

Cell-free DNA (cfDNA) and exosome profiling from a year-long human spaceflight reveals circulating biomarkers

Daniela Bezdan¹, Kirill Grigorev¹, Cem Meydan¹, Fanny A. Pelissier Vatter², Michele Cioffi², Varsha Rao³, Kiichi Nakahira⁴, Philip Burnham⁵, Ebrahim Afshinnkoo^{1,6,7}, Craig Westover¹, Daniel Butler¹, Chris Moszary¹, Matthew MacKay¹, Jonathan Fook¹, Tejaswini Mishra³, Serena Lucotti², Brinda K. Rana⁸, Ari M. Melnick⁹, Haiying Zhang¹⁰, Irina Matei², David Kelsen¹⁰, Kenneth Yu¹⁰, David C Lyden², Lynn Taylor¹¹, Susan M Bailey¹¹, Michael P. Snyder³, Francine E. Garrett-Bakelman^{12,13,14}, Stephan Ossowski¹⁵, Iwijn De Vlaminck¹⁶, Christopher E. Mason^{1,6,7,17}*

Affiliations:

¹Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA

²Children's Cancer and Blood Foundation Laboratories, Departments of Pediatrics, and Cell and Developmental Biology, Drukier Institute for Children's Health, Meyer Cancer Center, Weill Cornell Medical College, New York, NY, USA

³Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA

⁴Nara Medical University, Kashihara, Nara, Japan

⁵Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104

⁶The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY, USA

⁷The WorldQuant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, NY, USA

⁸Department of Psychiatry University of California, San Diego, La Jolla, CA, USA

⁹Department of Medicine, Weill Cornell Medicine, New York, NY, USA

¹⁰Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA

¹¹Department of Environmental & Radiological Health Sciences, Colorado State University, Fort Collins, CO, USA.

¹²Department of Medicine, University of Virginia School of Medicine, Charlottesville, VA, USA

¹³Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, Charlottesville, VA, USA

¹⁴University of Virginia Cancer Center, Charlottesville, VA

¹⁵Institute of Medical Genetics and Applied Genomics, University of Tübingen, Tübingen, Germany

¹⁶Nancy E. and Peter C. Meinig School of Biomedical Engineering, Cornell University, Ithaca, NY, USA

¹⁷The Feil Family Brain and Mind Research Institute, Weill Cornell Medicine, New York, NY, USA

*Corresponding Author

Christopher E. Mason

Weill Cornell Medicine

1305 York Ave., Y13-05

New York, NY 10021

Tel: 203-668-1448

E-mail: chm2042@med.cornell.edu

Keywords: NASA, cfDNA, liquid biopsy, mtDNA, mtRNA, exosomes, International Space Station, NASA Twins Study

Abstract

The health impact of prolonged space flight on the human body is not well understood. Liquid biopsies based on cell-free DNA (cfDNA) or exosome analysis provide a noninvasive approach to monitor the dynamics of genomic, epigenomic and proteomic biomarkers, and the occurrence of DNA damage, physiological stress, and immune responses. To study the molecular consequences of spaceflight we profiled cfDNA isolated from plasma of an astronaut (TW) during a year-long mission on the International Space Station (ISS), sampling before, during, and after spaceflight, and compared the results to cfDNA profiling of the subject's identical twin (HR) who remained on Earth, as well as healthy donors. We characterized cfDNA concentration and fragment size, and the positioning of nucleosomes on cfDNA, observing a significant increase in the proportion of cell-free mitochondrial DNA inflight, suggesting that cf-mtDNA is a potential biomarker for space flight-associated stress, and that this result was robust to ambient transit from the International Space Station (ISS). Analysis of exosomes isolated from post-flight plasma revealed a 30-fold increase in circulating exosomes and distinct exosomal protein cargo, including brain-derived peptides, in TW compared to HR and all known controls. This study provides the first longitudinal analysis of astronaut cfDNA during spaceflight, as well as the first exosome profiles, and highlights cf-mtDNA levels as a potential biomarker for physiological stress or immune system responses related to microgravity, radiation exposure, and other unique environmental conditions on the ISS.

Introduction

A wide range of physiological effects impact the human body during a prolonged stay in microgravity, such as headward fluid shift, atrophy of muscles, and decreases in bone density, which have been described for astronauts on the international space station (ISS)(Williams et al., 2009). In recent years, an increasing number of government and private space agencies have formed, and missions to the Moon and Mars are now planned for the late 2020s and 2030s (Iosim et al, 2020). These pending missions may span 30 months and require landing on a planet with almost no clinical infrastructure for medical monitoring or treatments. Yet, data on physiological changes of long-term missions (>6 months) is almost non-existent. These long-duration missions and the increasing exposure of humans to spaceflight-specific conditions necessitates the study of molecular changes in the human body induced by exposure to spaceflight stressors such as microgravity, radiation, noise, restricted diet, and reduced physical work opportunities. The NASA Twins study (Garrett-Bakelman et al., 2019) enabled interrogation of the impact of prolonged spaceflight on the human biology and cell-to-cell variations in the immune system (Gertz et al., 2020); however, there has never been a study on the impact of spaceflight on cell-free DNA (cfDNA).

Molecular signatures informative of human health and disease can be found in cfDNA and nucleic acids isolated from plasma, saliva, or urine (Heitzer et al., 2018; Hummel et al., 2018; Siravegna et al., 2017; Verhoeven et al., 2018; Volik et al., 2016). Non-invasive methods for monitoring health-related biomarkers in liquids such as plasma ('liquid biopsy') have already been successfully introduced in a wide range of contexts, including: prenatal testing for detection of trisomy and micro-deletions (Bianchi et al., 2014; Zhang et al., 2019), cancer diagnostics (Bettegowda et al., 2014; Diehl et al., 2008; Wang et al., 2017), monitoring of cancer therapies

(Birkenkamp-Demtröder et al., 2016; Wan et al., 2020), monitoring of the health of solid-organ transplants (Verhoeven et al., 2018; De Vlaminck et al., 2014), and screening for infections (Blauwkamp et al., 2019; Burnham et al., 2018; De Vlaminck et al., 2013). Hence, liquid biopsy is a potentially useful method for monitoring physiologic conditions of astronauts before, during and after spaceflight.

Indeed, cfDNA is extremely dynamic and responsive, providing strong indicators of DNA damage and tumor growth in distal tissues (Newman et al., 2016), immune response or infection (Zwirner et al., 2018), and RNA regulatory changes, with an innate capacity to reveal the cells of origin undergoing apoptosis or necrosis (Thierry et al., 2016). Various studies have reported changes in cfDNA concentration (Zwirner et al., 2018), cfDNA fragment length distribution (Mouliere et al., 2011; Underhill et al., 2016), mutation profiles and signatures (Newman et al., 2016), and cfDNA methylation (Shen et al., 2018) indicative of physiological conditions such as cancer. Mitochondrial DNA (mtDNA) can also be found in the extracellular space, circulating as short DNA fragments, encapsulated in vesicles and even as whole functional mitochondria (Amir Dache et al., 2020; Song et al., 2020). Several recent studies observed increased levels of cell-free mitochondrial DNA (cf-mtDNA) in psychological conditions (Lindqvist et al., 2016, 2018) and reduced cf-mtDNA levels in Hepatitis B infected patients associated with a higher risk of developing hepatocellular carcinoma (Li et al., 2016). However, since no such information exists for using these metrics for astronauts, we investigated the utility of cfDNA for the monitoring of the physiologic conditions of astronauts to spaceflight.

Of note, cfDNA comprises the footprints of nucleosomes, and these nucleosome features enable tracing of the tissue-of-origin for cfDNA in normal and disease states, through analysis of nuclear architecture, gene structure and expression (Murtaza and Caldas, 2016; Snyder et al., 2016). In particular, nucleosome positioning and depletion of short cfDNA sequences reveal footprints of transcription factor binding, promoter activity, and splicing, ultimately informing gene regulatory processes in the tissue/cell of origin (Snyder et al., 2016). Similar information can be revealed from exosomes, which are nano-sized vesicles (size 30–150nm) derived from perinuclear luminal membranes of late endosomes/multivesicular bodies and released into extracellular environment via multivesicular body fusion within the cell membrane ((Kalluri and LeBleu, 2020; Mathieu et al., 2019)) that can mediate long-range physiological crosstalk (Hoshino et al., 2015; Mathieu et al., 2019). Exosomes act as vehicles for horizontal transfer of information through their cargo: proteins, lipids, metabolites and DNA, as well as coding and non-coding RNAs (Valadi et al., 2007; Wortzel et al., 2019). Moreover, exosomes can be powerful mediators of responses to environmental stimuli as external and physiological stress impact their release, cargo and function, contributing to pathogenesis (Harmati et al., 2019; O'Neill et al., 2019; Qin et al., 2020). Since exosomes are abundant in plasma, they are critical components of liquid biopsies (Colombo et al., 2014; Hoshino et al., 2020) and analysis of their content can complement the information obtained from cfDNA, but there is no information about exosomes in astronauts.

To address this gap in knowledge, we profiled cfDNA isolated from plasma samples before, during, and after the one-year mission on the International Space Station (ISS) to evaluate the utility of cfDNA as a means to monitor physiological problems during extended missions in space. We also profiled the exosomes of both astronauts after the mission completion. While bulk RNA sequencing data have shown widespread gene expression changes in astronauts, including

mitochondrial RNA (mtRNA) spikes in flight samples from the One-Year Mission (Garrett-Bakelman et al., 2019), there has not yet been a study of astronauts that has leveraged cfDNA and exosomes. We focused on quantitative measures such as the levels of mitochondrial DNA, cfDNA fragment length, and the depletion of nucleosome signatures at transcription start sites. Together, our NGS results provide a “whole-body molecular scan”, which can provide a novel measurement of the impact of spaceflight on the human body, as well as serve as a continued metric of physiology and cellular stress for future long-duration missions.

Results

Study design and sample collection

We analyzed circulating cfDNA of a pair of male monozygotic twins over two years, starting when they were both 50 years old. During the NASA Twin Study, the flight subject (TW) was aboard the International Space Station (ISS) for 340 days, while his identical twin, the ground subject (HR), remained on Earth. We collected cfDNA at 12 time points from HR and 11 time points from TW. Of the latter, four samples were collected in-flight on board of the ISS or space shuttle. In addition, we profiled the cfDNA of an unrelated control subject (MS) to simulate the ambient return from the ISS. To control for ambient return (AR) effects (return of samples in the Soyuz capsule) on the molecular signatures of cfDNA, we subjected two MS samples and one HR sample to an extended shipping procedure (see Methods). Plasma and cfDNA were extracted using the same protocol for all samples (Methods). We observed a broad range of cfDNA concentration between 6.7 ng/ml and 79.9 ng/ml plasma (mean = 27.9 ng/ml, median = 23 ng/ml) across samples (**Table 1**). However, we found no significant difference in cfDNA concentrations between flight, ground or control subjects (ANOVA $p = 0.49$, **Supp. Fig. 1A**), TW and HR (Wilcoxon rank test $p = 0.65$), and flight and ground samples (Wilcoxon rank test $p = 0.352$). TW showed borderline significantly higher cfDNA concentration pre- and post-flight compared to in-flight (Wilcoxon rank test $p = 0.043$), however, this is not significant when comparing TW in-flight, TW ground, and HR/MS ground (ANOVA $p = 0.4$, **Supp. Fig 1B**). Complementary metadata on the health status of TW and HW during the mission has been previously published (Garret-Bakelman et al., 2019), and no deviations in medication or exercise regimen were noted in the medical records.

Cf-DNA fragment length distribution is influenced by the ambient return

It has previously been shown that cfDNA derived from tumor cells is shorter than cfDNA derived from healthy cells (Jiang et al., 2015; Mouliere et al., 2011). This effect can be explained by a change in nucleosome binding or by a degradation of nucleotides at the end of nucleosome loops. We therefore hypothesized that environmental stressors such as microgravity or radiation could also impact the length distribution of cfDNA. Indeed, we found a slight shift to longer cfDNA fragment lengths in TW in-flight samples (**Fig. 1A**). However, a similar shift was observed in ground samples subjected to ambient return simulation (**Fig. 1A**, boxplots with yellow border). Ambient return samples show a similar peak at the 300 to 400bp fragment length, which is only marginally visible for fresh samples (**Fig. 1B**). Thus, some proportion of long cfDNA fragments likely originate from blood cells damaged during return flight or transport from the ISS.

To examine how this might affect other cfDNA fractions, we next examined cell-free mitochondrial DNA (cf-mtDNA). Recent studies indicate that a prominent fraction of cf-mtDNA in the plasma is contained within intact, circulating mitochondria (Al Amir Dache, 2020) and that larger mtDNA fragments can also arise from blood cell degradation. However, our centrifugation

step largely removed intact mitochondria and our library preparation comprised mostly smaller DNA fragments (at least 75% are <350bp)(Supplemental Fig. 2), including an even smaller fraction (<10%) of the aligned reads (Fig. 2). Thus, the observed fractions of cf-mtDNA are mostly derived from shorter cf-mtDNA molecules and should represent cf-mtDNA that is randomly fragmented and sequenced across the entire mitochondria.

As further evidence of this, the mitochondrial genome showed continuous read coverage in all samples, ranging from 50x-200x coverage (Supplemental Fig. 3), regardless of the collection method. Indeed, the length distribution of cf-mtDNA is not affected by ambient return as observed for chromosomal cfDNA (Fig. 2B), and the average length does not change significantly in inflight samples or AR simulation samples. Even though cf-mtDNA amounts can significantly vary based on the donor profiles (Lindqvist et al, 2016) and degree intact vs. fragmented mitochondria, these NGS data showed that the total cf-mtDNA profiles show relative uniformity in both length and proportion of reads (Fig. 2).

Levels of cell-free mitochondrial DNA are increased during space flight

Next we investigated the fraction of cf-mtDNA relative to chromosomal cfDNA in plasma of TW, HR, and MS. In order to characterize the cfDNA originating from mitochondria during spaceflight, we normalized the count of NGS reads mapping to the mitochondrial chromosome (chrM) by chromosome length and the total number of reads in the library, generating a RPKM measurement. For comparison, we performed the same procedure with reads mapping to chromosome 21. We found a sharp increase of cf-mtDNA for subject TW for inflight samples (Fig. 3A) compared to TW ground samples (Wilcoxon rank test $p=0.012$), compared to HR ground samples (Wilcoxon rank test $p=0.018$), and compared to all ground samples of HR and TW (Wilcoxon rank test $p=0.0045$, ANOVA $p=0.00049$). In contrast, we found no significant increase in cfDNA mapping to chromosome 21 (Fig. 3B) in TW-inflight compared to ground samples of TW and HR.

