  |                 |                 |                             |
| up       | 1.115428        | Kpn2146_0162    | 212             | GlmZ RNA activator of glmS mRNA |
| up       | 1.34515         | Kpn2146_1637    | 166             | RIT RNA                     |
| up       | 1.527822        | Kpn2146_1634    | 170             | RIT RNA                     |
| up       | 1.722864        | Kpn2146_4340    | 157             | sok antitoxin               |
| up       | 2.218488        | Kpn2146_1640    | 169             | RIT RNA                     |
| up       | 2.243956        | Kpn2146_1642    | 169             | RIT RNA                     |
| down     | -2.14516        | Kpn2146_1562    | 162             | sroC RNA                    |
| ncRNAs in M strain vs. D strain |                  |                 |                 |                             |
| up       | 1.066489        | Kpn2146_1199    | 97              | Bacterial small signal recognition particle RNA |
| up       | 1.091426        | Kpn2146_0336    | 148             | srrL-Hfq binding RNA        |
| up       | 1.136683        | Kpn2146_0162    | 212             | GlmZ RNA activator of glmS mRNA |
| up       | 1.182549        | Kpn2146_1634    | 170             | RIT RNA                     |
| up       | 1.227269        | Kpn2146_1562    | 162             | sroC RNA                    |
| up       | 1.289375        | Kpn2146_3220    | 108             | RprA RNA                    |
| up       | 1.29224         | Kpn2146_0026    | 119             | Spot 42 RNA                 |
| up       | 1.305698        | Kpn2146_1642    | 169             | RIT RNA                     |
| up       | 1.442965        | Kpn2146_1640    | 169             | RIT RNA                     |
| up       | 1.508266        | Kpn2146_1637    | 166             | RIT RNA                     |
| up       | 2.056248        | Kpn2146_4589    | 184             | 6S/SsrS RNA                 |
| ncRNAs in T strain vs. M strain |                  |                 |                 |                             |
| up       | 1.089629        | Kpn2146_2396    | 1910            | group II intron S.ma.II     |
| up       | 1.117141        | Kpn2146_2146    | 1910            | group II intron S.ma.II     |
| up       | 1.370092        | Kpn2146_4340    | 157             | sok antitoxin               |
| down     | -1.02646        | Kpn2146_3275    | 132             | RIT RNA                     |
| down     | -1.26951        | Kpn2146_2450    | 118             | Fumarate/nitrate reductase regulator sRNA |
| down     | -1.37514        | Kpn2146_4084    | 148             | Glm Y RNA activator of glmS mRNA |
| down     | -1.38085        | Kpn2146_0026    | 119             | Spot 42 RNA                 |
| down     | -1.56665        | Kpn2146_4589    | 184             | 6S/SsrS RNA                 |
| down     | -1.73232        | Kpn2146_1199    | 97              | Bacterial small signal recognition particle RNA |
| down     | -1.90215        | Kpn2146_0595    | 87              | C4 antisense RNA            |
| down     | -2.2773         | Kpn2146_5143    | 66              | RyhB RNA                    |
Kpn2146_3273 (RtT RNA) and Kpn2146_3274 (tRNA-Tyr) (T-39A and C-37T). Those mutations within or nearby those transcriptional regulators and non-coding RNA regions could be associated with the regulation of metabolic pathways through up-regulating or down-regulating their targets in the *K. pneumoniae* strains during pathogen environmental adaptation.

The study of genomic variations as a function of pathogen adaptation to a specific niche could reveal important insights into how they sense and respond to varied habitats. Our comparative genomic analysis revealed only a few mutations specific to each strain, but a large number of mutations were shared by the D, M and T strains, suggesting the all three strains underwent adaptation genomic variations while being cultivated in different environments, though each strain were exposed to different combination of stressors and the strength of each stressors they encounter might be different. We also noticed that a large proportion of the mutations occurred in the intergenic regions or ncRNAs, suggesting that those mutations could play important regulatory roles on their down-stream genes or more distantly located genes, which could be involved in stress response and environmental adaptation. There were also some mutations occurring in hypothetical proteins such as Kpn2146-1328, Kpn2146-2411, and Kpn2146-2412, the roles of which in pathogen environmental adaption await further studies.

