

SCIENTIFIC REPORTS



OPEN

Meta-genomic analysis of toilet waste from long distance flights; a step towards global surveillance of infectious diseases and antimicrobial resistance

Received: 17 December 2014

Accepted: 17 April 2015

Published: 10 July 2015

Thomas Nordahl Petersen¹, Simon Rasmussen¹, Henrik Hasman², Christian Carøe¹, Jacob Bælum¹, Anna Charlotte Schultz², Lasse Bergmark², Christina A. Svendsen², Ole Lund¹, Thomas Sicheritz-Pontén¹ & Frank M. Aarestrup²

Human populations worldwide are increasingly confronted with infectious diseases and antimicrobial resistance spreading faster and appearing more frequently. Knowledge regarding their occurrence and worldwide transmission is important to control outbreaks and prevent epidemics. Here, we performed shotgun sequencing of toilet waste from 18 international airplanes arriving in Copenhagen, Denmark, from nine cities in three world regions. An average of 18.6 Gb (14.8 to 25.7 Gb) of raw Illumina paired end sequence data was generated, cleaned, trimmed and mapped against reference sequence databases for bacteria and antimicrobial resistance genes. An average of 106,839 (0.06%) reads were assigned to resistance genes with genes encoding resistance to tetracycline, macrolide and beta-lactam resistance genes as the most abundant in all samples. We found significantly higher abundance and diversity of genes encoding antimicrobial resistance, including critical important resistance (e.g. *bla*_{CTX-M}) carried on airplanes from South Asia compared to North America. Presence of *Salmonella enterica* and norovirus were also detected in higher amounts from South Asia, whereas *Clostridium difficile* was most abundant in samples from North America. Our study provides a first step towards a potential novel strategy for global surveillance enabling simultaneous detection of multiple human health threatening genetic elements, infectious agents and resistance genes.

Globally, infectious diseases are the cause of about 22% of all human deaths¹. Due to increasing population density, increasing use of antimicrobial agents, disruption of wildlife habitats, and an increase in global travel and trade this number has been estimated to increase in coming years^{2,3}. Rapid global surveillance systems for early detection of worldwide spread of zoonotic and human pathogens would enable outbreak control and epidemic prevention^{4,5}. In addition to the direct severe consequences for human health, infectious diseases may cause an increased financial burden on health systems and may imply restriction on travel, trade, etc. Global spread of antimicrobial resistance in bacterial pathogens is another major threat against human health, yet we don't know much about how these genes travel and spread worldwide⁶.

Current international disease surveillance system is mainly based on reports made by doctors after treatment of infected patients⁷. It has been attempted to establish regional and global networks, for the

¹Department of Systems Biology, Technical University of Denmark, Kgs. Lyngby, Denmark. ²National Food Institute, Technical University of Denmark, Kgs. Lyngby, Denmark. Correspondence and requests for materials should be addressed to F.M.A. (email: fmaa@food.dtu.dk)

exchange of information for single or few pathogens⁸. This approach suffers from numerous limitations, such as limited number of sites and target pathogens, as well as serious time-delays in reporting⁹. This has prompted attempts to initiate information platforms for exchange of data, based on reporting of clinical symptoms or Internet based data mining¹⁰. However, this will neither give pathogen nor gene specific data, and is still based on clinical disease symptoms.

As the current routine is mainly based on individual pathogens, not necessarily capturing all disease threats, for instance antimicrobial resistance, where the bulk of resistance genes may be present in the commensal bacterial flora¹¹. This is often insufficient and allows infectious agents and resistance genes to spread and infect large populations before they are even recognized^{4,5,12,13}. Furthermore, most current surveillance typically originates from developed countries and emerging outbreaks are often not detected until they appear there². All this results in a delayed and potentially biased global surveillance picture, where any taken actions often are too late².

As the cost of Next Generation Sequencing (NGS) has decreased this technology has become available for routine use¹⁴. The technology is applicable for complete sequencing of entire microbial communities¹⁵ and directly on clinical samples¹⁶. It has also been argued that the technology might have major implications for especially developing countries where NGS as a one-fits-all tool could enable these countries to avoid the establishment of more complex diagnostic methods as those already established in developed countries¹⁴. NGS combined with novel hot-spot sampling points could potentially provide simultaneous diagnosis of all known pathogens and resistance genes. As international airplane flights are known to be a major route of transmission for infectious diseases¹⁷, airports have previously been suggested as potential control points for limiting the spread of disease outbreaks¹⁸. In the past airline travel has been restricted to a narrow demographic. Due to dramatic improvements in cost, availability, logistics and airline networks, this is no longer the case, and airline travel is expanding rapidly to all segments of society and to regions previously considered remote. Clearly there are still biases in the population representation in airline traffic, however it is strategic to implement antimicrobial surveillance with modes of transport since they represent transmission of antimicrobial resistance and infectious diseases, and air travel is an important avenue to address, for which surveillance protocols are currently weak or non-existent. In 2012 approximately 1.11 billion passengers travelled on an international flight; a number expected to increase beyond 1.45 billion in 2016¹⁹. Thus, airports offer a potentially optimal site for global surveillance.

Here we investigated toilet waste from 18 international airplanes arriving in Copenhagen, Denmark, by complete sequencing of subsequent bioinformatics analysis quantifying the occurrence of all known antimicrobial resistance genes, as well as selected pathogens (Fig. 1). The prevalence of norovirus was also determined.