Notably, the mtDNA levels in whole blood increased steadily inflight while on the ISS. Indeed, TW had the highest fraction of cf-mtDNA within the first inflight timepoint (T4), including more than a 24-fold increase, when compared to ground samples (Fig. 3C, 3D). In the two later inflight time points, he had 4- and 8-fold increases compared to pre-flight levels. The normalized levels of chromosome 21 cfDNA were stable for both TW and HR for the duration of the mission (0.25-0.26 RPKM), revealing no obvious bias due to sample handling (Fig. 3D). Interestingly, a positive correlation between mtDNA copy number and telomere length in healthy adults has been previously reported, and telomere elongation in blood and urine was also observed during spaceflight for TW (Garrett-Bakelman et al., 2019, Luxton et al, 2020).

Given the previously discussed effects of AR on cfDNA lengths, we tested for potential bias in cf-mtDNA levels due to AR. To do this, we compared the cf-mtDNA fraction observed in the MS simulated-AR samples (2 samples) to the MS control sample. We found that cf-mtDNA levels were actually lower in AR than in FR samples (Fig. 3E), suggesting that the shipping procedure from the ISS is likely not causing the observed increase in cf-mtDNA levels seen in the inflight samples. In addition, the AR simulation of the ground subject (HR) did not show a significant increase of cf-mtDNA levels compared to other HR samples (Fig. 3F). Thus, these

data suggest that the cf-mtDNA fraction was significantly increased during space flight, and not due to the AR blood-from-ISS transport process.

Nucleosome positioning suggests a shift in cell of origin of cfDNA due to transport conditions

Given that nucleosome positions are associated with both cfDNA and gene expression (Jiang and Pugh, 2009), we computed the nucleosome depletion around nucleosomes at transcription start sites (TSS) to infer gene expression (**Fig. 4A**), as previously demonstrated by Ulz and colleagues (Ulz et al., 2016). Indeed, these data indicated that the strength of nucleosome depletion is correlated to bulk gene expression from RNA-seq of the same subjects (Garrett-Bakelman et al., 2019) (**Fig. 4A**), with a decreased coverage at the site of the transcriptional start site (TSS) for highly expressed genes. Second, we identified the nucleosome footprint of CTCF in gene bodies, hypothesizing that nucleosome positioning patterns could reveal broad changes in gene regulation during spaceflight. A t-SNE analysis of TW and HR samples showed no flight-specific clustering (**Fig 4B**), indicating that nucleosome positioning identified through cfDNA may not be sensitive enough to identify spaceflight-related gene expression changes.

However, based on the correlations between per-tissue gene expression values (Kim et al., 2014) and nucleosome positioning observed on cfDNA, clear tissue signals in cfDNA were inferred for all plasma samples. Higher values (Pearson's correlation coefficient) suggested higher gene expression and stronger tissue signal (**Fig. 5A**) for hematopoietic lineages (up to $\rho = 0.156$, $n = 1087411335$), mid-range for liver, adrenal gland, and the retina (0.04-0.07) and less so for other peripheral tissues (e.g. lung, esophagus, 0.00-0.01). These results are consistent with the expected cfDNA prevalence in blood and with previous findings (Snyder et al., 2016). Despite such clear signals on tissue of origin, strong clustering of samples was observed, due to the confounding effect of ambient return. This was seen in both the tissue-of-origin analysis (**Fig. 5A**) as well as TSS protection (**Fig. 5B**), highlighting the need for controls and correction for any degradation. Further, this analysis does not take into account the cf-mtDNA reads, and therefore may not reflect the tissue of origin for mitochondrial reads or heteroplasmy.

Analysis of plasma-circulating exosomes post-flight

To determine how prolonged space missions and Earth re-entry impact circulating exosomes, we analyzed exosomes from the plasma of TW three years post-return to Earth, and compared their size, number and proteomes to plasma-derived exosomes isolated from HR and 6 age-matched, healthy controls. Exosomes were isolated by differential ultracentrifugation and both the size and number of exosomes were characterized by nanoparticle tracking analysis (NTA) (**Fig. 6A-E**). While the median size of exosomes was similar between HR, TW and healthy controls (**Fig. 6A-D**), the number of particles was ~30 times higher in TW compared to HR and healthy controls (**Fig. 6 E**). Proteomic mass spectrometry analysis revealed that TW, HR and control exosomes packaged similar numbers of proteins, including a total of 191 exosomal proteins shared among all samples. HR's exosome catalog contained 26 unique proteins, TW exosomes contained 61 unique proteins, and healthy controls contained 105 unique proteins (**Fig. 6F**).

Hierarchical clustering of the exosomal proteins revealed distinct signatures of HR and TW, which clustered apart from the six controls. Interestingly, classification of the pathways using Metascape (GO processes, KEGG pathways, Reactome gene sets, canonical pathways, and CORUM complexes)(Zhou et al., 2019) revealed that TW exosomes were enriched in proteins

involved in proteasome pathways (**Fig. 6H**). TW exosomes also packaged CD14, a pro-inflammatory monocyte marker, consistent with the increase in CD14⁺ monocytes observed post-return to gravity in immune markers studied upon return to Earth (Gertz et al, 2020). Notably, basigin and integrin $\beta 1$ proteins, which are correlated with cancer progression and inflammation (Hoshino et al., 2015, 2020; Keller et al., 2009; Yoshioka et al., 2014), were also detected in TW exosomes, but not in HR or healthy control exosomes.

Consistent with previous findings demonstrating microgravity downregulating adaptive immunity, particularly B cells (Cao et al., 2019), both TW and HR exosomes contained fewer immunoglobulins compared to healthy controls (**Fig. 6G**). Surprisingly, two brain-specific proteins, Brain-specific angiogenesis inhibitor 1-associated protein 2 (BAIAP2) and Brain-specific angiogenesis inhibitor 1-associated protein 2-like protein 1 (BAIAP2L1), were found in TW plasma-derived exosomes (**Supplemental Table 1, Supplemental Fig. 4A**), yet were not detected in the plasma of HR or healthy controls. In contrast, HR exosomal cargo was enriched in proteins associated with regulation of apoptotic pathways (Théry et al., 2001) and ATP biosynthesis (**Fig. 6I**). Moreover, we observed that the 20S proteasome, but not the regulatory 19S proteasome, is found uniquely associated with the plasma-circulating exosomes in the flight subject (TW) 3 years after his return to Earth (**Fig. 6G**). Finally, both TW and HR exosomes, but not controls, were enriched in specific components of the humoral immune response and leukocyte migration, including the CD53 tetraspanin (**Supplemental Table 2, Supplemental Fig. 4B**) which could reflect either biology shared by the twins or changes associated with travel to space; however, analysis of plasma exosome samples from genetically unrelated astronauts would be required to distinguish between these possibilities.

Discussion

Our study focused on cfDNA and exosomes collected during the NASA Twins study, a longitudinal, multi-omic experiment examining the effects of long-term spaceflight on the human body. In particular, we revealed cf-mtDNA fraction to be a potential new biomarker of physiological stress during prolonged spaceflight, though the total cfDNA concentration is not significantly correlated with spaceflight. We further observed unique exosome and exosomal protein signatures within TW several years after the year-long mission, including an increased amount of exosomes and brain-specific proteins (BAIAP2 and BAIAP2L1). Of note, we identified multiple biases likely caused by ambient return (AR) blood draws from the ISS, including results of tissue of origin deconvolution through nucleosome positioning as well as cfDNA fragment length. As such, future studies will need to control for AR affects if they wish to examine these molecular dynamics. As an example, DNA could be extracted in space (Castro-Wallace *et al.*, 2017) and either cryopreserved to increase its stable during transport or directly sequencing inflight to minimize biases and obtain results faster (McIntyre *et al.*, 2016, McIntyre *et al.*, 2019).

Interestingly, analysis of plasma exosomes isolated post-return to Earth revealed unique alterations in TW relative to HR and healthy controls, such as a dramatic increase in the number of circulating particles as well as changes in the types of protein cargo. Since the majority of plasma circulating exosomes are derived from immune cells, it is likely that these alterations reflect immune dysfunction associated with space travel and return to gravity. Specifically, the reduction in TW exosomal immunoglobulin levels and the presence of CD14, a macrophage marker, may signal a shift towards innate immunity, as even short-term chronic exposure to

cosmic radiation and microgravity leads to a decrease in adaptive immune cells (Cao et al., 2019; Fernandez-Gonzalo et al., 2017). However, circulating exosomes also reflect systemic changes in homeostasis and physiology, as demonstrated by the packaging of brain-specific proteins in TW which were not seen in control exosomes, which may indicate long-term altered expression of exosomes from the brain after spaceflight. Previous studies had shown that microgravity affects tight junction protein localization within intestinal epithelial cells (Alvarez et al., 2019). It is conceivable that prolonged space travel could exert similar effects on tight junctions within the blood-brain barrier, allowing for more exosomes to enter the peripheral blood.

One remarkable finding of our study is that the 20S proteasome, but not the regulatory 19S proteasome, is found uniquely associated with the plasma-circulating exosomes in the flight subject 3 years post his return to Earth. Recent research has discovered the ubiquitin-independent proteolytic activity of the 20S proteasome and its role as the major degradation machinery under oxidizing conditions (Aiken et al., 2011; Deshmukh et al., 2019; Pickering and Davies, 2012). Elevated levels of 20S proteasome have been detected in the blood plasma from patients with various blood cancers, solid tumors, autoimmune diseases and other non-malignant diseases (Deshmukh et al., 2019; Sixt and Dahlmann, 2008). It is also reported that active 20S proteasomes within apoptotic exosome-like vesicles can induce autoantibody production and accelerate organ rejection after transplantation (Dieudé et al., 2015), reduce the amount of oligomerized proteins (Schmidt et al., 2020) and reduce tissue damage after myocardial injury (Lai et al., 2012), and are correlated with cancer and other pathological status such as viral infection and vascular injury (Dieudé et al., 2015; Gunasekaran et al., 2020; Tugutova et al., 2019). The elevated circulating exosomal 20S proteasome in the flight subject may reflect the increased physiological need to clear these proteins resulting from long-term blood, immune or other physiological disorders caused by various stress factors during the flight or return to gravity (Ben-Nissan and Sharon, 2014; Vernice et al., 2020). Study of plasma exosomes obtained from flight subjects at other time points including pre- and inflight will be necessary to further examine whether plasma exosomal proteasome can serve as biomarker for pathological processes associated with space flight.

There are limitations in the study design that prevent broad biological conclusions. First, the sample number is too small to control for all types of potential biases and results may be somewhat driven by individual health issues. Second, there is no comparable experimental data to date and the effect of return to gravity in the Soyuz capsule on the integrity of the sampled material is unknown. Third, the exosome samples have been taken post-flight and can only inform about long-term effects of extended spaceflight. However, this study stands as a demonstration of the applications and possibilities of utilizing cfDNA and exosome profiling to monitor astronaut health and can improve the study design of future missions and research (Iosim et al., 2019; Nangle et al., 2020).

In summary, we identified cell-free mitochondrial DNA (cf-mtDNA) as a novel biomarker of physiological stress during prolonged spaceflight, which is stable even during transport from the ISS. However, we demonstrated that transport-induced biases for cell-type deconvolution from cfDNA needs to be improved in order to be used as a “molecular whole body scan”, and/or deployment of more real-time methods (e.g. inflight sequencing). Also, we observed that exosome concentration in plasma and unique exosomal proteins such as 20S proteasomes, CD14,

and BAIAP2 demonstrate characteristic changes in the flight subject (TW), potentially caused by physiological stress during prolonged spaceflight. Overall, these data and methods provide novel metrics and data types that can be used in planning for future types of astronaut health monitoring, as well as help establish non-invasive molecular tools for tracking the impact of stress and spaceflight during future missions.

Acknowledgements

We would like to thank the Epigenomics Core Facility and the Scientific Computing Unit (SCU) at Weill Cornell Medicine, as well as the Starr Cancer Consortium (I9-A9-071) and funding from the Irma T. Hirsch and Monique Weill-Caulier Charitable Trusts, Bert L and N Kuggie Vallee Foundation, the WorldQuant Foundation, The Pershing Square Sohn Cancer Research Alliance, NASA (NNX14AH51G (all Twins Study principal investigators); NNX14AB01G (S.M.B.); and NNX17AB26G (C.E.M.), NNX14AH52G), the National Institutes of Health (R25EB020393, R01NS076465, R01AI125416, R01ES021006, R01AI151059, 1R21AI129851, 1R01MH117406), TRISH (NNX16AO69A:0107, NNX16AO69A:0061, NIH/NCATS KL2-TR-002385), the Bill and Melinda Gates Foundation (OPP1151054), the Leukemia and Lymphoma Society (LLS) grants (LLS 9238-16, Mak, LLS-MCL-982, Chen-Kiang).

Disclosure Statement

S.M.B. is a cofounder and Scientific Advisory Board member of KromaTiD, Inc., CEM is a cofounder and board member for Biotia, Inc. and Onegevity Health, Inc., as well as an advisor or grantee for Abbvie, Inc., ArcBio, Daiichi Sankyo, DNA Genotek, Karius, Inc., and Whole Biome, Inc. DB is a cofounder of Poppy Health, Inc. and Analog Llc.

Author Contributions

CEM, DB and DCL conceived the study, DB, CM, EA, SO, FAV, HZ, IM, KG and PB and, wrote the manuscript DB,BS, FEG, DBU, DPK, KHY, KN, TL, VR, FAV sample collection and/or processing DB, CM, SO, HZ, IM, JF, KG and PB Bioinformatic and Analytics, AM, BS, CEM, CM, CW, EA, FEG, IV, MC, MPS, RKB, SL, SO and TM review manuscript and guided interpretation. All authors read and approved the manuscript.

Methods

Sample collection

In the NASA Twin Study spanning 24 months we collected blood samples at 12 time points from the twin on earth (HR) and 11 time points from the twin in space (TW), as previously described (Garrett-Bakelman et al., 2019). From TW, samples were collected before the flight (PRE-FLIGHT), during the flight (FLIGHT) and after the flight (POST-FLIGHT). Specimens were processed as previously described (Garrett-Bakelman et al., 2019). Briefly, whole blood was collected in 4mL CPT vacutainers (BD Biosciences Cat # 362760,) per manufacturer's recommendations, which contained 0.1M sodium citrate, a thixotropic polyester gel and a FICOLL Hypaque solution. Hence our specimens were not exposed to heparin. Samples were mixed by inversion. Samples collected on ISS were stored at 4°C after processing and returned by the Soyuz capsule. There was an average of 35-37 hours from collection to processing, including repatriation

time. Plasma was obtained by centrifugation of the CPT vacutainers at 1800 X g for 20 minutes at room temperature, both for the ISS and for the ground-based samples. Finally, plasma was collected from the top layer in the CPT vacutainer and flash frozen prior to long term storage at -80°C.

