Exosomes Mini-Series

Independent human mesenchymal stromal cell–derived extracellular vesicle preparations differentially attenuate symptoms in an advanced murine graft-versus-host disease model

Rabea J. Madel1,2,#, Verena Börger3,#, Robin Dittrich3, Michel Bremer3, Tobias Tertel3, Nhi Ngo Thi Phuong1, Hideo A. Baba4, Lambros Kordelas5, Simon Staubach3, Frank Stein6, Per Haberkant6, Matthias Hackl7, Regina Grillari8, Johannes Grillari8,9, Jan Buer1, Peter A. Horn3, Astrid M. Westendorf1, Sven Brandau10, Carsten J. Kirschning1,*, Bernd Giebel3,**

1 Institute of Medical Microbiology, University Hospital Essen, University of Duisburg-Essen, Essen, Germany
2 Department of Infectious Diseases, West German Centre for Infectious Diseases, University Hospital Essen, University of Duisburg-Essen, Essen, Germany
3 Institute for Transfusion Medicine, University Hospital Essen, University of Duisburg-Essen, Essen, Germany
4 Institute of Pathology, University Hospital Essen, University of Duisburg-Essen, Essen, Germany
5 Department of Hematology and Stem Cell Transplantation, University Hospital Essen, University of Duisburg-Essen, Essen, Germany
6 Proteomics Core Facility, EMBL Heidelberg, Heidelberg, Germany
7 TamiRNA GMBH, Vienna, Austria
8 Evercyte GmbH, Vienna, Austria
9 University of Natural Resources and Life Science, Vienna, Austria
10 Department of Otorhinolaryngology, University Hospital Essen, University of Duisburg-Essen, Essen, Germany

ARTICLE INFO

Article History:
Received 19 November 2022
Accepted 15 March 2023
Available online xxx

Key Words:
exosomes
extracellular vesicles
graft-versus-host disease
heterogeneity
mesenchymal stem cells
mesenchymal stromal cells

ABSTRACT

Background aims: Extracellular vesicles (EVs) harvested from conditioned media of human mesenchymal stromal cells (MSCs) suppress acute inflammation in various disease models and promote regeneration of damaged tissues. After successful treatment of a patient with acute steroid-refractory graft-versus-host disease (GVHD) using EVs prepared from conditioned media of human bone marrow–derived MSCs, this study focused on improving the MSC-EV production for clinical application.

Methods: Independent MSC-EV preparations all produced according to a standardized procedure revealed broad immunomodulatory differences. Only a proportion of the MSC-EV products applied effectively modulated immune responses in a multi-donor mixed lymphocyte reaction (mdMLR) assay. To explore the relevance of such differences in vivo, a first mouse GVHD model was optimized.

Results: The functional testing of selected MSC-EV preparations demonstrated that MSC-EV preparations revealing immunomodulatory capabilities in the mdMLR assay also effectively suppress GVHD symptoms in this model. In contrast, MSC-EV preparations, lacking such in vitro activities, also failed to modulate GVHD symptoms in vivo. Searching for differences of the active and inactive MSC-EV preparations, no concrete proteins or miRNAs were identified that could serve as surrogate markers.

Conclusions: Standardized MSC-EV production strategies may not be sufficient to warrant manufacturing of MSC-EV products with reproducible qualities. Consequently, given this functional heterogeneity, every individual MSC-EV preparation considered for the clinical application should be evaluated for its therapeutic potency before administration to patients. Here, upon comparing immunomodulating capabilities of independent MSC-EV preparations in vivo and in vitro, we found that the mdMLR assay was qualified for such analyses.

© 2023 International Society for Cell & Gene Therapy. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

* Correspondence: Carsten Kirschning, Institute of Medical Microbiology, University Hospital Essen, University of Duisburg-Essen, Hufelandstr. 55, 45147 Essen, Germany.
** Correspondence: Bernd Giebel, Institute for Transfusion Medicine, University Hospital Essen, University of Duisburg-Essen, Virchowstr. 179, 45147 Essen, Germany.
E-mail addresses: carsten.kirschning@uk-essen.de (C.J. Kirschning), bernd.giebel@uk-essen.de (B. Giebel).
# These authors contributed equally to this work.

https://doi.org/10.1016/j.jcyt.2023.03.008
1465-3249/© 2023 International Society for Cell & Gene Therapy, Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Introduction

Allogeneic hematopoietic stem cell transplantation is considered as the only therapy providing curative perspectives in several diseases, including leukemia [1]. However, allogeneic hematopoietic stem cell transplantation is associated with severe side effects. Approximately 35% of patients receiving matched or up to 50% of those receiving unrelated or alternative donor transplants, respectively, develop mild (grade I and II) to severe (grade III and IV) graft-versus-host disease (GVHD) symptoms. In GVHD, donor-derived immune cells attack and destroy healthy tissues of the recipients. Severe acute graft-versus-host disease (aGVHD) is associated with high morbidity and mortality [2]. As first-line treatment, patients with GVHD regularly receive immunosuppressive corticosteroids. However, 50% of all patients with GVHD are steroid refractory and have dismal long-term prognoses, with 2-year survival rates as little as 5–30% [3]. For now, no guideline defined second-line treatment exists. Thus, novel treatment strategies for patients with steroid-refractory GVHD are needed.