Comparative transcriptomic analysis revealed that the gene expression patterns of the three strains were significantly different from each other. The transcriptomes of both the M and T strains indicated significant changes in microbial metabolism compared to the D strain. Many more genes were regulated in the T strain compared to the M strain, suggesting that the T strain experienced more stress during spaceflight than the M strain which was cultivated in a simulated space condition with $10^{-3} \text{g}$ microgravity. Furthermore, a large proportion of the differentially expressed genes in the T strain, as compared to both D and M strains, were involved in a variety of different metabolic pathways including the following COG categories: amino acid transport and metabolism, and carbohydrate transport and metabolism. These results suggest that in response to environmental stress, pathogenic bacteria including highly drug-resistant ones exhibit a great flexibility and adaptability to survive successfully through regulation of multiple physiological functions and cellular pathways, among which the regulation of metabolism pathways seems to play a very important role. Using RNA-Seq, we also identified a few previously reported or putative ncRNAs, which could play critical roles during the environmental adaptation of *K. pneumoniae*. As the RNA isolation procedure used in this study selected against small RNA molecules, it is likely that additional small ncRNAs not detected here could also be transcribed during stress responses of the

Figure 6 | Distribution of differentially expressed genes in COG functional categories. The y-axis represents the number of genes in each COG category.
**K. pneumoniae.** Nevertheless, the results from this study provide powerful evidence showing that RNA-Seq allow quantitative characterization of bacterial transcriptomes and provide a useful tool for exploring transcriptional regulatory networks in bacteria.

In summary, our comparative genomic and transcriptomic analysis of a NDM-1 *K. pneumoniae* strain being cultivated in different conditions support the possibility resistance or adaptation of the highly drug-resistant pathogenic bacteria towards the extreme environments such as the conditions in spacecrafts, which could have important impacts on the microbial ecology of the extraterrestrial space. The present study could also serve as a basis for future studies examining the complex network systems that regulate bacterial adaptation to extreme environments. Furthermore, the insights we gained from the observations of persistence and pathogenesis-related genomic variations and transcriptomic changes of the NDM-1 *K. pneumoniae* strain in response to spaceflight could also facilitate the development of novel or alternative therapeutic methods needed to treat the recalcitrant infections caused by those highly drug-resistant *K. pneumoniae* strains. Finally, the data obtained from this study could also be important for future infectious disease risk assessment and prevention during spaceflight missions and in general public health as well.

**Methods**

**Bacterial strains and culture conditions.** The reference *K. pneumoniae* strain KP ATCC BAA-2146 was used in the study. *K. pneumoniae* was cultivated in semisolid medium (with 0.5% agar) for 15 days at 21°C, and the culture were placed in the following conditions: on the ground (D strain), in simulated space condition with rotation at 30 rpm and 10⁻¹ g microgravity (M strain), and in the Shenzhou 10 spacecraft (T strain). For phenotypic analysis, the *K. pneumoniae* strains were routinely grown in Luria-Bertani (LB) broth. Agar was added to a final concentration of 1.5% when necessary. For growth curve measurement, the strains were grown at designated temperature (30°C, 37°C or 40°C) with shaking (180 rpm). The CDM used for growth curve measurement was described previously.

**Drug susceptibility testing.** Drug susceptibility testing (DST) for the *K. pneumoniae* strains was performed using the bioMérieux VITEK II-AST-GN13 system following manufacturer’s instructions as described previously. The following 18 drugs were tested: ampicillin (AMP), piperacillin/tazobactam (TZP), ampicillin/sulbactam (SAM), cefazolin (CFZ), ceftriaxone (CRO), cefadiazime (CAZ), ceftipime (FEP), cefotetan (CTT), etepramet (ETM), imipenem (IMP), aztreonam (ATM), ciprofloxacin (CIP), levofloxacin (LVX), gentamicin (GM), tobramycin (TOB), amikacin (AMK), trimethoprim-sulfamethoxazole (SXT), furadantin (FD). The ESBLs were detected by the bioMérieux VITEK 2 AST-GN13 test. E. coli strains ATCC 25922 and ATCC 35218, *K. pneumoniae* strain ATCC 700603 and *P. aeruginosa* strain ATCC 27853 were used as quality control strains for the DST.

**H₂O₂ sensitivity assays.** To measure the susceptibility of *K. pneumoniae* to oxidative stress, H₂O₂ sensitivity assays were conducted according to the method reported by Cumley et al. with minor modifications. Briefly, overnight-grown bacteria were diluted 100 folds in fresh LB medium and grown to early stationary phase at 37°C with vigorous shaking. Bacteria were resuspended in 0.1 M sodium phosphate buffer (pH 7.4) with increasing concentrations of H₂O₂ (0 mM, 1 mM, 5 mM) without agitation. After 2 h incubation at 37°C, cultures were diluted and plated onto LB plates. All experiments were repeated at least three times.