To simulate batch effects between fresh material (samples collected on earth) and ambient return material (samples collected during flight and returned via Soyuz capsule at 4°C), we generated 3 control samples (MS) representing fresh (FR) and ambient return (AR) material as described before (Garrett-Bakelman et al., 2019). Whole blood of a male volunteer of similar age and ethnicity as HR/TW was collected in three CPT tubes. Plasma was collected and stored as described for the TW and HR specimens. To generate the ambient return control (AR) two CPT vacutainers were shipped at ambient temperature (4°C) from Stanford University to Weill Cornell Medicine and back as air cargo. The returned CPT vacutainers were spun at 300 X g for 3 minutes and aliquoted. One aliquot from each tube was spun once more at 1800 X g for 3 minutes to completely clear the plasma of cell debris, resulting in the final AR controls. The aliquoted plasma was stored at -80°C.

cfDNA extraction and comparison of cfDNA concentrations in ground and flight subjects

Between 250ul and 1 ml plasma was retrieved from HR, TW and MS samples. The frozen plasma was thawed at 37°C for 5min and spun at 16000g for 10 minutes at 4°C to remove cryo-precipitates. The volume of each plasma sample was brought up to 1ml using sterile, nuclease-free 1X phosphate buffered saline pH 7.4. Circulating cell-free nucleic acid (ccfNA) was extracted using the Qiaamp Circulating Nucleic Acid kit (Qiagen, USA) following the manufacturer's protocol. ccfNA was extracted in 50ul AE buffer. Concentration and size distribution information was obtained by running 1ul of ccfNA on the Agilent Bioanalyzer using the High Sensitivity DNA chip (Agilent technologies, CA, USA). ~15ul aliquots were set aside for cell-free DNA or DNA methylation analyses and stored at -80°C. A range of extracted cfDNA of 1ng-38ng/mL plasma has been reported for healthy donors, while cancer patients often show higher levels of 30-50ng/ml (Table 1). In the HR, TW and control samples we extracted between 6.7 ng/ml and 79.9 ng/ml plasma (mean = 27.9 ng/ml, median = 23 ng/ml)(Table 1). We tested if there is a significant difference between HR, TW or MS as well as FR and AR samples using Wilcoxon rank test (R function wilcox) for pairwise comparisons and ANOVA (R function anova) for multi-group comparisons. We furthermore visualized the distributions of the groups HR, TW and MS as boxplots using the R package ggplot2.

Q-PCR Analysis of cfDNA

The frozen plasma was thawed at 37°C for 5min and spun at 16000g for 10 minutes at 4°C to remove cryo-precipitates. DNA level in samples was measured by SYBR Green dye-based qPCR assay using a PRISM 7300 sequence detection system (Applied Biosystems) as described previously (Nakahira PLoS Med. 2013, Garrett-Bakelman Science 2019, PMIDs: 24391478 and 30975860). The primer sequences were as follows: human NADH dehydrogenase 1 gene (hu mtNd1): forward 5'-ATACCCATGGCCAACCTCCT-3', reverse 5'-GGGCCTTTGCGTAGTTGTAT-3'. Plasmid DNA with complementary DNA sequences for human mtDNA was obtained from ORIGENE (SC101172). Concentrations were converted to copy number using the formula; mol/gram × molecules/mol = molecules/gram, via a DNA copy number calculator (<http://cels.uri.edu/gsc/cndna.html>; University of Rhode Island Genomics and Sequencing Center).

The thermal profile for detecting mtDNA was carried out as follows: an initiation step for 2 min at 50°C is followed by a first denaturation step for 10 min at 95°C and a further step consisting of 40 cycles for 15 s at 95°C and for 1 min at 60°C. MtDNA levels in all of the plasma analyses were expressed in copies per microliter of plasma based on the following calculation: $c = Q \times \text{VDNA} / \text{VPCR} \times 1 / \text{Vext}$; where c is the concentration of DNA in plasma (copies/microliter plasma); Q is the quantity (copies) of DNA determined by the sequence detector in a PCR; VDNA is the total volume of plasma DNA solution obtained after extraction; VPCR is the volume of plasma DNA solution used for PCR; and Vext is the volume of plasma extracted.

Library generation and sequencing

DNA libraries were generated using the NEBNext DNA Library Preparation Kit Ultra II (New England Biolabs, USA). Libraries were generated using 15ul of the ccfNA according to the manufacturer's instruction. Following end-repair and dA-tailing, adaptor ligation was performed using 15-fold diluted adaptors. After removal of free adaptors using Agencourt magnetic beads (Beckman Coulter, USA), the libraries were PCR-amplified for 12 cycles using primers compatible Illumina dual-index sequences. Following bead cleanup for primer removal, the libraries were run on the Agilent Bioanalyzer to estimate size and concentration. All libraries were pooled at equal concentration and sent to New England Biolabs, Ipswich MA for sequencing. Preliminary sequencing on Illumina Miseq indicated the presence of adaptor dimers in some of the libraries. Therefore, the individual libraries were subjected to an additional round of bead purification and size and concentration estimation. Subsequently, all libraries were pooled again and sequenced on the NovaSeq 6000 using an S2 flow cell and 200-cycle kits (2x100). We finally obtained 4.9 and 4.1 billion reads passing quality filters.

cfDNA sequence analysis

Samples were de-multiplexed using the standard Illumina tools. Low quality bases were trimmed and Illumina-specific sequences and low quality sequences were removed from the sequencing data using Trimmomatic-0.32. Filtered, paired-end reads were aligned using BWA-mem to the hg38 human reference genome with the bwa-postalt option to handle alternative alignments. The resulting BAM files were post-processed (e.g. sorted) using samtools. Duplicate sequences were removed and only reads aligning in concordant pairs were used for further analysis. The fragment length distribution was generated by plotting the distance between read 1 and 2 obtained from the BAM file of each sample. Histograms and boxplots of the fragment length distribution for the autosomes (Figure 1) and for the mitochondrial genome (Figure 2) for all samples were generated using the R package ggplot2.

Analysis of cell free mitochondrial DNA

cfDNA read counts by chromosome (including the mitochondrial genome labeled ChrMT) were extracted using a 'edtools coverage -a feature_file -b sample.bam -counts', where feature_file contains the definition (name, start, end) of all chromosomes. Read counts per feature were length-normalized using the well-established reads-per-kilobase per million formula frequently applied to normalize RNA-seq data (Mortazavi et al., 2008). We used the R package ggplot 2 to visualize differences in the normalized cfDNA fraction (RPKM) originating from the mitochondrial genome between HR, TW and MS and over time during the mission (longitudinal analysis). We used Wilcoxon rank test (R function Wilcox) to test if the measurements of two conditions (e.g. TW on

ground vs. TW in flight) are significantly different. To analyze the differences among multiple groups we applied ANOVA (R function anova).

Nucleosome positioning analysis

Filtered, paired-end reads were aligned using BWA-mem to the hg37 human reference genome and post-processed using samtools: duplicate sequences were removed, and only reads aligning in concordant pairs were used for final analysis. The sequence read coverage in 10-kbp windows (-5 kbp to 5kbp) around the transcription start sites of all genes was determined using the samtools depth function. From the positions of the reads, nucleosome occupancy was inferred, and its periodograms calculated. A list of transcription start sites organized by transcriptional activity (measured in FPKM) was used to assign activity (Ulz et al., 2016). The depth of coverage was summed across genes according to transcriptional activity category (as depicted in Figure 4A). The coverage was normalized by subtracting the mean value from the intervals [TSS-3 kbp, TSS-1 kbp] and [TSS+1 kbp, TSS+3 kbp]. According to the method in Snyder et al., 2015, FFT values for the periods of 193-199bp were correlated with the gene level expression matrix, and the resulting tissue-periodicity correlations ranked by the value of Pearson's correlation coefficient and clustered (Ward method with Euclidean distances) to investigate characteristics of tissue-of-origin dependent on sample type.

Purification and Mass spectrometry analysis of plasma-circulating exosomes

Blood plasma was collected from TW 3 years post return of TW to Earth. Blood was also collected from HR (within one day of blood collection from TW) and from six age-matched healthy controls. Exosomes were purified by sequential ultracentrifugation, as previously described (Hoshino et al., 2015). Plasma samples were centrifuged for 10 minutes at 500xg, 20 minutes at 3,000xg, 20 minutes at 12,000xg, and the supernatant was collected and stored at -80°C for exosome isolation and characterization by NTA (NanoSight NS500, Malvern Instruments, equipped with a violet laser (405 nm). Samples were thawed on ice and centrifuged at 12,000xg for 20 min to remove large microvesicles. Exosomes were collected by spinning at 100,000xg for 70min, washed in PBS and pelleted again by ultracentrifugation in a 50.2 Ti rotor, Beckman Coulter Optima XE or XPE ultracentrifuge. The final exosome pellet was resuspended in PBS, and protein concentration was measured by BCA (Pierce, Thermo Fisher Scientific). Mass spectrometry analyses of exosomes were performed at the Rockefeller University Proteomics Resource Center using 10 µg of exosomal protein as described previously (Hoshino et al., 2015; Zhang et al., 2018). Heatmap and complete Euclidean clustering was performed with Morpheus, (<https://software.broadinstitute.org/morpheus>). Pathway analysis was performed with Metascape (Zhou et al., 2019).

References

- Aiken, C.T., Kaake, R.M., Wang, X., and Huang, L. (2011). Oxidative Stress-Mediated Regulation of Proteasome Complexes. *Mol. Cell. Proteomics* 10, R110.006924.
- Amir Dache, Al, Z., Otandault, A., Tanos, R., Pastor, B., Meddeb, R., Sanchez, C., et al. (2020). Blood contains circulating cell-free respiratory competent mitochondria. *The FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, 34(3), 3616–3630.
- Alvarez, R., Stork, C.A., Sayoc-Becerra, A., Marchelletta, R.R., Prisk, G.K., and McCole, D.F. (2019). A Simulated Microgravity Environment Causes a Sustained Defect in Epithelial Barrier Function. *Sci. Rep.* 9.
- Ben-Nissan, G., and Sharon, M. (2014). Regulating the 20S proteasome ubiquitin-independent degradation pathway. *Biomolecules* 4, 862–884.
- Berezin, A.E. (2016). The Cell-Free Mitochondrial DNA: A Novel Biomarker of Cardiovascular Risk? *Transl. Biomed.* 7.
- Bettgowda, C., Sausen, M., Leary, R.J., Kinde, I., Wang, Y., Agrawal, N., Bartlett, B.R., Wang, H., Lubner, B., Alani, R.M., et al. (2014). Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies. *Sci. Transl. Med.* 6, 224ra24–224ra24.
- Bianchi, D.W., Parker, R.L., Wentworth, J., Madankumar, R., Saffer, C., Das, A.F., Craig, J.A., Chudova, D.I., Devers, P.L., Jones, K.W., et al. (2014). DNA sequencing versus standard prenatal aneuploidy screening. *N Engl J Med* 370, 799–808.
- Birkenkamp-Demtröder, K., Nordentoft, I., Christensen, E., Høyer, S., Reinert, T., Vang, S., Borre, M., Agerbæk, M., Jensen, J.B., Ørntoft, T.F., et al. (2016). Genomic Alterations in Liquid Biopsies from Patients with Bladder Cancer. *Eur. Urol.*
- Blauwkamp, T.A., Thair, S., Rosen, M.J., Blair, L., Lindner, M.S., Vilfan, I.D., Kawli, T., Christians, F.C., Venkatasubrahmanyam, S., Wall, G.D., et al. (2019). Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. *Nat. Microbiol.* 4, 663–674.
- Burnham, P., Dadhania, D., Heyang, M., Chen, F., Westblade, L.F., Suthanthiran, M., Lee, J.R., De Vlaminc, I., and Vlaminc, I. De (2018). Urinary cell-free DNA is a versatile analyte for monitoring infections of the urinary tract. *Nat. Commun.* 9, 2412.
- Cao, D., Song, J., Ling, S., Niu, S., Lu, L., Cui, Z., Li, Y., Hao, S., Zhong, G., Qi, Z., et al. (2019). Hematopoietic stem cells and lineage cells undergo dynamic alterations under microgravity and recovery conditions. *FASEB J.* 33, 6904–6918.
- Castro-Wallace SL, Chiu CY, John KK, Stahl SE, Rubins KH, McIntyre ABR, Dworkin JP, Lupisella ML, Smith DJ, Botkin DJ, Stephenson TA, Juul S, Turner D, Izquierdo F, Federman S, Stryke D, Somasekar S, Alexander N, Yu G, Mason CE, Aaron S Burton. “Nanopore DNA Sequencing and Genome Assembly on the International Space Station.” *Scientific Data*. 2017 Dec 21;7(1):18022.
- Cheng, A.P., Burnham, P., Lee, J.R., Cheng, M.P., Suthanthiran, M., Dadhania, D., and De Vlaminc, I. (2019). A cell-free DNA metagenomic sequencing assay that integrates the host injury response to infection. *Proc. Natl. Acad. Sci. U. S. A.* 116, 18738–18744.
- Colombo, M., Raposo, G., and Théry, C. (2014). Biogenesis, Secretion, and Intercellular Interactions of Exosomes and Other Extracellular Vesicles. *Annu. Rev. Cell Dev. Biol.* 30, 255–289.
- Deshmukh, F.K., Yaffe, D., Olshina, M.A., Ben-Nissan, G., and Sharon, M. (2019). The contribution of the 20s proteasome to proteostasis. *Biomolecules* 9.
- Diehl, F., Schmidt, K., Choti, M.A., Romans, K., Goodman, S., Li, M., Thornton, K., Agrawal, N., Sokoll, L., Szabo, S.A., et al. (2008). Circulating mutant DNA to assess tumor dynamics. *Nat. Med.* 14, 985–990.
- Dieudé, M., Bell, C., Turgeon, J., Beillevaire, D., Pomerleau, L., Yang, B., Hamelin, K., Qi, S., Pallet, N., Béland, C., et al. (2015). The 20S proteasome core, active within apoptotic exosome-like vesicles, induces autoantibody production and accelerates rejection. *Sci. Transl. Med.* 7.
- Fernandez-Gonzalo, R., Baatout, S., and Moreels, M. (2017). Impact of particle irradiation on the immune system: From the clinic to mars. *Front. Immunol.* 8.

Garrett-Bakelman, F.E., Darshi, M., Green, S.J., Gur, R.C., Lin, L., Macias, B.R., McKenna, M.J., Meydan, C., Mishra, T., Nasrini, J., et al. (2019). The NASA Twins Study: A multidimensional analysis of a year-long human spaceflight. *Science* 364.

Gunasekaran, M., Bansal, S., Ravichandran, R., Sharma, M., Perincheri, S., Rodriguez, F., Hachem, R., Fisher, C.E., Limaye, A.P., Omar, A., et al. (2020). Respiratory viral infection in lung transplantation induces exosomes that trigger chronic rejection. *J. Hear. Lung Transplant.* 39, 379–388.