At the turn of the millennium, mesenchymal stromal cells (MSCs) had been shown to mediate immunosuppression [4,5]. Subsequently, Le Blanc et al. [6] successfully explored the capability of bone marrow (BM)-derived MSCs to suppress GVHD symptoms in a patient with severe treatment-resistant grade IV aGVHD. Consequently, numerous centers started to apply MSCs for steroid-refractory GVHD treatment, with varying outcomes [7]. Although several groups confirmed a benefit of MSC application in a proportion of patients with aGVHD, a phase 3 clinical trial failed to show efficacy [8,9]. The successes of clinical MSC studies depend on multiple factors such as application schemes and dosing as well as the MSC product quality itself [10]. Apparently, not all administered MSC products have the capability to effectively suppress GVHD symptoms.

Following reports that MSCs may not act in cellular manners but via extracellular vesicles (EVs), such as exosomes or microvesicles [11,12], we started exploring the therapeutic capabilities of EVs prepared from MSC-conditioned media (CM). To this end, we developed a scalable and standardized protocol for the manufacturing of MSC-EV products enabling the preparation of EVs from up to 10 L of MSC-CM [13–15]. After confirming the capability of obtained MSC-EV preparations in modulating pro-inflammatory immune reactions in vitro, and given the history of a patient with therapy-refractory GVHD, we performed an individual treatment attempt with one of our MSC-EV preparations: over a period of 2 weeks, the patients received seven escalating doses of an allogeneic MSC-EV product without showing any side effects [14]. Remarkably, after treatment, GVHD symptoms were massively suppressed for more than 4 months [14]. In order to consecutively translate MSC-EVs into regular GVHD clinic, we progressively improved our MSC-EV production and characterization platform. Up to now, using primary MSCs of healthy BM donors as starting material, we have manufactured more than 50 different scaled MSC-EV products in a standardized manner. Because MSCs are limited in their expansion capabilities, over time the MSCs of different donors had to be used as starting material for the MSC-EV production. Being aware that MSCs are a very heterogeneous cell entity [16–19], we considered functional heterogeneity among different MSC-EV preparations right from the beginning. Initially, we compared the pro- and anti-inflammatory content of resulting MSC-EV preparations and of four different MSC-EV products applied, the MSC-EV preparation with the highest anti-inflammatory cytokine content was chosen for the aforementioned individual treatment attempt of the patient with therapy-refractory GVHD [14].

Learning from the MSC field, which just has started discussing impacts of MSC heterogeneities and mechanism of action modalities on outcomes of clinical MSC trials in a more detailed manner [20–22], functional testing of given MSC-EV products is considered to be mandatory before clinical application [23,24]. However, the detailed knowledge on the mechanism of action (MoA) of how MSC-EVs improve clinical symptoms remains sparse. Coupled to this, it is a major challenge to set up appropriate assays allowing for the in vitro potency testing of given MSC-EV products [24,25]. To identify MSC-EV preparations containing the potency to effectively suppress GVHD symptoms, we established and applied an advanced murine GVHD model here. Subsequently, we compared the capability of a collection of independently produced human MSC-EV preparations in modulating the experimentally induced GVHD symptoms. Furthermore, we correlated the pre-clinical effects with the capabilities of the selected MSC-EV preparations to modulate T-cell activation in a novel multi-donor mixed lymphocyte reaction (mdMLR) assay [26,27] and tried to identify potential surrogate markers on the protein and microRNA level. Collectively, our data highlight the urgent need for potency testing of each individual MSC-EV product intended to be used in patients.

Materials and Methods

MSC growth and expansion

Human BM aspirates from healthy donors were obtained following informed consent according to the Declaration of Helsinki. Their usage was approved by the ethics committee of the University of Duisburg-Essen (12-2595-BO). MSCs were raised from BM aspirates and expanded in 10% human platelet lysate (hPL)-supplemented Dulbecco’s Modified Eagle’s Medium low-glucose (PAN Biotech) media as described previously [14,28]. For the production of CMs, regularly starting at passage 3 to passage 7, approximately 1000 cells/cm² were seeded into 4-layer stack cell factory systems (Thermo Fischer Scientific, Waltham, MA, USA). Upon reaching densities of approximately 50% confluency, media were exchanged and CMs harvested every 48 h. Reaching 80–90% confluency, MSCs were passaged. For the preservation of the CMs, cells and larger debris were removed by 2000g centrifugation for 15 min (Rotor: JS-5.3; Beckman Coulter, Krefeld, Germany). MSC-free CMs were stored at –20°C until use. CMs were screened regularly for mycoplasma contamination (VenorGeM OneStep; Minerva Biolabs, Berlin, Germany).