**Biofilm formation assay.** Bacteria were diluted 100 folds in LB broth, and was then inoculated into each well of a 96-well polystyrene flat-bottom microtiter plates and statically incubated at 37°C for 24 h. Then the wells were washed with phosphate-buffered saline (PBS) to remove unattached cells. Crystal violet (0.1% w/v; Sigma) was used to stain the attached cells for 30 min at 25°C. The plates were then washed with PBS and left to dry for a further 10 min, and the stained biomass was solubilized in 1% (w/v) SDS. The absorbance of each well was determined at OD590 nm. All experiments were repeated at least three times.

**Scanning electron microscopy.** Bacterial cells, grown in LB medium, were fixed with 2.5% (vol/vol) glutaraldehyde in phosphate-buffered saline (PBS, pH 7.4) and subsequently post-fixed in 1% (wt/vol) osmium tetroxide. The samples were then dehydrated in acetone, critical-point dried, and coated with gold-palladium. The specimens were finally examined with a FEI Quanta 200 SEM scanning electron microscope (USA).

**Carbon source utilization assays and chemical sensitivity assays.** The Biolog GEN III MicroPlate was used to analyze the *K. pneumoniae* strains in 94 phenotypic tests, which including 71 carbon source utilization assay and 23 chemical sensitivity assays. Briefly, the bacterial culture were picked up from the surface of the BUG + agar plate (Biolog, CA, USA) using a sterile cotton-tipped swab and inoculated into the IF-A Inoculum (Biolog, CA, USA). The target cell density of Inoculum was set to 90–98% T by turbidimeter (BioMérieux, Lyon, France). Then 100 μl of the inoculum was added into each well of the 96 GEN III MicroPlate (Biolog, CA, USA). After incubating the culture for 24 hours at 37°C, the OD₅₉₀ readings were measured with a BIOLOG microplate reader automatically and further confirmed visually.

**Genome sequencing.** Genome sequencing was performed by Beijing Genomics Institute (BGI, China). Briefly, the genomic DNA for each bacterium was prepared by conventional phenol-chloroform extraction methods. A 500 bp paired-end library was constructed for each purified DNA sample following the standard Illumina paired-end protocol with a low-cycle polymerase chain reaction during the fragment enrichment, and sequencing was performed on the Illumina HiSeq 2000 with 90 cycles. Low quality reads were filtered using the DynamicTrim and LengthSort Perl scripts within SolexaQA. Short reads were assembled using SOAPdenovo version 2.04 and the gaps were closed by GapCloser version 1.12.

**Genetic mutations detection and phylogenetic analysis.** First, the short reads were aligned onto the *K. pneumoniae* ATCC BAA-2146 genome reference using the SOAP2 program. Second, SOAPnsnp was used to score SNPs from aligned reads. SOAPnsnp results were filtered as follows: (1) The read coverage of the SNP site was more than 30, (2) The Illumina quality score of either allele was more than 30, and (3) The count of all mapped best base was more than two times the count of all mapped second best base. In addition, BWA 0.6.2 and SAMTools 0.1.18 were used to confirm the SNP results. The Illumina reads were first aligned by BWA with default parameters for each sample. The aligned results were piped to SAMtools to perform SNP and Indel analysis.

**Electronic PCR for MLST genotyping and identification of drug resistance-associated genes.** MLST with seven genes (adk, infB, mdh, pgI, pheA, poiF and tonB) was performed on isolates according to the protocol described on the *K. pneumoniae* MLST website (www.pasteur.fr/mlst). STs were assigned by using the MLST database (www.pasteur.fr/mlst/Kpneumoniae.html). Electronic PCR was used to extract the drug resistance-associated genes, then the DNA sequences were annotated using the BLAST program at http://www.ncbi.nlm.nih.gov. Mutations in the *gyrA* and *parC* genes were identified by comparing the DNA sequences with *gyrA* and *parC* sequences of the *K. pneumoniae* (GenBank accession numbers DQ673325 and NC009648 for *gyrA* and *parC* respectively).

**RNA-Seq and comparative transcriptomic data analysis.** Total RNA isolation as well as construction and sequencing of cDNA libraries of the *K. pneumoniae* strains were conducted by BGI (China). Briefly, total RNA samples were isolated using a RNeasy Protect Bacteria Mini Kit (QiAGEN, Germany) according to the manufacturer’s instructions. Sequencing was carried out by running 90 cycles on the Illumina HiSeq 2000. Paired-end reads were mapped to *K. pneumoniae* ATCC BAA-2146 (CP006659) reference sequence using Bowtie2. The number of reads mapped to each gene was counted by using HTSeq-count (http://www.huber.embl.de/users/anders/HTSeq). The edgeR package was used to normalize the read count data and for differential gene expression analysis. Fold changes with FDR < 0.001 were considered to be statistically significant. RPKM (Reads Per kb Million reads) values were provided to enable comparison of relative transcript abundance among different samples.