Harmati, M., Gyukity-Sebestyen, E., Dobra, G., Janovak, L., Dekany, I., Saydam, O., Hunyadi-Gulyas, E., Nagy, I., Farkas, A., Pankotai, T., et al. (2019). Small extracellular vesicles convey the stress-induced adaptive responses of melanoma cells. *Sci. Rep.* 9.

Heitzer, E., Haque, I.S., Roberts, C.E.S., and Speicher, M.R. (2018). Current and future perspectives of liquid biopsies in genomics-driven oncology. *Nat. Rev. Genet.* 20, 1.

Hoshino, A., Costa-Silva, B., Shen, T.L., Rodrigues, G., Hashimoto, A., Tesic Mark, M., Molina, H., Kohsaka, S., Di Giannatale, A., Ceder, S., et al. (2015). Tumour exosome integrins determine organotropic metastasis. *Nature* 527, 329–335.

Hoshino, A., Kim, H.S., Bojmar, L., Gyan, K.E., Cioffi, M., Hernandez, J., Zambirinis, C.P., Rodrigues, G., Molina, H., Heissel, S., et al. (2020). Extracellular Vesicle and Particle Biomarkers Define Multiple Human Cancers. *Cell* 182, 1044-1061.e18.

Hummel, E.M., Hessas, E., Müller, S., Beiter, T., Fisch, M., Eibl, A., Wolf, O.T., Giebel, B., Platen, P., Kumsta, R., et al. (2018). Cell-free DNA release under psychosocial and physical stress conditions. *Transl. Psychiatry* 8, 236.

Iosim S, MacKay M, Westover C, Mason CE. Translating current biomedical therapies for long duration, deep space missions. *Precision Clinical Medicine*. 2019 Dec;2(4):259-269.

Jiang, C., and Pugh, B.F. (2009). Nucleosome positioning and gene regulation: Advances through genomics. *Nat. Rev. Genet.* 10, 161–172.

Jiang, P., Chan, C.W.M., Chan, K.C.A., Cheng, S.H., Wong, J., Wong, V.W.-S., Wong, G.L.H., Chan, S.L., Mok, T.S.K., Chan, H.L.Y., et al. (2015). Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. *Proc. Natl. Acad. Sci.* 112, E1317--E1325.

Kalluri, R., and LeBleu, V.S. (2020). The biology, function, and biomedical applications of exosomes. *Science* (80-.). 367.

Keller, S., König, A.K., Marmé, F., Runz, S., Wolterink, S., Koensgen, D., Mustea, A., Schouli, J., and Altevogt, P. (2009). Systemic presence and tumor-growth promoting effect of ovarian carcinoma released exosomes. *Cancer Lett.* 278, 73–81.

Kim, M.-S.S., Pinto, S.M., Getnet, D., Nirujogi, R.S., Manda, S.S., Chaerkady, R., Madugundu, A.K., Kelkar, D.S., Isserlin, R., Jain, S., et al. (2014). A draft map of the human proteome. *Nature* 509, 575.

Lai, R.C., Tan, S.S., Teh, B.J., Sze, S.K., Arslan, F., de Kleijn, D.P., Choo, A., and Lim, S.K. (2012). Proteolytic Potential of the MSC Exosome Proteome: Implications for an Exosome-Mediated Delivery of Therapeutic Proteasome. *Int. J. Proteomics*.

Li, L., Hann, H.-W., Wan, S., Hann, R.S., Wang, C., Lai, Y., Ye, X., Evans, A., Myers, R.E., Ye, Z., et al. (2016). Cell-free circulating mitochondrial DNA content and risk of hepatocellular carcinoma in patients with chronic HBV infection. *Sci. Rep.* 6, 23992.

Lindqvist, D., Fernström, J., Grudet, C., Ljunggren, L., Träskman-Bendz, L., Ohlsson, L., and Westrin, Å. (2016). Increased plasma levels of circulating cell-free mitochondrial DNA in suicide attempters: associations with HPA-axis hyperactivity. *Transl. Psychiatry* 6, e971.

Lindqvist, D., Wolkowitz, O.M., Picard, M., Ohlsson, L., Bersani, F.S., Fernström, J., Westrin, Å., Hough, C.M., Lin, J., Reus, V.I., et al. (2018). Circulating cell-free mitochondrial DNA, but not leukocyte mitochondrial DNA copy number, is elevated in major depressive disorder. *Neuropsychopharmacology* 43, 1557–1564.

Malakhova, L., Bezlepkin, V.G., Antipova, V., Ushakova, T., Fomenko, L., Sirota, N., and Gaziev, A.I. (2005). The increase in mitochondrial DNA copy number in the tissues of Γ -irradiated mice. *Cell. Mol. Biol. Lett.*

Mathieu, M., Martin-Jaular, L., Lavieu, G., and Théry, C. (2019). Specificities of secretion and uptake of exosomes and other extracellular vesicles for cell-to-cell communication. *Nat. Cell Biol.* 21, 9–17.

McIntyre ABR, Rizzardi L, Yu AM, Alexander N, Rosen GL, Botkin DJ, Stahl SS, John KK, Castro-Wallace SL, McGrath K, Burton AS, Feinberg AP, Mason CE. “Nanopore Sequencing in Microgravity.” *Nature Partner Journals (npj) Microgravity.* 2, 2016:16035.

McIntyre ABR, Alexander N, Grigorev K, Bezdan D, Sichtig H, Chiu CY, Mason CE. Single-molecule sequencing detection of N6-methyladenine in microbial reference materials. *Nature Communications.* 2019 Feb 4;10(1):579.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, nmeth.1226.

Mouliere, F., Robert, B., Peyrotte, E., Del Rio, M., Ychou, M., Molina, F., Gongora, C., and Thierry, A.R. (2011). High fragmentation characterizes tumour-derived circulating DNA. *PLoS One.*

Murtaza, M., and Caldas, C. (2016). Nucleosome mapping in plasma DNA predicts cancer gene expression. *Nat. Genet.* 48, 1105–1106.

Nangle SN, Wolfson MY, Hartsough L, Ma N, Mason CE*, Merighi M, Nathan V, Silver PA, Simon M, Swett J, Thompson DB, Ziesack M. The Case for Biotechnology on Mars. *Nature Biotechnology.* 2020 Apr;38(4):401-407.

Newman, A.M., Lovejoy, A.F., Klass, D.M., Kurtz, D.M., Chabon, J.J., Scherer, F., Stehr, H., Liu, C., Bratman, S. V, Say, C., et al. (2016). Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat. Biotechnol.* 34, 547–555.

O’Neill, C.P., Gilligan, K.E., and Dwyer, R.M. (2019). Role of extracellular vesicles (EVs) in cell stress response and resistance to cancer therapy. *Cancers (Basel).* 11.

Pickering, A.M., and Davies, K.J.A. (2012). Degradation of damaged proteins: The main function of the 20S proteasome. In *Progress in Molecular Biology and Translational Science*, (Elsevier B.V.), pp. 227–248.

Qin, Y., Long, L., and Huang, Q. (2020). Extracellular vesicles in toxicological studies: key roles in communication between environmental stress and adverse outcomes. *J. Appl. Toxicol.*

Shen, S.Y., Singhanian, R., Fehring, G., Chakravarthy, A., Roehrl, M.H.A., Chadwick, D., Zuzarte, P.C., Borgida, A., Wang, T.T., Li, T., et al. (2018). Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature* 563, 579–583.

Schmidt MA, Iosim S, Schmidt CM, Afshinnekoo E, Mason CE. The NASA Twins Study: The Effect of One Year in Space on the Genome and Molecular Phenotype of Long-Chain Fatty Acid Desaturases and Elongases. *Lifestyle Genomics.* 2020. May 6. 1: 1-15.

Siravegna, G., Marsoni, S., Siena, S., and Bardelli, A. (2017). Integrating liquid biopsies into the management of cancer. *Nat. Rev. Clin. Oncol.* 14, 531.

Sixt, S.U., and Dahlmann, B. (2008). Extracellular, circulating proteasomes and ubiquitin - Incidence and relevance. *Biochim. Biophys. Acta - Mol. Basis Dis.* 1782, 817–823.

Song X, Hu W, Yu H, Wang H, Zhao Y, Korngold R, Zhao Y. Existence of Circulating Mitochondria in Human and Animal Peripheral Blood. *Int J Mol Sci.* 2020 Mar 19;21(6):2122

Snyder, M.W., Kircher, M., Hill, A.J., Daza, R.M., and Shendure, J. (2016). Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. *Cell* 164, 57–68.

Théry, C., Boussac, M., Véron, P., Ricciardi-Castagnoli, P., Raposo, G., Garin, J., and Amigorena, S. (2001). Proteomic Analysis of Dendritic Cell-Derived Exosomes: A Secreted Subcellular Compartment Distinct from Apoptotic Vesicles. *J. Immunol.* 166, 7309–7318.

Thierry, A.R., El Messaoudi, S., Gahan, P.B., Anker, P., and Stroun, M. (2016). Origins, structures, and functions of circulating DNA in oncology. *Cancer Metastasis Rev.* 35, 347–376.

Tugutova, E.A., Tamkovich, S.N., Patysheva, M.R., Afanas’ev, S.G., Tsydenova, A.A., Grigor’eva, A.E., Kolegova, E.S., Kondakova, I. V., and Yunusova, N. V. (2019). Relation between tetraspanin- associated and tetraspanin- non- associated exosomal proteases and metabolic syndrome in colorectal cancer patients. *Asian Pacific J. Cancer Prev.* 20, 809–815.

Ulz, P., Thallinger, G.G., Auer, M., Graf, R., Kashofer, K., Jahn, S.W., Abete, L., Pristauz, G., Petru, E., Geigl, J.B., et al. (2016). Inferring expressed genes by whole-genome sequencing of plasma DNA. *Nat Genet.*

Underhill, H.R., Kitzman, J.O., Hellwig, S., Welker, N.C., Daza, R., Baker, D.N., Gligorich, K.M., Rostomily, R.C., Bronner, M.P., and Shendure, J. (2016). Fragment Length of Circulating Tumor DNA. *PLoS Genet* 12, e1006162.

Valadi, H., Ekström, K., Bossios, A., Sjöstrand, M., Lee, J.J., and Lötvall, J.O. (2007). Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* 9, 654–659.

Verhoeven, J.G.H.P.H.P., Boer, K., Van Schaik, R.H.N.N., Manintveld, O.C., Huibers, M.M.H.H., Baan, C.C., and Hesselink, D.A. (2018). Liquid Biopsies to Monitor Solid Organ Transplant Function. *Ther. Drug Monit.* 40, 515–525.

Vernice NA, Meydan C, Afshinnekoo E, Mason CE. Long-term spaceflight and the cardiovascular system. *Precision Clinical Medicine.* 2020. Jun 16. pbaa022.

De Vlaminck, I., Khush, K.K., Strehl, C., Kohli, B., Luikart, H., Neff, N.F., Okamoto, J., Snyder, T.M., Cornfield, D.N., Nicolls, M.R., et al. (2013). Temporal Response of the Human Virome to Immunosuppression and Antiviral Therapy. *Cell* 155, 1178–1187.

De Vlaminck, I., Valantine, H.A., Snyder, T.M., Strehl, C., Cohen, G., Luikart, H., Neff, N.F., Okamoto, J., Bernstein, D., Weisshaar, D., et al. (2014). Circulating Cell-Free DNA Enables Noninvasive Diagnosis of Heart Transplant Rejection. *Sci. Transl. Med.* 6, 241ra77–241ra77.

Volik, S., Alcaide, M., Morin, R.D., and Collins, C. (2016). Cell-free DNA (cfDNA): Clinical significance and utility in cancer shaped by emerging technologies. *Mol. Cancer Res.* 14, 898–908.

Wan, J.C.M.M., Heider, K., Gale, D., Murphy, S., Fisher, E., Mouliere, F., Ruiz-Valdepenas, A., Santonja, A., Morris, J., Chandrananda, D., et al. (2020). ctDNA monitoring using patient-specific sequencing and integration of variant reads. *Sci. Transl. Med.* 12, eaaz8084.

Wang, Y.K., Bashashati, A., Anglesio, M.S., Cochrane, D.R., Grewal, D.S., Ha, G., McPherson, A., Horlings, H.M., Senz, J., Prentice, L.M., et al. (2017). Genomic consequences of aberrant DNA repair mechanisms stratify ovarian cancer histotypes. *Nat Genet.*

Williams, D., Kuipers, A., Mukai, C., and Thirsk, R. (2009). Acclimation during space flight: Effects on human physiology. *CMAJ* 180, 1317–1323.

Wortzel, I., Dror, S., Kenific, C.M., and Lyden, D. (2019). Exosome-Mediated Metastasis: Communication from a Distance. *Dev. Cell* 49, 347–360.

Yakes, F.M., and Van Houten, B. (1997). Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress. *Proc. Natl. Acad. Sci. U. S. A.* 94, 514–519.

Yoshioka, Y., Kosaka, N., Konishi, Y., Ohta, H., Okamoto, H., Sonoda, H., Nonaka, R., Yamamoto, H., Ishii, H., Mori, M., et al. (2014). Ultra-sensitive liquid biopsy of circulating extracellular vesicles using ExoScreen. *Nat. Commun.* 5, 3591.

Zhang, H., Freitas, D., Kim, H.S., Fabijanic, K., Li, Z., Chen, H., Mark, M.T., Molina, H., Martin, A.B., Bojmar, L., et al. (2018). Identification of distinct nanoparticles and subsets of extracellular vesicles by asymmetric flow field-flow fractionation. *Nat. Cell Biol.* 20, 332–343.

Zhang, J., Li, J., Saucier, J.B., Feng, Y., Jiang, Y., Sinson, J., McCombs, A.K., Schmitt, E.S., Peacock, S., Chen, S., et al. (2019). Non-invasive prenatal sequencing for multiple Mendelian monogenic disorders using circulating cell-free fetal DNA. *Nat. Med.* 1–9.

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10, 1523.