Results

Sequencing reads and phylogenetic clustering. An average of 18.6 Gigabytes (Gb) (ranging between 14.8 and 25.7 Gb) of raw Illumina paired end sequence data was generated. The sequencing reads were cleaned, trimmed and mapped against reference sequence databases (Tables S1–4).

An average of 106,839 (0.06%) reads were assigned to resistance genes, while on average 48.7%, 9.0%, 0.17% and 0.24% were assigned to MetaHit draft genomes, complete bacterial genomes, plasmids and human, respectively. An average of 41.8% of the reads could not be assigned to any reference sequence. Rarefaction of the reads mapping to complete bacterial genomes showed a strong degree of saturation in the sequence data suggesting that sufficient amount of data has been generated (Figure S2).

Clustering of samples based on bacterial abundance profiles showed separation of the airplane samples into several clusters, but with an overall association according to geography (Fig. 2A). Tetracycline, macrolide and beta-lactam resistance genes were the most abundant in all samples (Fig. 2B) and clustering according to resistance gene class abundance was also to some degree associated with geography. At the individual gene level (Figure S3) clustering also indicated an overall association according to geography. As expected the *strA* and *strB* genes, normally being genetically linked, clustered together, while clustering of *floR* and *sulI* genes can be explained through the association with class I integrons. The grouping of *tet(M)*, *tet(O)*, *tet(Q)* and *tet(W)* is likely due to the general high abundance of these genes across all samples.

Abundance of antimicrobial resistance genes. Compared to North America we found significantly higher abundance of antibiotic resistance genes among samples from South Asia, North Asia and all Asian samples combined (p-values 0.003; 0.005; 0.001) (Fig. 3A). Additionally the resistance gene richness was higher in samples from South Asia and all Asian samples combined, compared to North America (p-values 0.04 and 0.03) (Fig. 3A). In general an association between resistance gene abundance and richness was observed (Pearson Correlation Coefficient 0.69, Fig. 3B). These observations suggest a general independence of the individual genes and gene classes and an overall selection of a higher prevalence and diversity of resistance genes in mainly South Asian compared to American samples.

In-depth analysis of the abundance of individual resistance genes showed that in total 31 resistance genes were found to have significant differences (FDR < 0.05) based on geographical origin (Table S5). Comparing South Asian versus North American samples showed significantly higher abundances of *bla*_{CTX-M}, *qnrS* and the 16S methyltransferase *npmA*. A high abundance was also observed of *floR*