MSCs were analyzed according to the criteria of the International Society of Cell and Gene Therapy [29]. In summary, as described previously, MSCs were analyzed with fluorescently labeled with anti-CD14, anti-CD31, anti-CD34, anti-CD44, anti-CD45, anti-CD73, anti-CD90 and anti-CD105 antibodies (supplementary Table 1) by flow cytometry (Cytoflex; Software Cyteexpert 2.3; Beckman Coulter). Using passage 3 MSCs, the MSCs’ osteogenic and adipogenic differentiation potentials were confirmed by conventional MSC differentiation assays [14,18].

Preparation of EVs

For EV harvesting, CMs of individual MSC propagations were thawed and further cleared from solid components by 45 min of 6800g centrifugation (rotor: JS-5.3) and subsequent filtration through 0.22-μm rapid-flow filters (Nalgene; Thermo Fisher Scientific). Of note, CMs of independent MSC propagations were never pooled or mixed, even when cells originated from aliquots of the same MSC stocks. EVs were prepared from pooled CMs of individual MSC with an established polyethylene glycol precipitation method followed by ultracentrifugation (PEG-UC) exactly as described previously [14,15]. Obtained EV pellets were re-suspended in 10 mmol/L HEPES/0.9% NaCl buffer (Thermo Fisher Scientific). The concentration of prepared MSC-EVs was adjusted so that a 1-ml final sample contained the EV yield prepared from CMs of approximately 4 × 107 MSC equivalents. MSC-EV preparations were stored at –80°C until use. Repetitive thawing and freezing cycles were avoided. For control purposes, fresh hPL-supplemented media were processed in parallel (including
incubation for 48 h at 37°C, 5% CO₂, saturated water vapor atmosphere). During the course of the project three independent pooled hPL batches (50–70 individual units per pool) were used for the MSC propagation. Consequently, EVs of fresh hPL-supplemented media were prepared for each of the hPL batches, resulting in the PL₁-EV, PL₂-EV and PL₃-EV preparations.

Physical and protein–biochemical analyses of MSC-EV preparations

MSC-EV preparations were characterized according to the minimal information for studies of extracellular vesicles 2018 (MISEV2018) commitment [30]. The particle concentration and their average sizes within obtained MSC-EV preparations were determined by nanoparticle tracking analysis on a ZetaView PMX-120 BASIC platform (ParticleMetrix, Meerbusch, Germany) as described previously [15,31]. The device was calibrated using a polystyrene bead standard (100 nm; Thermo Fisher Scientific). Samples were loaded and videos were recorded at all 11 positions, with five repetitions. Additional settings included sensitivity 75, shutter 75, minimum brightness 20, minimum size 5 and maximum size 200. The median value (>50) for size was used for data analysis.

The protein contents of the MSC-EV preparations were determined using the bicinchoninic acid assay (Pierce, Rockford, IL, USA) in 96-well plates according to the manufacturer’s recommendations.

Western blotting

Western blot analyses were performed as described previously [15]. The following antibodies were used: anti-Syntenin (clone EPR8102; Abcam, Cambridge, UK), anti-Prohibitin (clone II-14-10; Thermo Fisher Scientific), anti-Calnexin (ab10286; Abcam), anti-Cytochrome C (clone JS-81; BD Biosciences, San Jose, CA, USA), anti-C9 (clone VJ1/20.3.1; kindly provided by Francisco Sánchez, Madrid, Spain) and anti-C6D3 (HSC6; BioLegend, San Diego, CA, USA). Band intensities were analyzed using Image J (National Institutes of Health, Bethesda, MD, USA).

Transmission electron microscopy

Transmission electron microscopy analyses were performed exactly as described recently [32]. Images were taken on a JEM 1400Plus electron microscope (JEOL, Tokyo, Japan) equipped with a 4096 × 4096-pixel CMOS camera (TemCam-F416; TVIPS, Gauting, Germany) and run at an operating voltage of 120 kV. Image acquisition software EMMENU (Version 4.09.83) was used for taking 16-bit images. Image post-processing was performed with the software ImageJ (Version 1.52b; National Institutes of Health).

mdMLR assay

To test allogeneic immune responses, a novel mdMLR was used [27]. EV preparations to be tested (regularly 25 µg) were applied to 6x10⁵ cells of a mixed peripheral blood mononuclear cell pool of 12 different donors and cultured in a final volume of 200 µL per well within 96-well u-bottom shape plates (Corning, Kaiserslautern, Germany) for 5 days. Thereafter, cells were harvested, stained with a collection of specifically selected fluorescently labelled antibodies (CD4-BV785 [300554, Clone: RPA-T4, BioLegend], CD8-BV650 [344730, Clone: SC-1, BioLegend], CD25-PE [12-0259-42, Clone: BC-96, Thermo Fisher Scientific] and CD54-AP700 [A7-429-T100, Clone: 1H4, EXB10]) and analyzed on a CytoFLEX flow cytometer (Software CyExpert 2.3, Beckman-Coulter). Activated and non-activated CD4 T cells were discriminated by means of their CD25 and CD54 expression, respectively.