**Functional annotation and enrichment analysis.** KOBASE 2.0 (KEGG Orthology Based Annotation System) was used to identify metabolic pathways and to calculate the statistical significance of each pathway. The COG annotation was performed using the Blastall software against the Cluster of Orthologous Groups (COG) database. COG enrichment analysis was determined by comparing the prevalent categories of differentially expressed genes assigned to a specific COG category to the prevalence of genes in the whole genome assigned to that COG category with a Fisher’s exact test.

**Quantitative RT-PCR of selected targets.** To validate whether RNA-Seq provides reliable quantitative estimates of transcript levels, qRT-PCR analysis was conducted for 10 randomly selected genes. cDNA samples were analyzed by quantitative PCR with KAPA SYBR FAST qPCR Kit (KAPA Biosystems) on ABI 7300 system (Applied Biosystems). Data were analyzed by the 2⁻ΔΔCt method and normalized to the *gyrA*.

Each experiment was performed in triplicates and repeated at least three times.

1. Little, M. L., Qin, X., Zerr, D. M. & Weissman, S. J. Molecular epidemiology of colonizing and disease-causing *Klebsiella pneumoniae* in pediatric patients. *J Med Microbiol* 63, 610–616, doi:10.1099/jmm.0.06335-0 (2014).
2. Shon, A. S., Bajwa, R. P. & Russo, T. A. Hypervirulent (hypermucoviscous) *Klebsiella pneumoniae*: a new and dangerous breed. *Virology* 4, 107–118, doi:10.4161/viru.22718 (2013).
28. Li, W.

27. Ortega, A. D., Gonzalo-Asensio, J. & Garcia-del Portillo, F. Dynamics of Klebsiella pneumoniae: Tracking Molecular Epidemiology and Outcomes through a Regional Network. Antimicrob Agents Chemother, doi:10.1128/AAC.02636-14 (2014).

25. Cavanagh, A. T. & Wassarman, K. M. 6S RNA, A Global Regulator of cell wall peptidoglycan biosynthesis. J Bacteriol 193, 2044–2051, doi:10.1128/JB.00064-11 (2011).

23. Kim, H., Goo, E., Kang, Y., Kim, J. & Hwang, I. Regulation of universal stress response in non-growing bacteria located inside Bacillus subtilis spores for a prolonged period of time in real space conditions. Astrobiology 13, 89–99, doi:10.1089/ast.2012.1104 (2013).

22. Bushnell, G., Mitrani-Gold, F. & Mundy, L. M. Emergence of New Delhi metallo-β-lactamase type 1-producing enterobacteriaceae and non-enterobacteriaceae: global case detection and bacterial surveillance. Int J Infect Dis 17, e164–e170, doi:10.1016/j.ijid.2013.03.010 (2013).

20. Qi, Y. et al. Genomic and transcriptomic analysis of NDM-1-producing Klebsiella pneumoniae isolates. J Med Microbiol 63, 101–108, doi:10.1099/jmm.0.05447-0 (2014).

19. Sangurdekar, D. P., Zhang, Z. & Khodursky, A. B. The association of DNA damage response regulator with the Cdc6 protein genes by quorum sensing and RpoS in Staphylococcus aureus. J Biol Chem 289, 4354–4366, doi:10.1074/jbc.M113.486721 (2014).

18. Ramirez, M. S. & Tolmasky, M. E. Aminoglycoside modifying enzymes. Antimicrob Agents Chemother 47, 308–314, doi:10.1128/AAC.47.1.308-314.2003 (2003).

17. Robicsek, A., Sahm, D. F., Strahilevitz, J., Jacoby, G. A. & Hooper, D. C. Broader surveillance for aminoglycoside resistance: importance of quinolone resistance and the role of plasmid-mediated enzymes in Klebsiella pneumoniae. J Antimicrob Chemother 65, 1054–1061, doi:10.1093/jac/dkt301 (2010).

15. Frye, J. G. & Jackson, C. R. Genetic mechanisms of antimicrobial resistance identified in Salmonella enterica, Escherichia coli, and Enterococcus spp. isolated from U.S. food animals. Front Microbiol 4, 135, doi:10.3389/fmicb.2013.00135 (2013).

14. Lee, K. W. et al. Biofilm development and enhanced stress resistance of a model, mixed-species community biofilm. ISME J 8, 994–1007, doi:10.1038/ismej.2014.194 (2014).