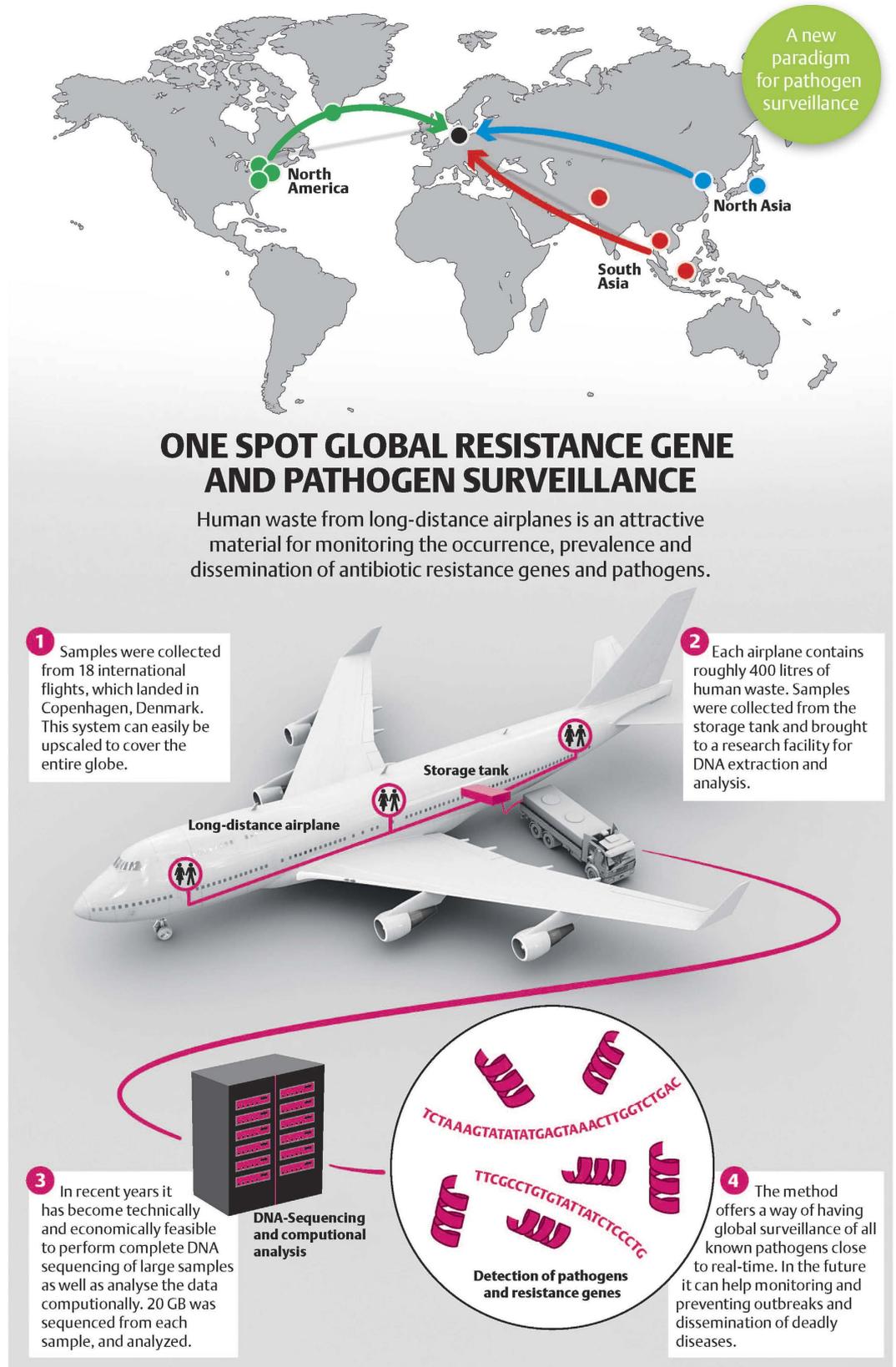


Figure 1. Origin of the 18 long-distance flights with destination being the international airport in Copenhagen, Denmark, as well as the analytic procedure applied. Figure created in Adobe Illustrator.

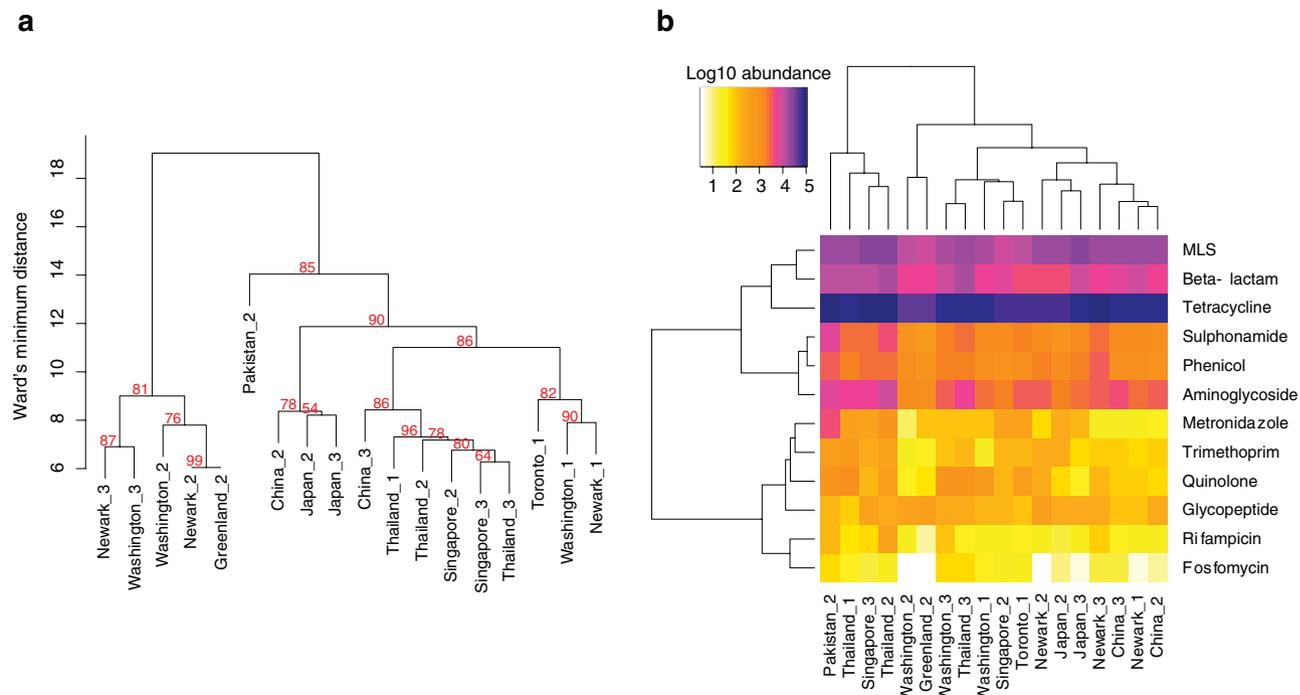


Figure 2. Geographical clustering of flight samples based on adjusted read abundance. a. Hierarchical clustering of flight samples based on normalized abundance of complete bacterial genomes and genomes from the human gut, bootstrap values are indicated in red and calculated using pvclust. Branch distance in the tree is Ward's minimum distance calculated using the Lance-Williams formula. b. Heatmap showing hierarchical clustering based on the normalized abundance of resistance gene classes. The abundance is in log10 scale from white (low), yellow, orange (intermediate), magenta, blue (high). MLS: macrolide, lincosamide, streptogramin.