Mouse breeding and experimentation

Inbred C57Bl/6J-2Kb (C57Bl/6) strain mice of specific genotypes were bred in house [33]. Wild-type female Balb/c:H-2Kd (Balb/c) mice were purchased from Jackson Laboratories (Bar Harbor, ME, USA) or Charles River Laboratories (Sulzfeld, Germany). Specifically, 12- to 14-week-old Balb/c mice were used as recipients and C57Bl/6 mice as grafts of bone marrow to be purified and transplanted. All mice were housed in a specific pathogen-free facility and had access to autoclaved food and drinking water ad libitum. Five days before preconditioning ionizing irradiation, drinking water and food pellets were mixed/soaked with antibiotics (neomycin, ampicillin, vancomycin and metronidazole, each at 0.33 g/L) and provided until termination of the experiments. All animal procedures were performed in accordance with the international guidelines for good laboratory practice and approved by the North Rhine-Westphalia State Agency for Nature, Environment and Consumer Protection (LANUV, reference numbers 84-02.04.2011.A319 and 84-02.04.2014.A494).

BM T-cell depletion

BM cells of C57Bl/6 donor mice were collected by flushing tibias and femurs by using 10-mL syringes (Terumo, Irvine, CA, USA) filled with cell culture medium (Dulbecco’s Modified Eagle’s Medium, 10% fetal calf serum [FCS], 1% penicillin-streptomycin, 1% antibiotic-antimycotic; Thermo Fisher Scientific) [33]. After red blood cell lysis by incubation in lysis buffer (Carl Roth, Karlsruhe, Germany) 1:10 for 3 min at room temperature, Thy1.2 cells were depleted via negative selection using the CD90.2 MicroBeads mouse Kit (Miltenyi Biotec, Bergisch Gladbach, Germany) according to the manufacturer’s instructions.

Naïve T-cell purification from spleens

For isolation of naïve CD4 T cells, spleens from C57Bl/6 mice were mashed through a 70-µm cell strainer (Thermo Fisher Scientific). After red blood cell lysis (1:10 dilution for 3 min at room temperature; Carl Roth), naïve CD4 T cells were purified via negative selection using the Naïve CD4 T-cell Isolation mouse Kit (Miltenyi Biotec) according to the manufacturer’s recommendations.

Experimental GVHD model

Recipient Balb/c mice were fed with antibiotic containing drinking water and antibiotic soaked chow starting 5 days (d5) before preconditioning. To induce GVHD at d0, recipient mice total bodies were subjected to a minimally lethal dose of 8 Gy of X-rays (source RS 320, X-Strahl Ltd., operated at 300 kV, 10 mA at a dose rate of 161.55 cGy/min). Unless stated otherwise, on d1 recipient mice were subjected to a minimally lethal dose of 8 Gy of X-rays (source RS 320, X-Strahl Ltd., operated at 300 kV, 10 mA at a dose rate of 161.55 cGy/min). Unless stated otherwise, on d1 recipient mice received intravenously 5 × 10⁶ T-cell–depleted C57Bl/6 BM cells and 0.7 × 10⁶ purified C57Bl/6 splenic naïve CD4 T cells. The symptomatic mice with GVHD were treated with aliquots of selected MSC-EV preparations at 3 consecutive days, at d7, d8 and d9, applied intravenously as 300-µL sterile saline suspensions.

Mouse weights and clinical scores were documented daily until principal experiment termination at d11. Towards clinical scoring, five parameters were monitored [34], namely, weight change, posture (hunching), activity, fur texture and skin integrity, wherein each value’s maximal score was 2. Animals reaching the cumulatively maximal score of 10 were sacrificed immediately by enforced isoflurane anesthesia whereas otherwise sacrifice was performed at d11. Blood was sampled immediately by retro-orbital bleeding and collected in both, ethylenediaminetetraacetic acid tubes (VWR, Radnor, PA, USA) for blood cell counting and heparin tubes (BD Biosciences) to gain plasma for cyto- and chemokine content analysis. Specifically, bead-based Luminex assay systems (R&D Systems, Minneapolis, MN,
USA) specific to keratinocytes-derived chemokine, tumor-necrosis factor alpha, interleukin-6 and granulocyte colony-stimulating factor were applied according to the manufacturer’s protocols on a Luminex 200 instrument equipped with the XPenon software (Luminex,ustin, TX, USA).