Zwirner, K., Hilke, F.J., Demidov, G., Ossowski, S., Gani, C., Rieß, O., Zips, D., Welz, S., and Schroeder, C. (2018). Circulating cell-free DNA: A potential biomarker to differentiate inflammation and infection during radiochemotherapy. *Radiotherapy Oncology.*

Figures, Tables, and Supplementary Tables/Figures for:

Cell-free DNA (cfDNA) and exosome profiling from a year-long human spaceflight reveals circulating biomarkers

Daniela Bezdan¹, Kirill Grigorev¹, Cem Meydan¹, Fanny A. Pelissier Vatter², Michele Cioffi², Varsha Rao³, Kiichi Nakahira⁴, Philip Burnham⁵, Ebrahim Afshinnkoo^{1,6,7}, Craig Westover¹, Daniel Butler¹, Chris Moszary¹, Matthew MacKay¹, Jonathan Foon¹, Tejaswini Mishra³, Serena Lucotti², Brinda K. Rana⁸, Ari M. Melnick⁹, Haiying Zhang¹⁰, Irina Matei², David Kelsen¹⁰, Kenneth Yu¹⁰, David C Lyden², Lynn Taylor¹¹, Susan M Bailey¹¹, Michael P. Snyder³, Francine E. Garrett-Bakelman^{12,13,14}, Stephan Ossowski¹⁵, Iwijn De Vlaminck¹⁶, Christopher E. Mason^{1,6,7,17}*

Figures

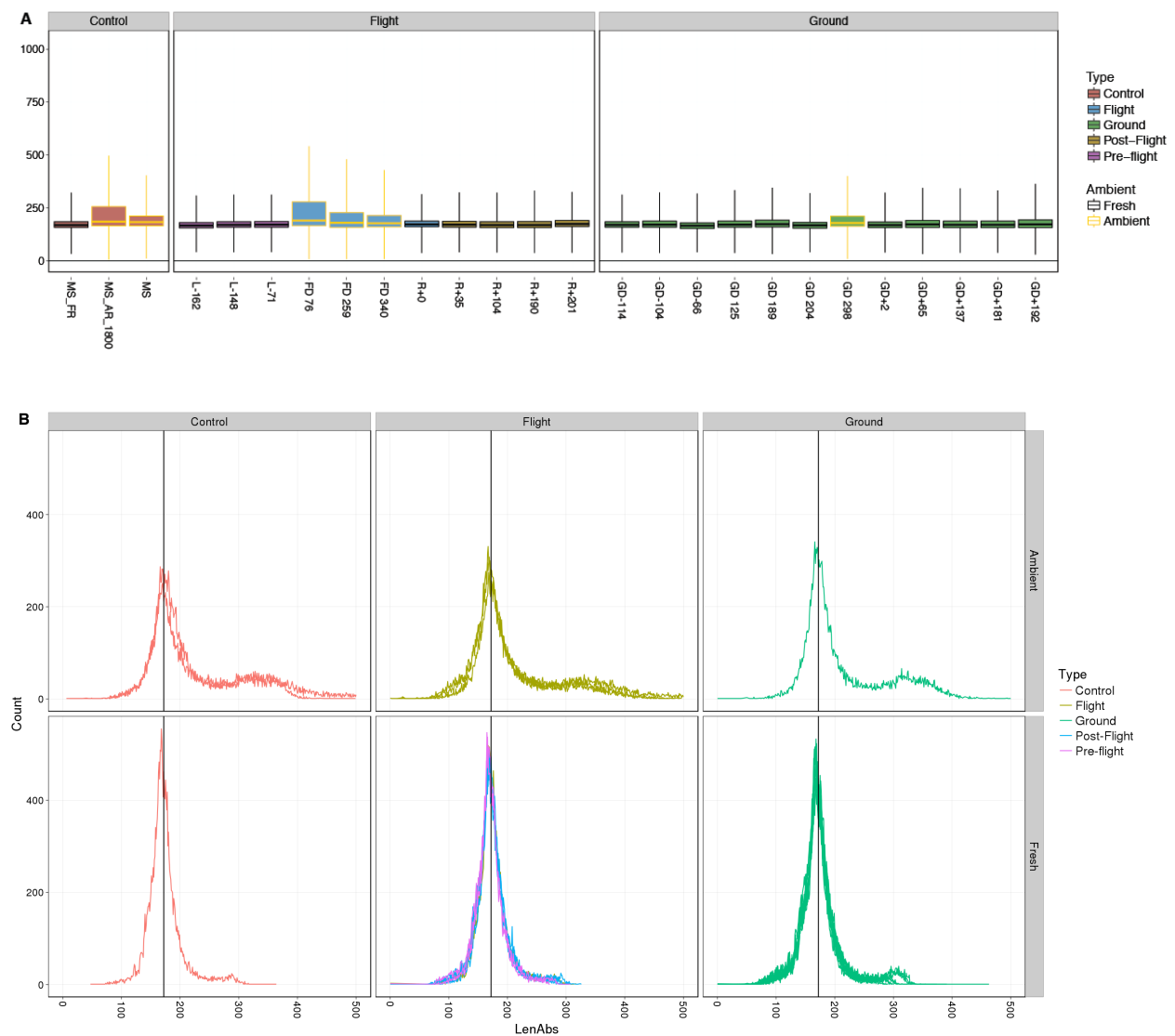


Figure 1. Size distribution of cfDNAs in ambient return, ambient return simulation and fresh samples. (A). Ambient return simulation samples (control and ground samples with yellow border) show a highly similar pattern as observed for inflight samples (blue box with yellow border). Long cfDNA fragments likely originate from blood cells damaged during transport. (B) Ambient return samples show an increased fraction of cfDNA with fragment length > 300bp compared to fresh samples. Our experimental procedure does only allow interrogation of DNA fragments up to a

length of 500bp, thus the content of long mtDNA fragments contained in intact circulating mitochondria is not reflected in this analysis.

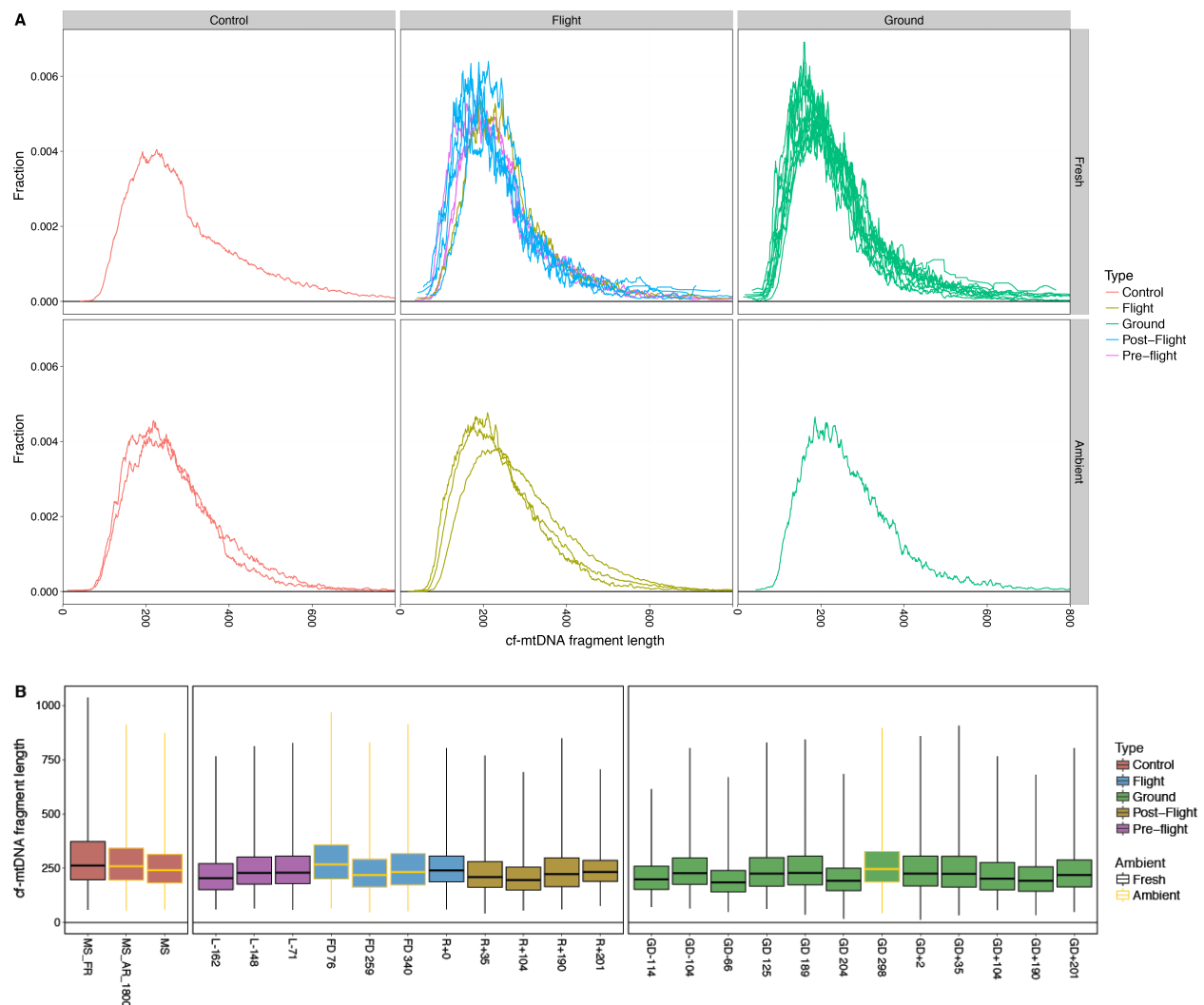


Figure 2. Size distribution of cf-mtDNAs. We observed a wider range of cf-mtDNA lengths compared to total cfDNA (from 100 to 600bp). (A) cf-mtDNA size distributions are similar in ground, flight and control samples, and are not affected by ambient return (AR) or AR simulation. (B) Average length of cf-mtDNA is significantly longer than the average length reported for chromosomal cfDNA (~250bp vs. ~160bp). The average length of cf-mtDNA is not affected by sample type (control, flight, ground) or sample handling (fresh, AR, AR simulation).

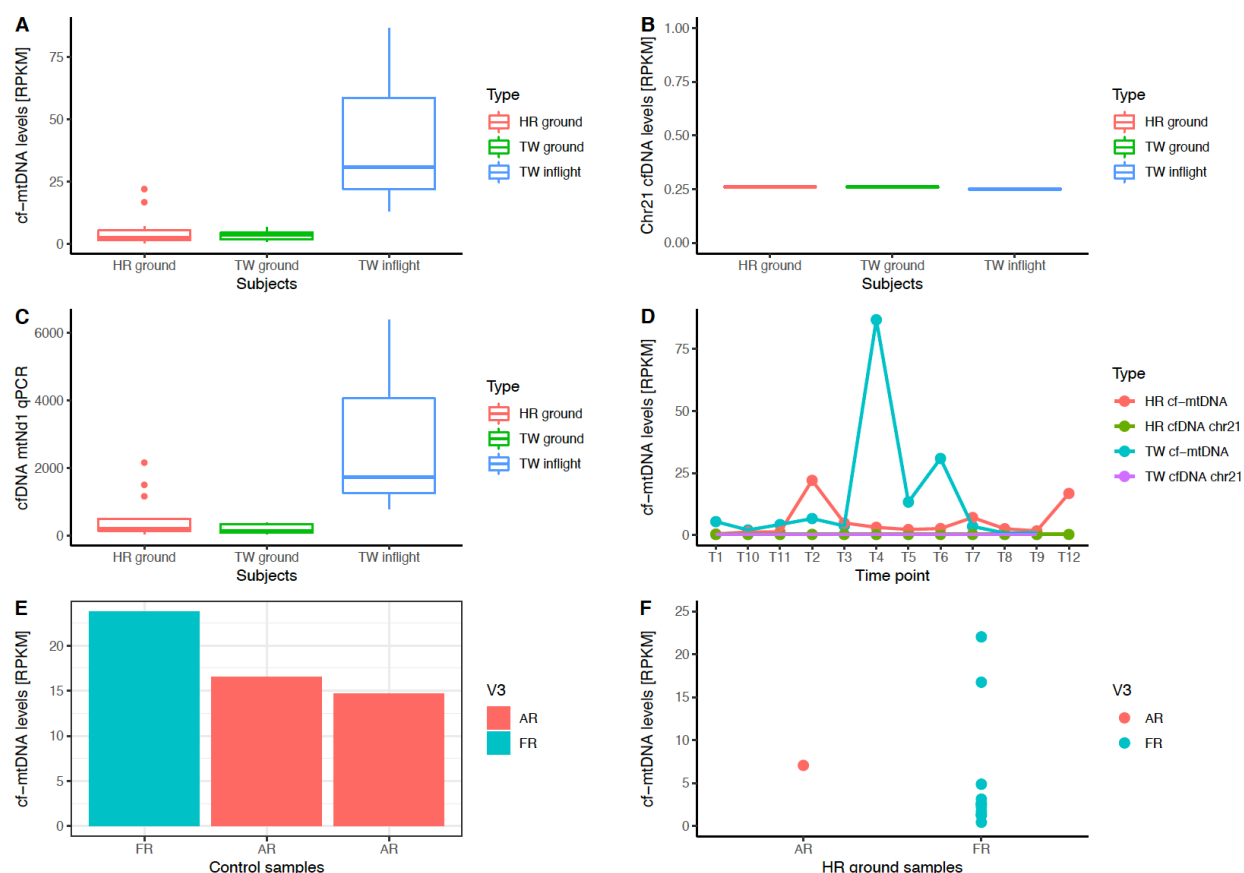


Figure 3. Analysis of normalized cfDNA read counts by chromosome, including the mitochondrial genome. (A) TW exhibits a significant increase in cell-free mtDNA during space flight compared to TW and HR ground samples. Counts are reads per kilobase per Million reads, or RPKM (B) Chromosomes do not show any change in RPKM during space flight, as exemplified using chr21. (C) Q-PCR based validation of increased cf-mtDNA fraction in plasma during space flight. (D) Normalized cf-mtDNA fraction and fraction of reads mapping to chr21 for 12 time point during the mission (T4-T6 = space flight). The highest increase in cf-mtDNA fraction is observed during the first months on ISS. (E) Ambient return simulation using two control samples showed no increase in cf-mtDNA compared to fresh samples, but a slight reduction. (F) Ambient return simulation (AR) using one HR ground sample did not show a significant increase in cf-mtDNA fraction. Two outliers within the fresh samples (FR) indicate that other conditions (e.g. stress, disease, immune reaction) could have influenced cf-mtDNA levels of HR on the ground.