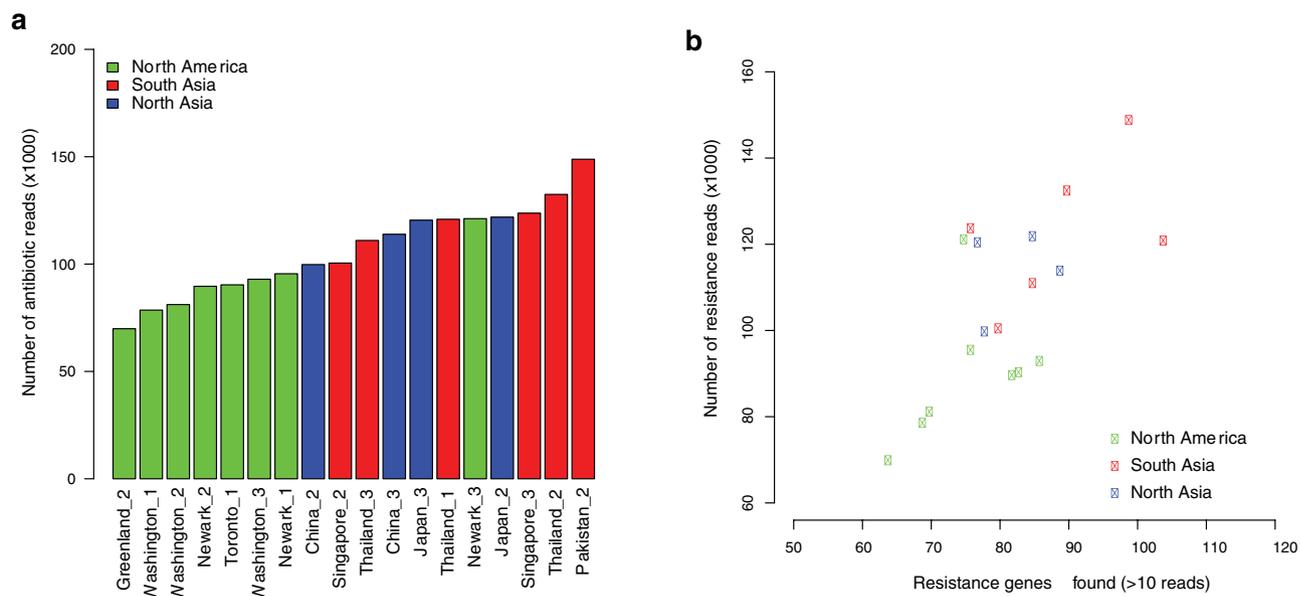


Figure 3. Antibiotic resistance genes detected in samples from different long-distance flights. A. Adjusted number of resistance reads identified in the flight samples. Green: North America, red: South Asia, blue: North Asia. B. The association between number of resistance genes found with more than 10 reads and the adjusted abundance of resistance reads per sample. Legend as in A.

(florphenicol/chloramphenicol resistance), *sul2* and *sul3* (sulphonamide resistance), *bla*_{OXA} and *bla*_{TEM} (beta-lactam resistance) and *tet(A)* and *tet(B)* (tetracycline resistance) in samples from South Asia. Additionally, the abundance of *erm(A)* and *erm(C)* (macrolide resistance) was significantly higher in North Asia compared to South Asia and of *erm(B)* significantly higher in North Asia compared to North America.

Abundance of selected pathogens. Unique reads mapping to the human pathogens *S. enterica*, *C. difficile* and *C. jejuni* were observed in all samples (Table S6). No regional differences could be observed for *C. jejuni*, but the abundance of *S. enterica* was higher among samples from South Asia compared to both North Asia ($P=0.053$) and North America ($P=0.031$). In contrast, *C. difficile* was less abundant in samples from South Asia compared to North America ($P=0.0067$) and North Asia ($P=0.089$). Quantitative real time PCR was used to detect and quantify human noroviruses genogroup (G) I and II (Supplementary materials). Ten and 14 samples were positive for GI and GII, respectively. Significantly higher values were observed for GII norovirus in South Asian compared to North Asian and North American samples, whereas no differences were observed for GI (Table S7).

Discussion

Antimicrobial resistance constitutes a huge global health problem⁶. It is well known that changes in diet, microbial intake and antibiotic exposure can rapidly and drastically alter the gut microbiome^{20,21}. In addition, several recent studies have shown major differences in gut composition across geography^{22–26}, probably related to environmental or diet differences. This suggests that the gut microbiome samples from a long-distance flight, to a large extent, reflect the microbiomes acquired at the site of boarding and not a normal “nationally” acquired flora. Strangely enough studies into the changes of the gut composition during travel are scarce, even though it has been indicated that major changes in both the gut composition and occurrence of antimicrobial resistance may occur²⁷ and that foreign travel is a major risk factor for acquisition of selected resistance genes²⁸. Unfortunately no relevant data for comparison to the current study exist, however, in general our observations are in agreement with recent data from WHO⁶.

Some of the most recent examples on rapid global spread of antimicrobial resistance, have been the emergence of plasmid mediated quinolone resistance, extended spectrum beta-lactamases mainly the CTX-M enzymes, and high-level aminoglycoside resistance due to 16S methyltransferases^{29–31}. CTX-M is considered endemic to Europe and Asia²⁹, whereas reports from North America remain sporadic. In agreement with this we found a higher frequency of CTX-M and *qnrS* genes among the samples from Asia compared to North America.

The patterns observed for the selected pathogens are also supported by previous observations^{32–35}. Thus, travellers to North America have less risk of disease caused by *Salmonella* or norovirus compared to travellers to Asia, whereas, *C. difficile* infections seems to be mainly associated with North America and Europe. However, as for resistance genes no current global surveillance exists, making comparison to other studies difficult.