Colon analyses

Colons of d11 sacrificed mice were prepared and flushed with PBS to remove faces. The colons were carefully flattened with a thin wooden stick, rolled up into a Swiss Role as described before [35] and fixed in 4% formalin in a tissue cassette. The samples were then embedded in paraffin and 5-μm thin sections cut via a microtome. These tissue sections were then stained with hematoxylin and eosin and analyzed by light microscopy (Axiovert 40C; Zeiss, Jena, Germany). The pathology was evaluated and graded from 0 to 4, with 0 indicating no pathology and 4 indicating severe pathology, based on crypt and microvilli integrity (histological score).

Regulatory T cells (Treg), within colons were investigated as reported previously [36]. In brief, prepared colons were flushed with phosphate-buffered saline (PBS). cut into pieces and washed with PBS containing 3 mmol/L ethylenediaminetetraacetic acid (Sigma-Aldrich, St. Louis, MO, USA) for 10 min at 37°C. Thereafter, two washing steps with RPMI-1640 ( Gibco) containing 1% FCS, 1 mmol/L ethyleneglycol tetraacetic acid and 1.5 mmol/L MgCl2 (Sigma-Aldrich) were performed. Subsequently, the colonic pieces were digested in RPMI-1640 medium containing 20% FCS and 100 μM collagenase IV (Sigma-Aldrich). Single cells were separated from the remaining tissue by filtration through 40-μm cell strainers and washed with RPMI-1640 thereafter. Cells were surface-stained with Fixable Viability Dye (FVD) eFluor 780 (Thermo Fisher Scientific) and anti-CD4-PE antibodies (clone H129.19; BD Biosciences). Afterwards, intracellular staining with anti-FoxP3-FITC antibodies (clone: FJK-16; Thermo Fisher Scientific) was performed with the BD Cytofix/Cytoperm Kit (BD Biosciences) according to the manufacture instructions. Labeled cells were analyzed on a FACS Canto flow cytometer (BD Biosciences). Tregs were identified as FVD-CD4+FoxP3-cells.

Proteomic analyses

For proteomic profiling, MSC-EV sample volumes containing 5 μg of protein were used and subjected to in-solution tryptic digestions using a modified version of the Single-Pot Solid-Phase-enhanced Sample Preparation (SP3) protocol [37,38]. 1% SDS-containing lysates were added to Sera-Mag Beads (Thermo Fisher Scientific) in 10 μL of 15% formic acid and 30 μL of ethanol. Binding of proteins was achieved by shaking for 15 min at room temperature. Sodium dodecyl sulfate was removed by four subsequent washes with 200 μL of 70% ethanol. Proteins were digested overnight at room temperature with 0.4 μg of sequencing-grade modified trypsin (Promega, Walldorf, Germany) in 40 μL of HEPES/NaOH, pH 8.4 in the presence of 1.25 mmol/L tris (2-carboxyethyl) phosphine and 5 mmol/L chloroacetamide (Sigma-Aldrich). Beads were separated, washed with 10 μL of an aqueous solution of 2% dimethyl sulfoxide (DMSO), and the combined eluates were dried down. Peptides were reconstituted in 10 μL of H2O and reacted for 1 h at room temperature with TMT16pro labeling reagent (Thermo Fisher Scientific). To this end, 50 μg of TMT16pro label reagent was dissolved in 4 μL of acetonitrile and added to the peptides. Excess TMT reagent was quenched by the addition of 4 μL of an aqueous 5% hydroxylamine solution (Sigma-Aldrich). Peptides were reconstituted in 0.1% formic acid, and equal volumes were mixed. Mixed peptides were purified by a reverse phase clean-up step (OASIS HLB 96-well μElution Plate; Waters GmbH, Eschborn, Germany). Peptides were subjected to an off-line fractionation under high pH conditions [37]. The resulting 12 fractions were then analyzed by LC-MS/MS on an Q Exactive Plus mass spectrometer (Thermo Fisher Scientific).