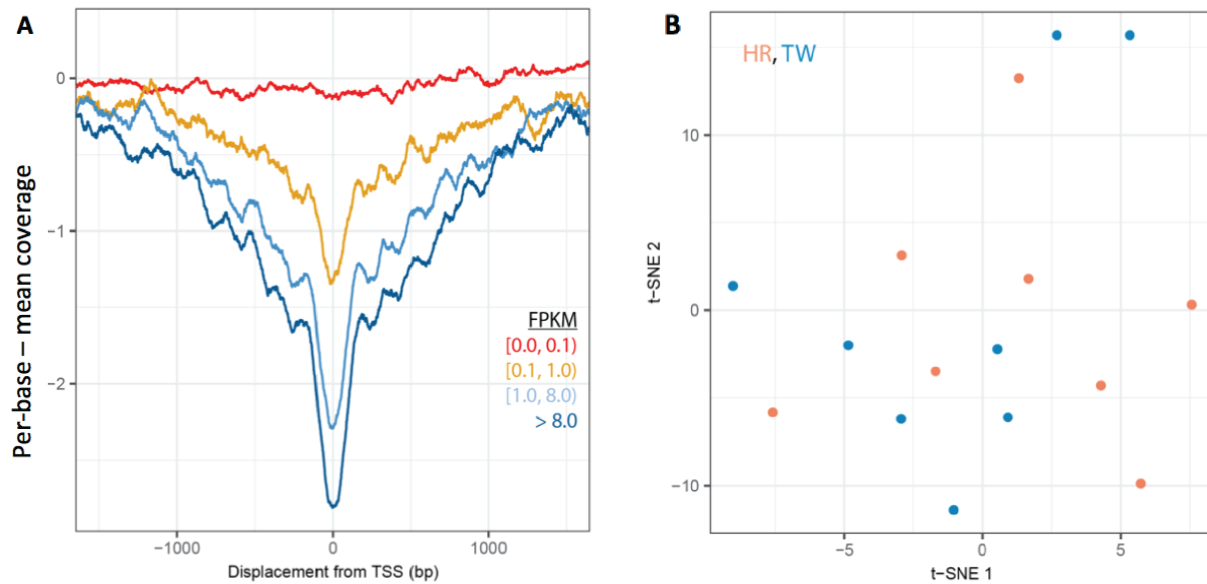
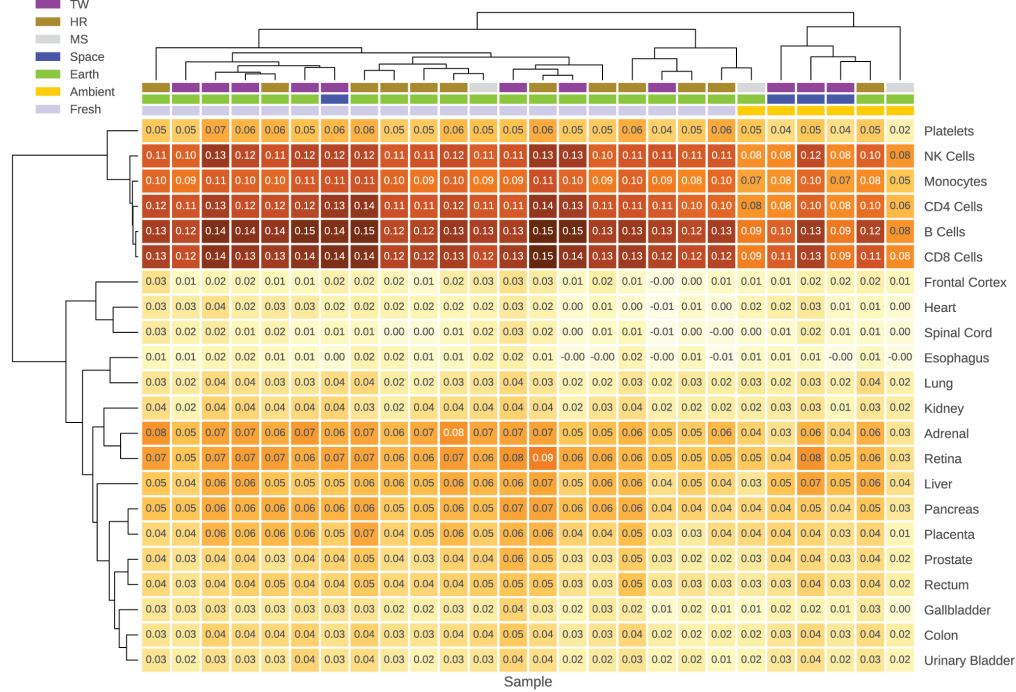


Figure 4. cfDNA nucleosome footprinting. (A) Nucleosome depletion in cfDNA around transcription start sites (TSS) is highly correlated with the expression of the respective genes and can therefore be used to estimate promoter activity and gene expression. (B) t-SNE based on genome-wide promoter nucleosome footprint of cfDNA samples reveals no clustering of flight subject and ground subject samples.

A



B

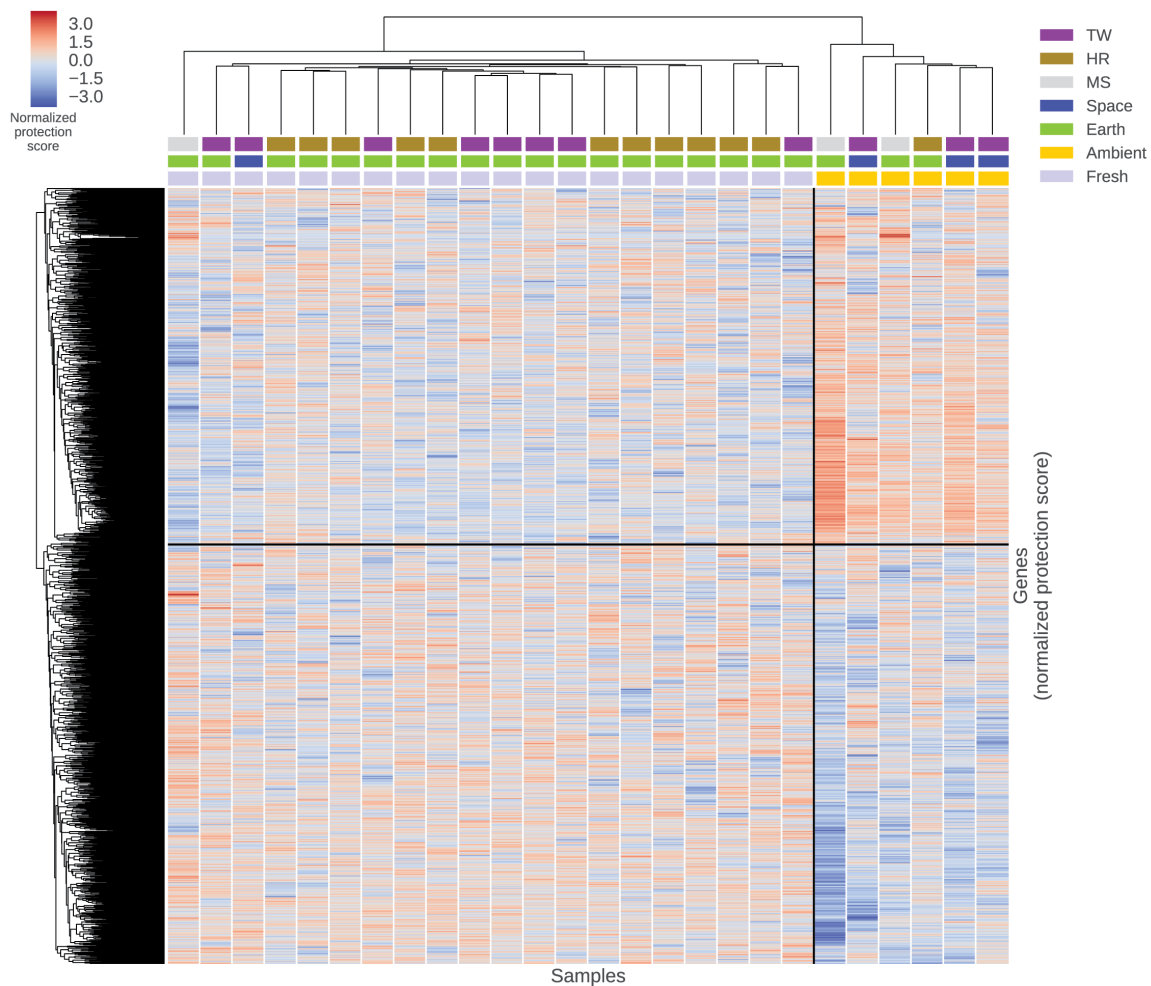


Figure 5. Tissue of origin deconvolution. (A) Correlation coefficients (multiplied by -1) for each tissue in each sample, clustered by sample and by tissue. The highest signals are, expectedly, from cells of hematopoietic origin. Spaceflight-dependent dynamics of tissue signal are confounded by the effect of ambient return, as suggested by ambient return samples tending to cluster together regardless of other features. (B) Clustering of samples using TSS protection in cfDNA as a measure of gene expression (lower protection correlates to higher expression). Ambient return samples cluster tightly together and uncover two major clusters of genes whose expression differs significantly from other samples, suggesting transport-related degradation processes or nucleosome detachment. Distribution of mean TSS protection per gene in ambient return and fresh samples is significantly different (t-test $p < 1e-3$).

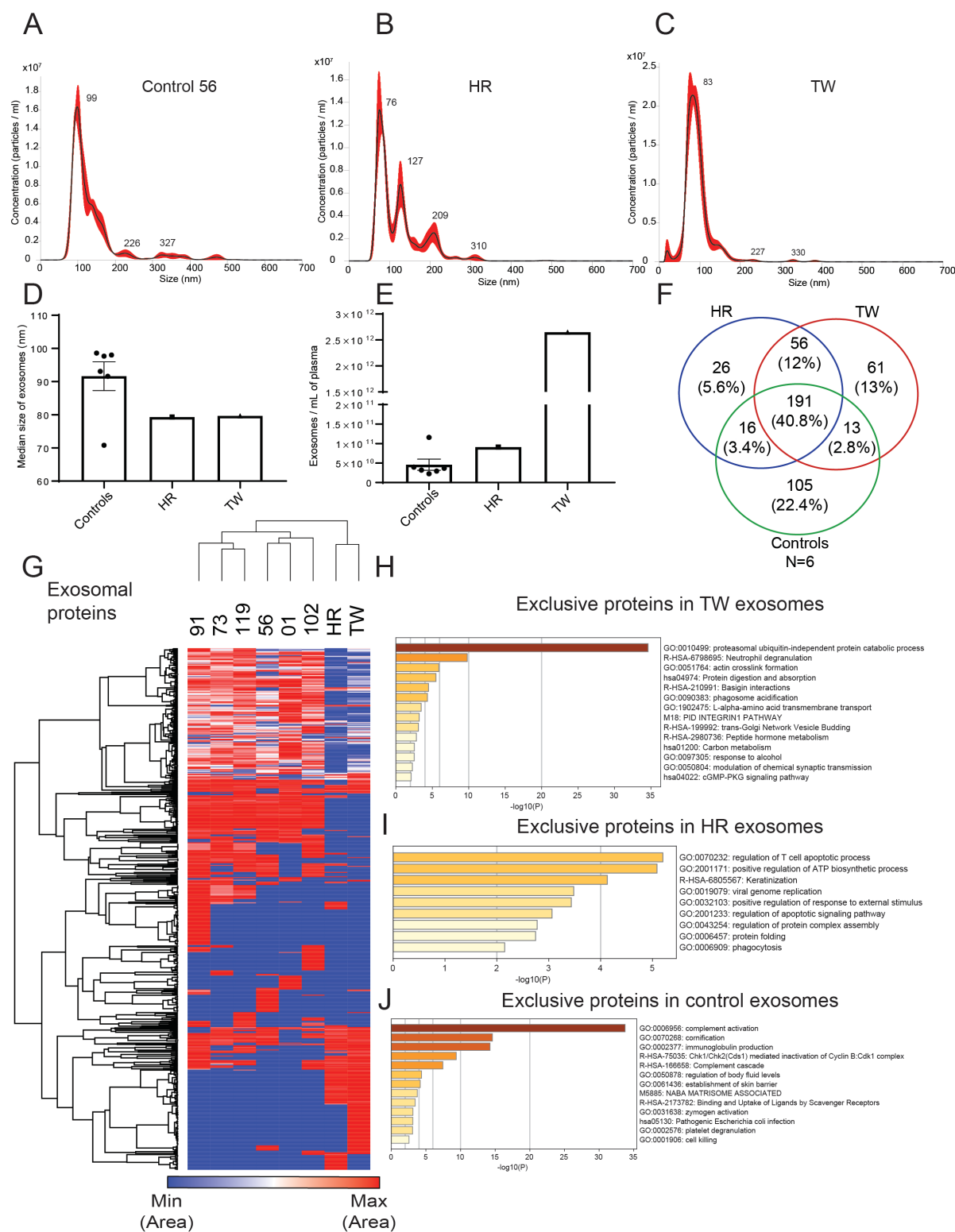
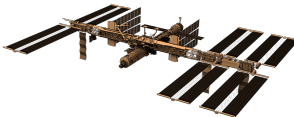
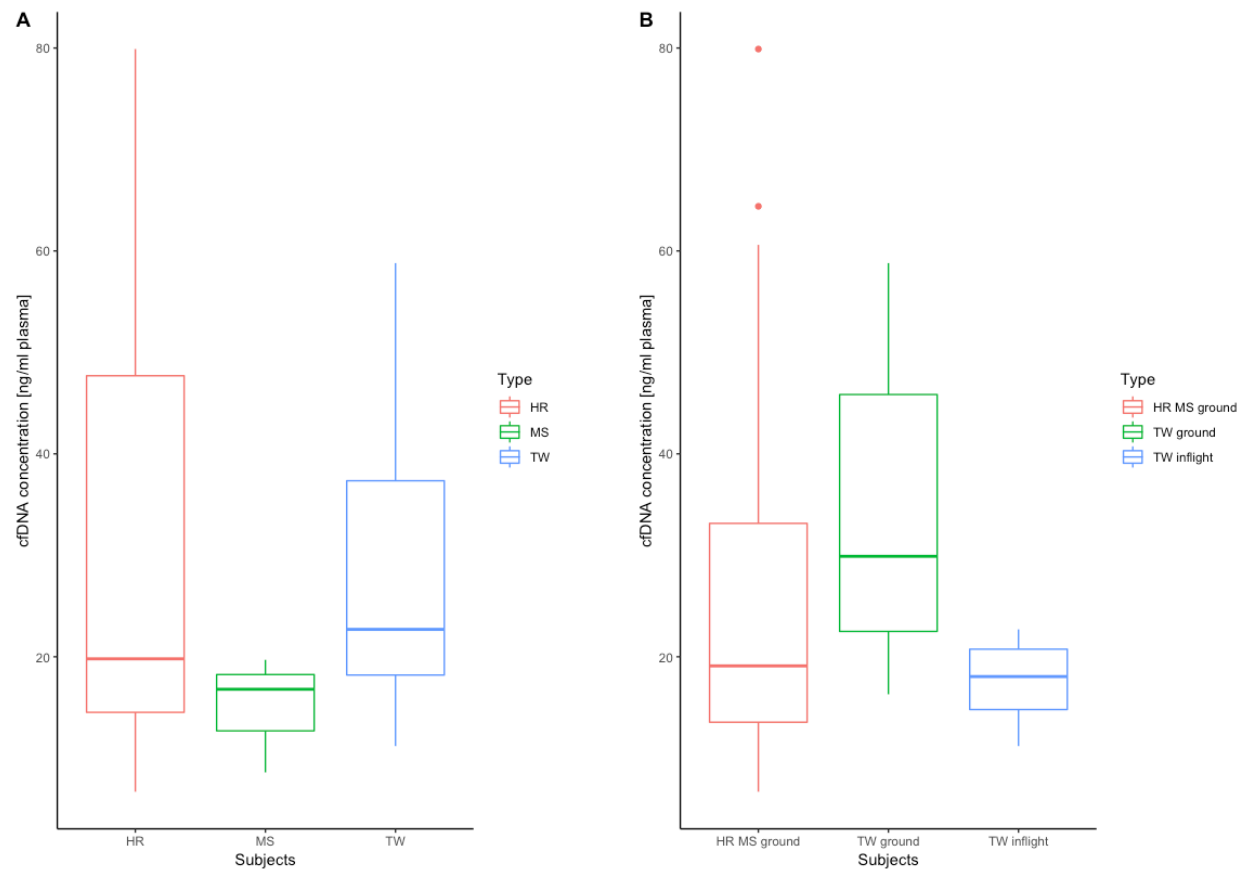