13. Lee, K. W. et al. Analysis of drug resistance determinants in Klebsiella pneumoniae isolates from a tertiary-care hospital in Beijing, China. PLoS One 7, e24220, doi:10.1371/journal.pone.0024220 (2012).

12. Andrade, L. N. et al. Expansion and Evolution of a Virulent, Extensively Drug-Resistant (Polymyxin B-Resistant), QnrS1-, CTX-M-2- and KPC-2-producing Klebsiella pneumoniae ST11 International High-Risk Clone. Clin Microbiol Infect 18, 128–131, doi:10.1111/j.1469-0691.2011.03647.x (2012).

11. Fynes, B. et al. Genetic analysis of antimicrobial resistance determinants on plasmids and transposons in Escherichia coli. Drug Resist Update 13, 151–157, doi:10.1016/j.dru.2010.08.003 (2010).

10. Sangurdekar, D. P., Zhang, Z. & Khodursky, A. B. The association of DNA damage response and nucleotide level modulation with the antibacterial mechanism of the anti-folate drug trimethoprim. BMC Genomics 12, 583, doi:10.1186/1471-2164-12-583 (2011).

9. Qi, Y. et al. ST11, the dominant clone of KPC-producing Klebsiella pneumoniae. J Antimicrob Chemother 66, 307–312, doi:10.1093/jac/dkt313 (2011).

8. Reid, G. et al. Effects of ciprofloxacin, norfloxacin, and ofloxacin on in vitro adhesion and survival of Pseudomonas aeruginosa AK1 on urinary catheters. Antimicrob Agents Chemother 47, 3077–3083, doi:10.1128/AAC.47.8.3077-3083.2003 (2003).

7. Bushnell, G., Mitrani-Gold, F. & Mundy, L. M. Emergence of New Delhi metallo-β-lactamase type 1-producing enterobacteriaceae and non-enterobacteriaceae: global case detection and bacterial surveillance. Int J Infect Dis 17, e164–e170, doi:10.1016/j.ijid.2013.03.010 (2013).

6. Al Bahry, S., Sivakumar, N. & Al-Khambashi, M. Effect of naldixic acid on the morphology and protein expression of Pseudomonas aeruginosa. Asian Pac J Trop Med 5, 265–269, doi:10.1016/S1995-7645(12)60037-6 (2012).

5. Kim, H., Goo, E., Kang, Y., Kim, J. & Hwang, I. Regulation of universal stress protein genes by quorum sensing and RpoS in Burkholderia glumae. J Bacteriol 194, 982–992, doi:10.1128/JB.00396-11 (2012).

4. van Duin, D. et al. Surveillance of Carbenapenem-Resistant Klebsiella pneumoniae: Tracking Molecular Epidemiology and Outcomes through a Regional Network. Antimicrob Agents Chemother, doi:10.1128/AAC.02636-14 (2014).

3. Lippmann, N., Lubbert, C., Kaisers, T., Kaisers, U. X. & Rodloff, A. C. Clinical epidemiology of Klebsiella pneumoniae carbapenemases. Lancet Infect Dis 14, 235–242, doi:10.1016/S1473-3099(14)70054-4 (2014).

2. van Duin, D. et al. Surveillance of Carbenapenem-Resistant Klebsiella pneumoniae: Tracking Molecular Epidemiology and Outcomes through a Regional Network. Antimicrob Agents Chemother, doi:10.1128/AAC.02636-14 (2014).

1. Munoz-Price, L. S. et al. Clinical epidemiology of the global expansion of Klebsiella pneumoniae carbapenemases. Lancet Infect Dis 13, 785–796, doi:10.1016/S1473-3099(13)70190-7 (2013).

Acknowledgments

Financial support was provided by the National Basic Research Programs of China (2014CB744400 and 2012CB518700), National Natural Science Foundation of China (81371769), and the Ministry of Health and the Ministry of Science and Technology, China (2013ZX10003006), and the Chinese Academy of Sciences (KJZD-EW-L02). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

J.L., F.L. and Q.W. contributed equally to this paper. C.H.L. and C.L. designed and wrote the manuscript. All other authors. All authors read and approved the final manuscript.

Additional information

Accession Codes: The whole genome sequencing and RNA-Seq reads data have been deposited in GenBank under accession SRP042268.

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests. How to cite this article: Li, J. et al. Genomic and transcriptomic analysis of NDM-1 Klebsiella pneumoniae in spaceflight reveal mechanisms underlying environmental adaptability. Sci. Rep. 4, 6216; DOI:10.1038/srep06216 (2014).