To summarize, peptides were separated using an UltiMate 3000 RSLC (Thermo Fisher Scientific) equipped with a trapping cartridge (Precolumn; C18 PepMap 100, 5 μm, 300 lm i.d. × 5 mm, 100 Å) and an analytical column (Waters nanoEase HSS C18 T3, 75 lm × 25 cm, 1.8 lm, 100 Å). Solvent A: aqueous 0.1% formic acid and 3% DMSO; Solvent B: 0% formic acid and 3% DMSO in acetonitrile (all solvents were of liquid chromatography—mass spectrometry grade). Peptides were loaded on the trapping cartridge using solvent A for 4 min with a flow of 30 μL/min. Peptides were separated on the analytical column with a constant flow of 0.3 μL/min applying a 1-h gradient of 2–80% of solvent B in A. Peptides were directly analyzed in positive ion mode applying with a spray voltage of 2.2 kV and a capillary temperature of 275°C using a Nanospray-Flex ion source and a Pico-Tip Emitter 360 lm OD × 20 lm ID; 10 lm tip (New Objective). MS spectra with a mass range of 375–1200 m/z were acquired in profile mode using a resolution of 70,000 (maximum fill time of 250 ms or a maximum of 3e5 ions [automatic gain control, AGC]). Fragmentation was triggered for the top 10 peaks with charge 2–4 on the MS scan (data-dependent acquisition) with a 30-s dynamic exclusion window (normalized collision energy was 30). Precursors were isolated with a 0.7 m/z window, and MS/MS spectra were acquired in profile mode with a resolution of 35,000 (maximum fill time of 120 ms or an AGC target of 2e5 ions). For data analysis, protein output file of IsobarQuant (protein.txt – files) was processed using the R programming language (ISBN 3-900051-07-0). Only proteins that were quantified with at least two unique peptides were considered for the analysis. The raw TMT reporter ion intensities (signal_sum columns) were first cleaned for batch effects using limma [39] and further normalized using variance stabilization normalization [40]. Proteins were tested for differential expression using the limma package. The replicate information was added as a factor in the design matrix given as an argument to the ‘limFit’ function of limma. A protein was annotated as a hit with a false discovery rate (fdr) smaller 5% and a fold-change of at least 100% (ratio of 2) and as a candidate with a fdr below 20% and a fold-change of at least 50% (ratio of 1.5) [41]. The proteomic data are shared to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD032744”. Considering only proteins in the downstream analysis that have been identified with at least two unique peptides, a total of 991 proteins were identified.

miRNA analyses

Total RNA was extracted from 100-μL MSC-EV preparation aliquots. Before using the mirNeasy Mini Kit (Qiagen, Hilden, Germany), sample volumes were adjusted to 200 μL with 100 μL of nuclease-free water (Qiagen). Samples were mixed with 1000 μL of Qiazol, to which 1 μL of reconstituted spike-in controls (miRCURY; Qiagen) was added. EV samples were homogenized by vigorous mixing on a vortex for 15 s. Following incubation at room temperature for 5 min, 200 μL of chloroform was added to each lysate. Next, samples were centrifuged at 12 000g for 15 min at 4°C and 650 μL of the upper aqueous phase transferred to a miRNAeasy mini column. Samples were supplemented with 7 μL of glycogen (5 mg/mL) to a final concentration of 50 μg/mL. Following addition of 975 μL ethanol, RNA was precipitated followed by automated washing with RPE and RWT buffer in a Qiacube liquid handling robot (Qiagen). Total RNA was eluted in 30 μL of nuclease-free water and stored at −80°C until further analysis.

Due to the presence of heparin in the cell culture media and hence in EV preparations, for reverse transcription (RT)-quantitative polymerase chain reaction (PCR)-based microRNA screening a heparinase digest was performed using 10× heparinase buffer (200 mmol/L Tris, 500 mmol/L NaCl, 40 mmol/L CaCl2). Heparinase (200 U/mL; Thermo
Fisher Scientific) treatment was conducted for 3 h at 25°C by mixing 10 μL of obtained RNA samples with 10 μL of heparinase solution. Heparinase-treated total RNA samples were used for cDNA synthesis using the miRCURY RT kit (Qiagen). Reaction conditions were set according to recommendations by the manufacturer. In total, 18 μL of total RNA was used per 90 μL of RT reaction. To monitor RT efficiency and the presence of impurities with inhibitory activity, a synthetic RNA spike-in (cel-miR-39-3p) was added to the RT reaction. PCR amplification was performed in a 384-well plate format using a Roche LC480 II instrument (Roche, Germany) and miRCURY SYBR Green mastermix (Qiagen). Plates were incubated at 4°C for 1 h before PCR amplification, using the following settings: 95°C for 2 min, 45 cycles of 95°C for 10 s and 56°C for 60 s, followed by melting curve analysis. To calculate the cycle of quantification values (Cq-values), the second derivative method was used. Cq-values were normalized to the geometric mean of all five spike-in controls by subtracting the individual miRNA Cq-value from the Cq mean calculated for that sample. Subsequently, normalized dCq-values were adjusted for differences in the EV particle concentration. Principal component analysis was performed using ClustVis v2.0. Differential expression analysis was performed in Prism version 9.2.0 (GraphPad Software, San Diego, CA, USA) using multiple unpaired t tests assuming individual variance for each row. Multiple comparisons were adjusted for using fdr with a desired fdr of <.20%.

Statistical analysis

Prism 8 (GraphPad Software) was used for data presentation and statistical analysis. Differences between groups were determined by one-way analysis of variance (one-way analysis of variance) with Tukey, Dunnett or Kruskal–Wallis multiple comparisons test, comparing the mean of each group with the mean of every other group. P-values of less than 0.05 were considered to represent statistical significance.