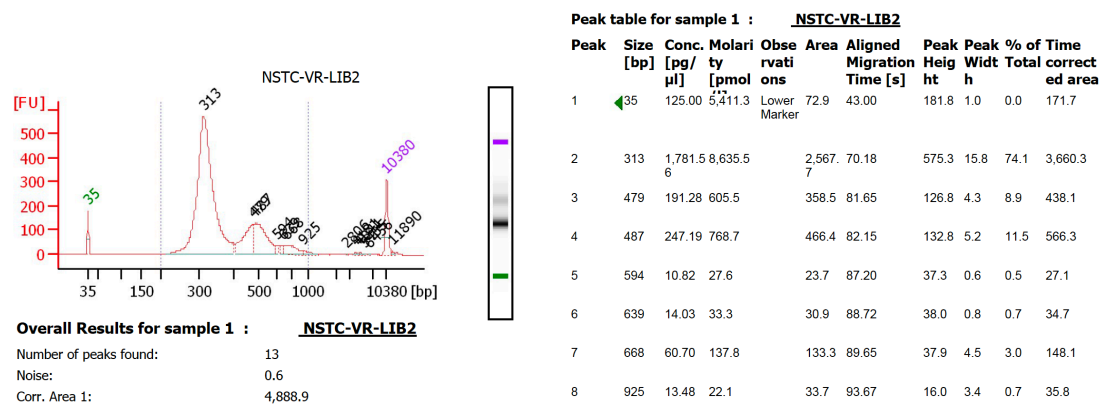
Figure 6. Characterization of plasma-derived exosomes isolated from HR and TW. Plasma samples were collected 3 years (TW) and 9 years (HR) post-flight. Nanosight profiles showing size distribution for exosomes isolated from the plasma of (A) Control, (B) HR, and (C) TW. Median size of exosomes (D) and exosome concentration (E) in TW (n=1), HR (n=1), and controls (n=6). (F) Venn diagram of exosomal proteins identified by mass spectrometry in plasma isolated from HR, TW and age-matched healthy controls. (G) Heatmap of plasma-derived exosomal proteins for HR, TW, and age-matched healthy controls. Pathway analysis of exclusive plasma-derived exosomal proteins from (H) TW, (I) HR, and (J) age-matched healthy controls.

Table 1. Overview of all plasma samples obtained during the 1-year mission. Subjects for this mission included the ground subject HR (blue), flight subject TW (green) and control subject MS (yellow). Samples taken on the ISS are highlighted in red. The last two columns show the concentration of cfDNA per ml plasma and the Q-PCR results for the mitochondrial transcript mtNd1 in copy/μl plasma.

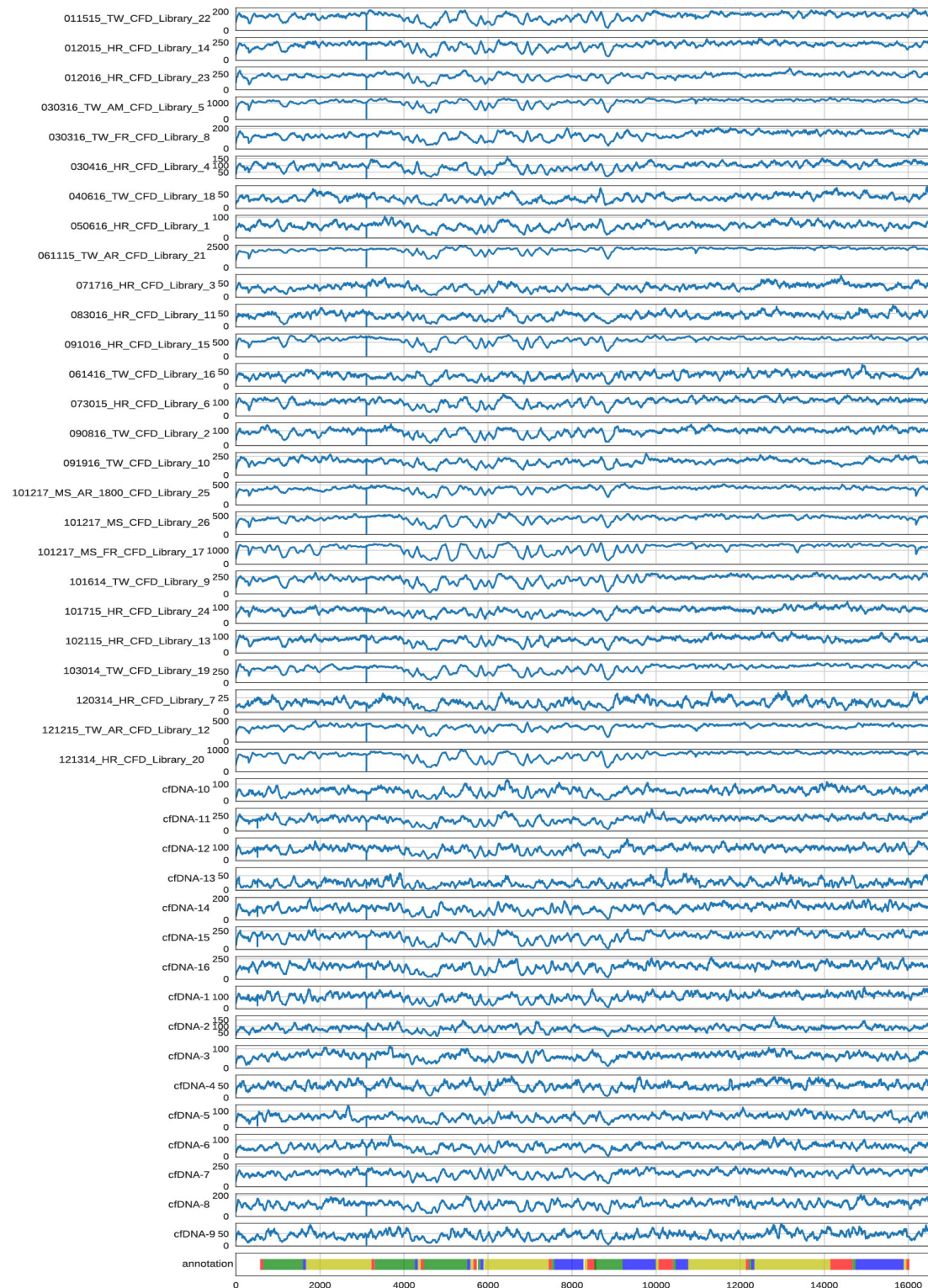
Time	Subject	Sample name	Total per plasma [ng/ml plasma]	mtNd1 Q-PCR [cp/μl plasma]	
PRE-FLIGHT		GD-114	15.5	44	
		GD-104	79.9	2159.6	
		GD-66	11.6	251.7	
FLIGHT		GD 125	6.7	277	
		GD 189	20.5	215.1	
		GD 204	64.4	193.7	
		GD 298	43.4	1502.8	
POST-FLIGHT		GD+2	60.6	157.7	
		GD+65	16.1	136.3	
		GD+137	19.1	68.7	
		GD+181	7.1	147.3	
		GD+192	22.9	1165.2	
AMBIENT		MS	8.6	3080.6	
RETURN		MS_AR	19.7	3590.8	
CONTROL		MS_AR_1800	16.8	1330.4	
PRE-FLIGHT		L-162	44.8	543.1	
		L-148	16.3	737.6	
		L-71	46.9	466	
FLIGHT		FD 76	20.1	6379.7	
		FD 259	11.2	786.9	
		FD 340	22.7	1735.3	
POST-FLIGHT		R+0	16	374.7	
		R+35	23.5	86.4	
		R+104	21.5	37.7	
		R+190	58.8	138.5	
		R+201	29.9	349.1	



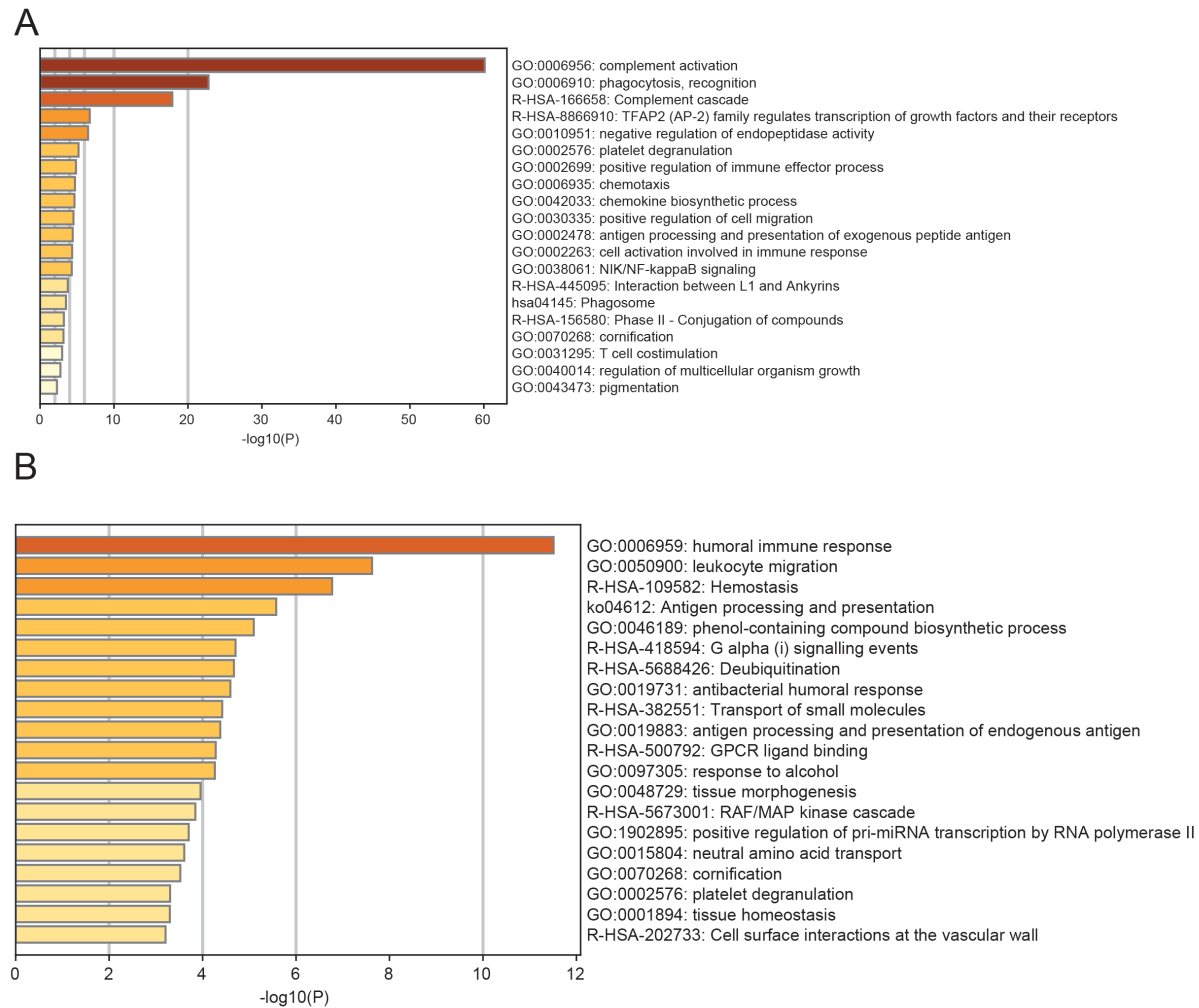
Supplemental Figure 1: cfDNA concentration in plasma of ground subject (HR), flight subject (TW) and ground controls (MS). (A) Comparison of cfDNA concentrations between HR, MS and TW samples. (B) Comparison of TW pre- and -post-flight to TW inflight and to the combined ground samples of HR and MS. No significant differences were observed.



Supplemental Figure 2. Library Fragment distribution for the cell-free DNA libraries. Pooled libraries were run on the Agilent Bioanalyzer 2100, with the entire fragment range area estimated to be 4,888.9. The first fraction peak was estimated to be at 313bp and the second peak was at 466bp. Given that the Illumina adapters add 120bp to each fragment size, this means that the estimated size of the first fragment set is 193bp and the second set is 346bp. The total area of the first peak represents 74.8% (3,660.3/4,888.9) of the signal.



Supplemental Figure 3: Read coverage distribution across the mitochondrial genome for all samples analyzed in this study. We observed continuous coverage of the complete mitochondrial genome in all samples.



Supplemental Figure 4: (A) Pathway analysis of inclusive plasma-derived exosomal proteins from TW (n=1), HR (n=2), and (J) age-matched healthy controls (n=6). (B) Pathway analysis of inclusive plasma-derived exosomal proteins from TW and HR excluding age-matched healthy controls

Supplementary Tables

Supplemental Table 1: List of inclusive proteins in exosomes isolated from the plasma of TW, HR and age-matched healthy controls.