There is currently only very limited information on the global occurrence and transmission of antimicrobial resistance and infectious diseases. Here we showed the feasibility of collecting and analysing long-distance airplane toilet waste for determining the occurrence of antimicrobial resistance genes and selected human pathogens. In addition to insight into the global macro-transmission of infectious agents and antimicrobial resistance, this may also open a novel avenue for simultaneous global surveillance of all infectious agents and quantification of transmission into different countries. Rather than the current surveillance approach where pathogen specific networks are established at laboratories in a limited number of countries, it might be more rapid and cost-efficient with sampling and analysis of the human waste that is transmitted around the World. Such a system is also flexible towards identifying and sampling putative hot-spots and will make it easier to obtain samples from a new area. A number of obstacles do however, need to be overcome to make this happen. Thus, in order to provide data for action it would be necessary with large scale sequencing facilities close to the sampling site, as well as a global database for real-time exchange of data. The current study included only 18 samples and it would even with the current technologies be a challenge to sequence samples from hundreds of flights on a weekly basis. In addition, more studies are needed to determine how representative the results are of the country of origin. In general airline travel is expanding rapidly and the demographic is not as biased at present compared to years ago. A growing group of people is travelling and they can of-course act as carriers of diseases and antimicrobial resistance genes around the World.

In future efforts, the large-scale DNA sequencing performed here should be supplemented with complete sequencing of RNA viruses. However, this will still only detect the current presence of infectious agents. Detection of specific antibodies may predict previous or recent presence of infectious agents in a given population. More antibodies are generated and excreted in feces and urine that are produced in the remaining body, which recently have led to the exploration of such samples for diagnosis of diseases^{36,37}. In the future this may also be combined with microbial sequencing to provide an even more complete picture of the prevalence and global transmission of diseases.

Methods

Sampling. Pooled air-toilet samples were obtained from 18 long-distance flights from Scandinavia Airline System (SAS) (www.sas.com) arriving to Copenhagen airport from nine different cities (Bangkok, Beijing, Islamabad, Newark, Kangerlussuaq, Tokyo, Toronto, and Washington DC), representing eight countries, in three different regions (North America, North and South Asia) (Fig. 1). At the sites of origin toilets were emptied, rinsed with water and disinfected prior to departure. Idu-Flight was used as disinfection agent, which is a deodorizing liquid, based on glutaraldehyde and benzalkoniumchloride. On arrival to Copenhagen the air flight waste containers were emptied by the SAS Cleaning and Service. For our purpose, one car was designated to empty the airplane waste containers, which were roughly rinsed with water prior to collecting samples. From each of the individual airplanes 6–8 waste toilets were emptied under high vacuum pressure into the special container at the service car. This procedure mixes the content thoroughly and no individual fecal clumps or toilet paper can be visually identified. Three individual ½ L samples were then collected using sterile tubes from the cars waste container and placed in a refrigerator. Subsequently, a container was rinsed with water before collecting toilet waste from another airplane. Between one and seven flights were collected during a maximum of 12 hours and immediately transported in closed containers for hazardous goods to the Technical University of Denmark and processed immediately.

Sample handling and DNA purification. On arrival the samples were handled as outlined in Figure S1. Each individual ½ litre tube was mixed as much as possible in the tube. Four tubes, one of them supplemented with RNAlater, of 10 mL each were collected from each ½ litre tube and frozen at -80°C . Twenty-five mL were collected from each of the three ½ L tubes obtained from a single airplane and mixed. Again four tubes of 10 mL were collected and stored as described above. The remaining approximately 35 mL combined waste was transferred to a 50 mL centrifuge tube and centrifuged at 1500 rpm for 2 minutes to remove large debris and large cells. A total of 1.2 mL was removed for conventional counting of bacteria and examination for antimicrobial resistance. A total of 30 mL supernatant was transferred to a new 50 mL tube. The supernatant was centrifuged at 10000 G for 10 min. Ten mL of the supernatant was collected and frozen at -80°C ; the remaining was discarded. The pellet with the bacterial cells was dissolved in 500 μL phosphate buffered saline and transferred to Eppendorf Safelock™ tubes and DNA purified using a protocol including both lysozyme and lysostaphin to increase cell lysis³⁸. Each of the DNA purifications were dissolved in 300 μL TE buffer and extracted with 2 rounds of phenol/chloroform, followed by precipitation to obtain highest possible purity. The purified DNA was frozen at -80°C for later sequencing.