Results

Co-transplantation of allogeneic T-cell–depleted BM and splenic naïve T cells robustly induces GVHD pathology

Intending to optimize the production of MSC-EVs for the treatment of patients with steroid-refractory acute GVHD, at first a murine GVHD model was improved for the preclinical testing of obtained MSC-EV preparations. To mimic the clinical allogeneic BM transplantation setting while controlling the severity of the subsequent GVHD responses very robustly, a strategy was chosen combining the stratifying aspects of previously described protocols. To this end, it was decided to co-transplant allogeneic T-cell–depleted BM [42] with a defined number of allogeneic isolated naïve T cells into myeloablated mice [43].

Balb/c mice were chosen as recipients and C57Bl/6 mice as MHC depleted BM and splenic naïve T cells. In order to trigger experimental GVHD by allogeneic BM transplantation in a controlled manner, 5 × 10^6 T-cell–depleted C57Bl/6 BM cells were co-transplanted with none, 0.125, 0.25, 0.5 or 1.0 × 10^6 spleen-extracted naïve T cells of the same donor mice, respectively. Then, 24 h before receiving the allogeneic transplants, Balb/c mice were myeloablated by total body irradiation. For controlling the irradiation effectiveness, additional mice were irradiated without receiving any transplants.

After transplantation, mice (five per group) were scored every 24 h in respect to body weight, habitus and activity as well as fur and skin appearance. As expected, the numbers of co-transplanted T cells correlated with GVHD symptom severities (Figure 1A). In detail, mice suffering from aGVHD showed weight loss, a hunched position, low activity, loss of fur and scurf on hair free areas. The severity of the symptoms was summarized as clinical score. Irradiated control mice not receiving any transplant suffered from severe weight lost and had to be sacrificed preterm at d9 after irradiation. In contrast, irradiated control mice that were only transplanted with T-cell–depleted BM recovered completely from a temporal weight reduction until d11. In mice receiving BM as well as T cells, the GVHD severity, as indicated by the clinical core and weight loss, increased with the number of co-transplanted T cells (Figure 1A). Host animals co-transplanted with the maximal number of applied T cells (1 × 10^6) developed such strong GVHD symptoms that they had to be sacrificed before the scheduled experimental endpoint at d11 (Figure 1A).

The number of co-transplanted T cells correlated with the severity of the colon GVHD. Specifically, loss of crypts and microvilli as well as the numbers of infiltrating cells and thus intestinal tissue swelling correlated with the co-transplanted T-cell dosages. Mice receiving the maximal dose of 1 × 10^6 CD4 T cells lacked virtually all crypts and microvilli (Figure 1B,C). Blood analyses pointed in the same direction in that serum keratinocytes-derived chemokine (CXCL1), tumor-necrosis factor alpha, interleukin-6 and granulocyte colony stimulating factor concentrations increased with the number of co-transplanted T cells (Figure 1D). An inverse correlation was observed for the number of platelets per blood volume (Figure 1E). Overall, the obtained results qualify the allogeneic co-transplantation of T-cell depleted BM and defined numbers of purified splenic naïve T cells as a tunable strategy to robustly and reproducibly induce GVHD symptoms in mice.

Applied MSC-EV preparations can modulate the severity of the GVHD symptoms in the established aGVHD model

To test whether MSC-EV preparations can modulate the severity of GVHD symptoms in the established mouse model, 5.0 × 10^6 T-cell–depleted C57Bl/6 BM cells were transplanted together with 0.7 × 10^6 pure naïve splenic T cells (Figure 2A). For the proof-of-principle experiments, an MSC-EV preparation (41.5-EV_52) was applied, which exhibited therapeutic efficacy in a murine ischemic stroke model [32] and CD73 activity in vitro [26]. Specifically, at d7, d8 and d9 post-transplantation mice were either treated with PBS or with aliquots of the 41.5-EV_52 preparation containing the EV-harvest from CM of 1.2 × 10^6 MSCs (three mice per group). Mice were continuously monitored until sacrifice at d11 post-transplantation (Figure 2A). The phenotypic appearances of animals were documented and blood samples as well as gut biopsies analyzed. In contrast to PBS-treated mice with aGVHD, 41.5-EV_52-treated mice with aGVHD increased their weight successively and displayed an appearance similar to merely BM transplanted and thus healthy-appearing mice (Figure 2B). The colon morphologies of 41.5-EV_52-treated mice were massively improved compared to those of PBS-treated mice with aGVHD (Figure 2C,D). Furthermore, the blood platelet number of 41.5-EV_52-treated mice was increased in comparison to that of PBS-treated mice with aGVHD (Figure 2E). Thus, the aGVHD symptoms of mice co-transplanted with BM and allogeneic T cells were successfully suppressed by treatment with the given MSC-EV preparation.