Inclusive in TW, HR and control exosomes		
35 kDa inter-alpha-trypsin inhibitor heavy chain H4	Fibrinogen gamma chain	Platelet glycoprotein 4
Actin, cytoplasmic 1	Fibulin 1	Polymeric immunoglobulin receptor
Adiponectin	Ficolin-2	Protein AMBP
Afamin	Ficolin-3	Protein IGHV1-46
Alpha-1-acid glycoprotein 1	Galectin-3-binding protein	Protein IGHV1OR15-1
Alpha-1-acid glycoprotein 2	Gelsolin	Protein IGHV2-26
Alpha-1-antichymotrypsin	Glutathione peroxidase	Protein IGHV3-13
Alpha-1-antitrypsin	Glyceraldehyde-3-phosphate dehydrogenase	Protein IGHV3-15
Alpha-1B-glycoprotein	Haptoglobin	Protein IGHV3-21
Alpha-2-antiplasmin	Haptoglobin-related protein	Protein IGHV3-35
Alpha-2-HS-glycoprotein	HCG2043239	Protein IGHV3-38
Alpha-2-macroglobulin	Heat shock cognate 71 kDa protein	Protein IGHV3-43
Angiotensinogen	Hemoglobin subunit alpha	Protein IGHV3-49
Antithrombin-III	Hemoglobin subunit beta	Protein IGHV3-64
Apolipoprotein A-I	Hemoglobin subunit delta	Protein IGHV3-73
Apolipoprotein A-II	Hemopexin	Protein IGHV3OR15-7
Apolipoprotein A-IV	Heparin cofactor 2	Protein IGHV3OR16-12
Apolipoprotein B-100	Histidine-rich glycoprotein	Protein IGHV3OR16-13
Apolipoprotein C-III	Ig alpha-1 chain C region	Protein IGHV3OR16-9
Apolipoprotein C-IV	Ig alpha-2 chain C region	Protein IGHV4-28
Apolipoprotein D	Ig delta chain C region	Protein IGHV4-34
Apolipoprotein E	Ig gamma-1 chain C region	Protein IGHV4-4
Apolipoprotein L1	Ig gamma-2 chain C region	Protein IGHV5-51
Apolipoprotein M	Ig gamma-3 chain C region	Protein IGHV6-1
Apolipoprotein(a)	Ig gamma-4 chain C region	Protein IGKV1-16
Band 3 anion transport protein	Ig heavy chain V-I region 5	Protein IGKV1-17
Beta-2-glycoprotein 1	Ig kappa chain V-IV region	Protein IGKV1-33
C4b-B	Ig lambda-2 chain C regions	Protein IGKV2-30
C4b-binding protein alpha chain	Ig mu chain C region	Protein IGKV2-40
C4b-binding protein beta chain	IgGfC-binding protein	Protein IGKV2D-24
Carboxypeptidase N catalytic chain	Immunoglobulin heavy variable 3-43D	Protein IGKV3D-20
Carboxypeptidase N subunit 2	Immunoglobulin heavy variable 4-38-2	Protein IGLV1-47
CD5 antigen-like	Immunoglobulin kappa constant	Protein IGLV2-11
CD81 antigen	Immunoglobulin lambda variable 1-51	Protein IGLV2-14
Ceruloplasmin	Immunoglobulin lambda variable 3-10	Protein IGLV3-19
Clusterin	Immunoglobulin lambda variable 8-61	Protein IGLV3-27
Coagulation factor V	Immunoglobulin lambda-like polypeptide 5	Protein IGLV7-43
Coagulation factor XII	Integrin alpha-IIb	Protein IGLV7-46
Coagulation factor XIII A chain	Inter-alpha-trypsin inhibitor heavy chain H1	Protein IGLV9-49
Coagulation factor XIII B chain	Inter-alpha-trypsin inhibitor heavy chain H2	Protein S100-A8
Collectin-11	Keratin, type I cuticular Ha1	Protein S100-A9
Complement C1q subcomponent subunit A	Keratin, type I cuticular Ha3-II	Proteoglycan 4
Complement C1q subcomponent subunit B	Keratin, type I cuticular Ha5	Prothrombin
Complement C1q subcomponent subunit C	Keratin, type I cuticular Ha6	Ras-related protein Rap-1b
Complement C1q tumor necrosis factor-related protein 3	Keratin, type I cytkeletal 10	Reelin
Complement C1r subcomponent	Keratin, type I cytkeletal 14	Retinol binding protein 4, plasma, isoform CRA_b
Complement C1s subcomponent	Keratin, type I cytkeletal 9	Serotransferrin
Complement C3	Keratin, type II cytkeletal 1	Serum albumin
Complement C4 beta chain	Keratin, type II cytkeletal 2 epidermal	Serum amyloid P-component
Complement C5	Keratin, type II cytkeletal 4	Serum paraoxonase/arylesterase 1
Complement component C6	Keratin, type II cytkeletal 5	Solute carrier family 2, facilitated glucose transporter member 1
Complement component C7	Keratin, type II cytkeletal 6B	Thrombospondin-1
Complement component C8 alpha chain	Kininogen-1	Transferrin receptor protein 1
Complement component C8 beta chain	Lipopolysaccharide-binding protein	Transferrin
Complement component C8 gamma chain	Lysozyme C	Truncated apolipoprotein C-I
Complement component C9	Mannan-binding lectin serine protease 1	Vitamin D-binding protein
Complement factor B	Mannan-binding lectin serine protease 2	Vitamin K-dependent protein S
Complement factor H	Oncoprotein-induced transcript 3 protein	Vitronectin
Complement factor H-related protein 1	Phosphatidylinositol-glycan-specific phospholipase D	von Willebrand factor
Complement factor H-related protein 5	Pigment epithelium-derived factor	Zinc-alpha-2-glycoprotein
Complement factor I light chain	Plasma kallikrein heavy chain	TTR
Corticosteroid-binding globulin	Plasma protease C1 inhibitor	VTN
Erythrocyte band 7 integral membrane protein	Plasminogen	VWF
Extracellular matrix protein 1	Platelet factor 4	

Supplemental Table 2: List of unique proteins in exosomes isolated from the plasma of TW, HR and age-matched healthy controls.

Unique in TW exosomes		Unique in HR exosomes	
Acetyl-CoA carboxylase 1	Nephrilysin	Alpha-enolase	
Alpha-soluble NSF attachment protein	Prehnycysteine oxidase 1	Cell division control protein 42 homolog	
Aminopeptidase N	Proteasome subunit alpha type	Clathrin heavy chain 1	
Angiotensin-related protein 6	Proteasome subunit alpha type-2	Coagulation factor VIII	
Atypical chemokine receptor 1	Proteasome subunit alpha type-3	Creatine kinase M-type	
Basement membrane-specific heparan sulfate proteoglycan core protein	Proteasome subunit alpha type-5	EMILIN-1	
Basigin	Proteasome subunit alpha type-6	Galectin-3	
Beta-Ala-His dipeptidase	Proteasome subunit alpha type-7	Galectin-9	
Brain-specific angiogenesis inhibitor 1-associated protein 2	Proteasome subunit beta type	Heat shock protein 75 kDa, mitochondrial	
Brain-specific angiogenesis inhibitor 1-associated protein 2-like protein 1	Proteasome subunit beta type-1	Immunoglobulin heavy variable 2-70D	
Calmodulin-like protein 5	Proteasome subunit beta type-2	Kallicrein-1	
Catalase	Proteasome subunit beta type-3	Keratin, type I cuticular Ha3-I	
Cation-dependent mannose-6-phosphate receptor O	Proteasome subunit beta type-4	Keratin, type II cuticular Hb2	
CD82 antigen	Proteasome subunit beta type-5	Keratin-associated protein 11-1	
Collagen alpha-1(VI) chain	Proteasome subunit beta type-6	Keratin-associated protein 2-2	
Complement factor H-related protein 4	Proteasome subunit beta type-7	Peptidyl-prolyl cis-trans isomerase A-like 4H	
Dipeptidyl peptidase 4	Proteasome subunit beta type-8	Plectin	
Epidermal growth factor receptor kinase substrate 8	Protein IGKV1-8	Prolactin-inducible protein	
Erythroid membrane-associated protein	Ras-related protein Rab-5C	Protein-glutamine gamma-glutamyltransferase 2	
Gamma-glutamyltranspeptidase 1	Ras-related protein Rab-7a	Putative keratin-87 protein	
G-protein coupled receptor family C group 5 member C	Rib e-5-ph phosphate isomerase	Sialic acid-binding Ig-like lectin 16	
Heat shock protein beta-1	Septin-5	Tandem C2 domains nuclear protein	
Histone H2A	Sodium/potassium-transporting ATPase subunit alpha	Thym in beta-4	
Immunoglobulin kappa variable 1-6	Suprabasin	Transitional endoplasmic reticulum ATPase	
Immunoglobulin lambda variable 5-39	Syntenin-1	Transmembrane protein 198	
Integrin beta-1	Tri eph phosphate isomerase	Tripeptidyl-peptidase 2	
Lactadherin	Trypsin-1		
Large neutral amino acids transporter small subunit 1	Vasodilator-stimulated ph phoprotein		
Large neutral amino acids transporter small subunit 3	Vesicle transport protein		
Monocyte differentiation antigen CD14	V-type proton ATPase 16 kDa proteolipid subunit		
	V-type proton ATPase catalytic subunit A		
Unique in Control exosomes			
14-3-3 protein epsilon	HLA class I histocompatibility antigen, B-73 alpha chain	Protein IGHV2-70	
14-3-3 protein eta	HLA class I histocompatibility antigen, B-8 alpha chain	Protein IGHV3-30	
14-3-3 protein gamma	HLA class I histocompatibility antigen, B-82 alpha chain	Protein IGHV3-33	
14-3-3 protein sigma	HLA class I histocompatibility antigen, Cw-18 alpha chain	Protein IGHV3-48	
14-3-3 protein theta	HLA class I histocompatibility antigen, Cw-4 alpha chain	Protein IGHV3-53	
Actin, alpha cardiac muscle 1	Homerin	Protein IGHV3-74	
Actin, cytoplasmic 2	Hyaluronan-binding protein 2	Protein IGHV3OR16-10	
Beta-actin-like protein 2	Ig heavy chain V-III region 23	Protein IGHV3OR16-8	
Carboxypeptidase B2	Ig kappa chain V-III region VG	Protein IGHV4-31	
Cholinesterase	Immunoglobulin heavy variable 1-8	Protein IGHV4-59	
Collectin-10	Immunoglobulin heavy variable 3-20	Protein IGKV1-27	
Complement C1r subcomponent-like protein	Immunoglobulin J chain	Protein IGKV1-39	
Complement C2	Immunoglobulin kappa joining 1	Protein IGKV2-24	
Complement factor H-related protein 2	Immunoglobulin kappa variable 2-29	Protein IGKV2D-30	
Complement factor H-related protein 3	Insulin-like growth factor-binding protein complex acid labile subunit	Protein IGKV3-7	
HCG1745306, isoform CRA_a	Integrin beta-3	Protein IGKV6D-21	
Hemoglobin subunit gamma-1	Inter-alpha-trypsin inhibitor heavy chain H3	Protein IGLV10-54	
Hepatocyte growth factor activator	Keratin, type I cuticular Ha2	Protein IGLV1-44	
Hepatocyte growth factor-like protein	Keratin, type I cuticular Ha7	Protein IGLV3-1	
HLA class I histocompatibility antigen, A-11 alpha chain	Keratin, type I cuticular Ha8	Protein IGLV3-21	
HLA class I histocompatibility antigen, A-2 alpha chain	Keratin, type I cytkeletal 15	Protein IGLV3-25	
HLA class I histocompatibility antigen, A-23 alpha chain	Keratin, type I cytkeletal 16	Protein IGLV4-60	
HLA class I histocompatibility antigen, A-25 alpha chain	Keratin, type I cytkeletal 17	Protein SAA2-SAA4	
HLA class I histocompatibility antigen, A-30 alpha chain	Keratin, type I cytkeletal 18	Protein tyr ine phosphate, receptor type, C	
HLA class I histocompatibility antigen, A-34 alpha chain	Keratin, type II cytkeletal 1b	Putative HLA class I histocompatibility antigen, alpha chain H	
HLA class I histocompatibility antigen, A-34 alpha chain	Keratin, type II cytkeletal 6A	Putative V-set and immunoglobulin domain-containing-like protein IGHV4OR15-8	
HLA class I histocompatibility antigen, A-68 alpha chain	Keratin, type II cytkeletal 75	Pyruvate kinase isozymes M1/M2	
HLA class I histocompatibility antigen, A-74 alpha chain	Keratin, type II cytkeletal 79	Radixin	
HLA class I histocompatibility antigen, B-39 alpha chain	Keratin, type II cytkeletal 80	Selenoprotein P	
HLA class I histocompatibility antigen, B-41 alpha chain	Latent-transforming growth factor beta-binding protein 1	Serum amyloid A-1 protein	
HLA class I histocompatibility antigen, B-42 alpha chain	Platelet factor 4 variant	SWISS-PROT:P15636 Protease I precursor Lysyl endopeptidase Achromobacter lyticus.	
HLA class I histocompatibility antigen, B-44 alpha chain	Pregnancy zone protein	Tetranectin	
HLA class I histocompatibility antigen, B-55 alpha chain	Properdin	THBS3 protein	
HLA class I histocompatibility antigen, B-59 alpha chain	Protein IGHV1-2	Thyroxine-binding globulin	
HLA class I histocompatibility antigen, B-7 alpha chain	Protein IGHV1-69	Uncharacterized protein	
Shared by TW and HR exosomes (not in controls)			
4F2 cell-surface antigen heavy chain	Immunoglobulin heavy variable 5-10-1		
60 kDa heat shock protein, mitochondrial	Immunoglobulin J chain		
A disintegrin and metalloproteinase with thrombopodin motifs 13	Immunoglobulin kappa variable 3-15		
Actin, gamma-enteric smooth muscle	Immunoglobulin kappa variable 3-20		
Aldehyde dehydrogenase family 16 member A1	Keratin, type I cytoskeletal 13		
APOC4-APOC2 readthrough (NMD candidate)	Keratin, type II cuticular Hb5		
Bone morphogenetic protein 1	Keratin, type II cuticular Hb6		
Carbonic anhydrase 1	Leukosialin		
Cathelicidin antimicrobial peptide	Leukocyte surface antigen CD53		
CD99 antigen	Major vault protein		
Cholesteryl ester transfer protein	Metalloredutase STEAP3		
Dermcidin	Neutral amino acid transporter B(0)		
Equilibrative nucleide transporter 1	Peroxisedoxin-2		
Ferritin heavy chain	Probable sodium-coupled neutral amino acid transporter 6		
Ferritin light chain	Protein IGHV3-72		
Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	Protein IGKV2D-29		
Guanine nucleotide-binding protein G(t) subunit alpha-1	Protein IGLV2-8		
HCG2041221	Protein TFG		
Heat shock 70 kDa protein 1B	Protein TRAJ61		
High affinity cationic amino acid transporter 1	Ras-related protein Rab-10		
Histone H2B type 1-K	Ribonuclease 4		
Histone H3.3	Secreted ph phoprotein 24		
Histone H4	Selenoprotein P		
HLA class I histocompatibility antigen, A-3 alpha chain	Serum amyloid A-4 protein		
HLA class I histocompatibility antigen, B-73 alpha chain	Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1		
Immunoglobulin heavy variable 1-3	Syntaxin-7		
Immunoglobulin heavy variable 3-64D	Ubiquitin-60S ribosomal protein L40		
Immunoglobulin heavy variable 3-7	V-type proton ATPase subunit d 1		