Whole community sequencing. Samples were diluted up to a final volume of 100 μL and fragmented using a Diagenode Bioruptor, with a 30'/30' on/off program of 20 cycles. 500 ng of fragmented DNA from each sample was used to convert the extracts into Illumina-compatible DNA libraries using NEBNext library building kit for second-generation sequencing (New England Biolabs, Ipswich, MA; Cat#. E6070L). Libraries were prepared according to manufacturer's directions with slight modifications as follows: Following end repair with incubation of 20 minutes at 12°C followed by 15 minutes at 37°C the samples were purified on a Qiagen Minelute silica spin-column following manufacturer's instructions. The DNA was ligated to Illumina blunt-end adapters as described by Meyer *et al.*³⁹, for 20 minutes at 20°C and again purified on a Qiagen qiaquick silica spin-column following manufacturer's instructions. Subsequently, the fill-in procedure was performed for 20 minutes at 65°C followed by a heat inactivation step for 20 minutes at 80°C . Libraries were then used directly for post-library indexing PCR with unique indexing for each sample³¹. Indexing PCR was done in 100 μL reactions using Phusion High-Fidelity PCR Master Mix (Thermo Fisher Scientific, Waltham, MA) and 10 μL template DNA library under the following conditions: 1) 30 seconds at 98°C , 2) 20 seconds at 98°C , 3) at 30 seconds at 60°C , 4) 30 seconds at 72°C , 5) 5 minutes at 72°C , 6) hold at 4°C . Step 2–4 was repeated for 4 cycles. Subsequently, the PCR products were cleaned with Qiagen qiaquick spin columns. Concentrations were measured on a Qubit fluorometer 2.0 (Life Technologies) using the dsDNA HS assay and subsequently analyzed on a Agilent 2100 Bioanalyzer (Agilent technologies). All 18 samples were pooled into one pool in equimolar concentrations for a final concentration of 9 nM. Sequencing was performed by the National High-throughput DNA Sequencing Centre at the University of Copenhagen, Denmark, on an Illumina HiSeq 2000 instrument by 100 cycles paired end for a total of 8 lanes.

Concentration Read processing and alignment. Paired end read data was processed by mapping airtoilet fastq samples against several reference sequence databases, which can be downloaded as described in Table S2. Initially, trimming and removal of adaptor sequences was done using *cutadapt*⁴⁰ with settings for minimum read length being 30 bp and a minimum Phred quality score of 30, to trim low-quality reads before adaptor removal (*cutadapt* parameter - quality-cutoff). Subsequently reads were processed through a mapping approach build on *bwa mem*⁴¹ and *samtools*⁴² software. *Bwa mem* (ver. 0.7.7-r441) was used with default settings to map reads against reference sequence databases. Throughout the mapping approach, only the most reliable hits were accepted i.e. properly paired reads were accepted

provided that each read maps with an alignment length being at least 80% of the read length. As part of the initial cleaning process, potential PhiX i.e. PhiX174 control reads were removed using the procedure described above and all remaining reads are likely to have a biological origin. The read statistics for the pre-processing steps are shown in Table S1 for each of the 18 samples available. Next, reads were mapped to a number of reference sequence databases following one of two possible routes, being either chainmode or fullmode mapping. A flowchart of the two mapping procedures is shown in Figure S4. All samples were mapped against the databases that can be created from public available reference sequences. In Fullmode a resistance gene database (ResFinder) was used. In Chainmode the follow ordered list of databases were used: MetaHitAssembly, Bacteria, Plasmid, Human, Invertebrates, Protozoa and Virus. Reads that mapped to the Bacteria database were subsequently used to extract hits of pathogen specific origin. Three pathogens with the highest read counts were considered as shown in Table S6 (*Salmonella enterica* subsp. *enterica* serovar Typhimurium DT104, *Clostridium difficile* R20291, *Campylobacter jejuni* subsp. *jejuni* IA3902). In this table we only show the most confidently mapped reads and refer to them as unique meaning those where the *bwa* flags 'AS' and 'XS' differ. The flag 'AS' is the alignment score and 'XS' is the score for the second best alignment.

Normalized abundance estimation, clustering and significance testing. The raw read counts were normalized by the total number of reads after PhiX filtering, summed to species level. All species with an abundance of more than 1,000 reads were transformed by log10 and clustered in R using the package pvclust using the ward method euclidean distance measure. The raw read counts of resistance genes were normalized by the total bacterial count i.e. sum of hits that could be assigned either of the to bacterial databases 'Bacteria' downloaded from NCBI complete genomes and 'MetaHitassembly' which is a collection of assembled draft genomes identified from an analysis of fecal samples from 124 European individuals¹⁵. Conversion tables from ResFinder identifiers i.e anti-microbial resistance genes, to gene and class level are available in Table S8. Significance testing was performed using geographical region as levels and tested individually using a Wilcoxon Rank Sum test. False Discovery Rate (FDR) was determined by using the average of three Monte Carlo simulations. The significant testing of sample-wide abundance and richness were performed using Wilcoxon Rank Sum test.

Extraction of viral RNA and real-time RT-PCR detection of noroviruses. Five ml portions of sampled abattoir material (toilet content including glutaraldehyde and benzalkoniumchloride, pH 9.5) were mixed by vortexing with 35 ml PBS and left for shaking for 30 min at room temperature to obtain a homogeneous suspension as well as allowing viruses to detach from sample material. The mixture was centrifuged at $10,000 \times G$ for 30 min to precipitate as much non-viral material, such as toilet paper, tissue and bacteria, as possible and the recovered supernatant were adjusted to pH 7.5 using HCL. Viruses were then precipitated from the supernatant by incubation with polyethylene glycol 8000 (80 g/L) and NaCl (17.5 g/L NaCl) during agitation (350 rpm) overnight at 4 °C. Pellet was resuspended in 0.75 ml PBS and the suspension was clarified by chloroform/butanol extraction (1:1 vol/vol), during which the mixture was vortexed for 30 sec followed by 5 min storage at room temperature before centrifugation at $10,000 \times G$ for 15 min at 4 °C. Nucleic acids were extracted from the remaining water phase using BioMerieux reagents and miniMag apparatus (BioMerieux, Herlev, Denmark) according to the protocol of the manufacturer except for eluting the nucleic acids in 100 µl elution buffer. One-step real-time RT-PCR for the detection of NoV genogroup I and II in 2.5 µl extracted RNA was carried out as described elsewhere⁴³ using a StepOnePlus™ System real-time PCR machine (Life Technologies Europe BV, Naerum, Denmark). Quantification of NoV GI and GII genome copies was done by interpolation to standard curves derived from NoV GI.3b and GII.1 RNA transcripts. Inhibition of amplification during RT-PCR detection of NoV GI and GII was evaluated by adding GI.3b and GII.1 transcripts, respectively, as internal amplification controls to each sample RNA extract and to nucleic acid-free water. The amplification efficiency was calculated by dividing the RNA transcripts recovered from the sample RNA by the RNA transcripts recovered from the nucleic acid-free water and was accepted for values above 50%. The extraction efficiency of viral nucleic acid was evaluated in all samples using Mengovirus, strain vMC₀ (ATCC VR-1597) as internal process control. Prior to extraction, approximately 1×10^4 plaque forming units of MC₀ were spiked into the samples and nucleic acid-free water. After extraction, the levels of MC₀ were determined by realtime RT-PCR^{43,44}. The extraction efficiency was calculated by dividing the MC₀ RNA recovered from the abattoir matrix by the MC₀ RNA recovered from the nucleic acid-free water and was accepted for values above 1%. The extraction efficiency was only used to evaluate the method performance and not incorporated into the calculation of virus concentration in the air flight samples. Significant testing was performed using Wilcoxon Rank Sum test.

Ethics. This study was conducted in accordance with the Danish Act on scientific ethical treatment of health research, as administrated and confirmed by the Research Ethics Committees of the Capital Region of Denmark (www.regionh.dk), Journal nr.: H-14013582.

Sequence data for the flight samples are available for download through ENA.

References

- Lozano, R. *et al.* Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2095–2128 (2012).
- Jones, K. E. *et al.* Global trends in emerging infectious diseases. *Nature* **451**, 990–993 (2008).
- Carroll, S. P. *et al.* Applying evolutionary biology to address global challenges. *Science* **346**, 1245993; doi: 10.1126/science (2014).
- Braden, C. R., Dowell, S. F., Jernigan, D. B., Hughes, J. M. Progress in global surveillance and response capacity 10 years after severe acute respiratory syndrome. *Emerg. Infect. Dis.* **19**, 864–869 (2013).
- Newell, D. G. *et al.* Food-borne diseases - the challenges of 20 years ago still persist while new ones continue to emerge. *Int. J. Food Microbiol.* **139**, S3–15 (2010).
- World Health Organization. Antimicrobial resistance: global report on surveillance. http://apps.who.int/iris/bitstream/10665/112642/1/9789241564748_eng.pdf?ua=1. (2014) Date of access: 12/12/2014.
- Gibbons, C. L. *et al.* Burden of Communicable diseases in Europe (BCoDE) consortium. Measuring underreporting and underascertainment in infectious disease datasets: a comparison of methods. *BMC Public Health* 2014; **14**: 147 (2014).
- Arita, I. *et al.* Role of a sentinel surveillance system in the context of global surveillance of infectious diseases. *Lancet Infect. Dis.* **4**, 171–177 (2004).
- Chan, E. H. *et al.* Global capacity for emerging infectious disease detection. *Proc. Natl. Acad. Sci. USA* **107**, 21701–21706 (2010).
- Brownstein, J. S., Freifeld, C. C., Madoff, L. C. Digital disease detection--harnessing the Web for public health surveillance. *N. Engl. J. Med.* **360**, 2153–2155 (2009).
- Wright, G. D. The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat. Rev. Microbiol.* **5**, 175–186 (2007).
- Johnson, A. P., Woodford, N. Global spread of antibiotic resistance: the example of New Delhi metallo- β -lactamase (NDM)-mediated carbapenem resistance. *J. Med. Microbiol.* **62**, 499–513 (2013).
- Petty, N. K. *et al.* Global dissemination of a multidrug resistant *Escherichia coli* clone. *Proc. Natl. Acad. Sci. USA* **111**, 5694–5699 (2014).
- Aarestrup, F. M. *et al.* Integrating genome-based informatics to modernize global disease monitoring, information sharing, and response. *Emerg. Infect. Dis.* **18**, e1 (2012).
- Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
- Hasman, H. *et al.* Rapid whole-genome sequencing for detection and characterization of microorganisms directly from clinical samples. *J. Clin. Microbiol.* **52**, 139–146 (2014).
- Khan, K. *et al.* Entry and exit screening of airline travellers during the A(H1N1) 2009 pandemic: a retrospective evaluation. *Bull. World Health Organ.* **91**, 368–376 (2013).
- UNWTO – world tourism barometer. http://dtxqt4w60xqpw.cloudfront.net/sites/all/files/pdf/unwto_barom13_01_jan_excerpt_0.pdf. (2013) Date of access: 26/11/2014.
- World Health Report 2007 – a safer future: global public health security in the 21st century. http://www.who.int/whr/2007/whr07_en.pdf. (2007) date of access: 26/11/2014.
- David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563 (2014).
- Power, S. E., O'Toole, P. W., Stanton, C., Ross, R. P., Fitzgerald, G. F. Intestinal microbiota, diet and health. *Br. J. Nutr.* **111**, 387–402 (2014).
- Blaser, M. J. *et al.* Distinct cutaneous bacterial assemblages in a sampling of South American Amerindians and US residents. *ISME J.* **7**, 85–95 (2013).
- De Filippo, C. *et al.* Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl. Acad. Sci. USA* **107**, 14691–14696 (2010).
- Escobar, J. S., Klotz, B., Valdes, B. E., Agudelo, G. M. The gut microbiota of Colombians differs from that of Americans, Europeans and Asians. *BMC Microbiol.* **14**, 311 (2014).
- Lin, A. *et al.* Distinct distal gut microbiome diversity and composition in healthy children from Bangladesh and the United States. *PLoS One* **8**, e53838 (2013).
- Yatsunencko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222–227 (2012).
- Gaarslev, K., Stenderup, J. Changes during travel in the composition and antibiotic resistance pattern of the intestinal *Enterobacteriaceae* flora: results from a study of mecillinam prophylaxis against travellers' diarrhoea. *Curr. Med. Res. Opin.* **9**, 384–387 (1985).
- von Wintersdorff, C. J. *et al.* High rates of antimicrobial drug resistance gene acquisition after international travel, the Netherlands. *Emerg. Infect. Dis.* **20**, 649–657(2014).
- Cantón, R., Coque, T. M. The CTX-M beta-lactamase pandemic. *Curr. Opin. Microbiol.* **9**, 466–475 (2006).
- Fritsche, T. R., Castanheira, M., Miller, G. H., Jones, R. N., Armstrong, E. S. Detection of methyltransferases conferring high-level resistance to aminoglycosides in enterobacteriaceae from Europe, North America, and Latin America. *Antimicrob. Agents Chemother.* **52**, 1843–1845 (2008).
- Robicsek, A., Jacoby, G. A., Hooper, D. C. The worldwide emergence of plasmid-mediated quinolone resistance. *Lancet Infect. Dis.* **6**, 629–640 (2006).
- Patel, M. M. *et al.* Systematic literature review of role of noroviruses in sporadic gastroenteritis. *Emerg. Infect. Dis.* **14**, 1224–1231 (2008).
- Ekdahl, K., de Jong, B., Wollin, R., Andersson, Y. Travel-associated non-typhoidal salmonellosis: geographical and seasonal differences and serotype distribution. *Clin. Microbiol. Infect.* **11**, 138–144 (2005).
- Paschke, C. *et al.* Controlled study on enteropathogens in travellers returning from the tropics with and without diarrhoea. *Clin. Microbiol. Infect.* **17**, 1194–1200 (2011).
- He, M. *et al.* Emergence and global spread of epidemic healthcare-associated *Clostridium difficile*. *Nat. Genet.* **45**, 109–113 (2013).
- Carvalho, J. J. *et al.* Non-invasive monitoring of immunization progress in mice via IgG from feces. *In Vivo* **26**, 63–69 (2012).
- Sawangsoda, P. *et al.* Diagnostic values of parasite-specific antibody detections in saliva and urine in comparison with serum in opisthorchiasis. *Parasitol. Int.* **6**, 196–202 (2012).
- Aarestrup, F. M., Wegener, H. C., Rosdahl, V. T. Evaluation of phenotypic and genotypic methods for epidemiological typing of *Staphylococcus aureus* isolates from bovine mastitis in Denmark. *Vet. Microbiol.* **45**, 139–150 (1995).
- Meyer, M., Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc.* 2010, [pdb.prot5448](http://dx.doi.org/10.1101/pdb.prot5448); doi: 10.1101/pdb.prot5448 (2010).
- Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
- Li, H., Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler Transform. *Bioinformatics* **26**, 589–595 (2010).
- Li, H. *et al.* The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Le Guyader, F. S. *et al.* Detection and quantification of noroviruses in shellfish. *Appl. Environ. Microbiol.* **75**, 618–624 (2009).
- Costafreda, M. I., Bosch, A., Pintó, R. M. Development, evaluation, and standardization of a real-time TaqMan reverse transcription-PCR assay for quantification of hepatitis A virus in clinical and shellfish samples. *Appl. Environ. Microbiol.* **72**, 3846–3855 (2006).

Acknowledgements

Tom Sørensen and SAS cleaning service are thanked for assistance in obtaining samples. We are grateful to Carsten Bidstrup, Hanne Nørgaard Nielsen and Maria Louise Johansen for technical assistance. This study has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 643476 and The Villum Foundation (VKR023052). The norovirus detection was partly funded by Aquavalens (EU FP7-KBBE-2012-6).

Author Contributions

F.M.A. and T.S.P. conceived the idea. F.M.A. ensured funding and samples. F.M.A., H.H., T.S.P., O.L., T.N.P., J.B. and S.R. planned the experiments and analytic procedures. T.N.P. performed the bioinformatics analysis in collaboration with S.R. and J.B. C.A.S., C.C. and L.B. performed the laboratory analysis. A.C.S. designed and performed the analysis for quantification of noroviruses. F.M.A. and T.N.P. wrote the paper with input from all co-authors.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Nordahl Petersen, T. *et al.* Meta-genomic analysis of toilet waste from long distance flights; a step towards global surveillance of infectious diseases and antimicrobial resistance. *Sci. Rep.* 5, 11444; doi: 10.1038/srep11444 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>