Independent MSC-EV preparations differ in their in vitro immunomodulatory capabilities

In the past, we have improved up-stream and down-stream processes for the manufacturing of MSC-EV products from up to 10 L of MSC-CM [13,14]. Applying our scaled production strategy, we have prepared EVs from CM of a variety of different donor MSCs. Regularly, starting with samples of the same donor MSC stock, independent MSC-EV batches of several donor MSCs were produced. Several of the obtained MSC-EV preparations have been applied to various disease models. In most of our published studies, we focused on MSC-EV...
preparations, which were considered as therapeutically active. Indeed, applying such MSC-EV preparations, therapeutic effects were observed in all models tested [28,44].

More recently, inspired by an advanced concept for a functional assay [48], we have set up an mdMLR assay [26], whose details are being reported by a back-to-back publication of this manuscript [27]. Over time, we have evaluated several of our MSC-EV preparations in this assay. The results revealed that some MSC stocks allow production of active MSC-EV preparations with a higher probability than others. For example, although eight of 10 EV preparations of the MSC 41.5 stock (internal identification numbers related to donors) modulated immune responses in the mdMLR assay, only two of 10 independent EV preparations of MSC 16.3 stock showed comparable activities (data not shown). Notably, the activities of MSC-EV preparations from other MSC stocks appear as more balanced.

Intending to evaluate the therapeutic potential of various MSC-EV preparations in the established GVHD model and to explore potential correlations between immunomodulatory in vitro and therapeutic in vivo capabilities, we decided to compare the in vivo therapeutic activity of different preparations of the MSC 41.5 stock, two of which showed activities in the mdMLR assay and one that was lacking this activity, as well as another active and two inactive MSC-EV preparations, which were considered as therapeutically active. Indeed, applying such MSC-EV preparations, therapeutic effects were observed in all models tested [28,44].

Fig. 1. Experimental murine GVHD severity depends on the amount of splenic allogenic T cells co-transplanted with T-cell-depleted BM cells. (A) Clinical score and percentage weight change of experimental groups (n = 5 animals per group) that received indicated numbers of splenic CD4 T cells. The irradiation effect was controlled on irradiated mice not receiving allogeneic BM transplants (IIR only). (B) Hematoxylin and eosin–stained colon sections of a representative animal of each transplantation group; arrows point on colon crypts with structural integrity (two left panels) or areas in which colon structures had been replaced by infiltrating cells (two right hand panels; scale bars represent 50 µm). (C) Cumulative histology scoring result of the gut analyses (n = 3 per experimental group). (D) Plasma cytokine concentration and (E) whole-blood platelet count in whole blood (n = 3 per group; IIR, ionizing irradiation; T, numbers of co-transplanted allogenic T cells; BM, allogeneic T-cell-depleted bone marrow cell transplantation; BM+T, bone marrow and T cell co-transplantation; PLT, platelets; mean values ± standard error of the mean; analysis of variance ***P < 0.001; **P < 0.01).
preparations of MSC stocks MSC 87 or MSC 16.3 and MSC 84, respectively. For the sake of coherence, the MSC-EV preparation nomenclature involved an index reflecting the mdMLR interfering activity score. Accordingly, “a” in EVa indicates the capability, and “i” in EVi the incapability to modulate immune responses in the mdMLR assay [26]. The six MSC-EV preparations studied in this manuscript are 41.5-EVa2, 41.5-EVb3, 41.5-EVc2, 16.3-EVd, 84-EVe, and 87-EVf (of note 41.5-EVa1 and 41.5-EVb1 were studied previously [26] but not in this study).

For increasing the comparability of all data in this manuscript, metric analyses as recommended in the community [30,49] as well as the mdMLR analyses were repeated for the selected MSC-EV preparations in parallel to each other. The MSCs characteristics, e.g., their cell surface phenotype and their osteogenic and adipogenic differentiation capability, were analyzed as well (supplementary Figure 1).

Particle concentrations of the MSC-EV preparations as measured by nanoparticle tracking analysis that varied from 1.2×10^8 to 3.8×10^8 particles per μL of MSC-EV preparation, which corresponded to the particle yield of supernatants of 4×10^7 MSCs (Table 1). The protein concentration of MSC-EV preparations varied from 4.80 to 7.84 μg/μL, resulting in particle concentrations per μg of protein ranging from 2.1×10^7 to 6.7×10^7 (Table 1). Thus, the particle and protein concentrations of obtained MSC-EV preparations are all in a similar range. The average size distribution of prepared particles ranged from 114.2 to 132.2 nm (Table 1). As confirmed by western blot, all MSC-EV preparations tested contained the tetraspaninns CD9, CD63 and CD81 as well as syntenin and according to the expectations lacked calnexin (supplementary Figure 2).

Although none of the parameters appeared ideal for the normalization of sample amounts to be applied into the mdMLR assays, due to their comparable metrics (Table 1), volumes of the selected MSC-EV preparations containing 25 μg of proteins were applied. As MSCs are raised in hPL-supplemented media (three independent hPL batches were used in the course of the project), which had not been EV depleted, EV preparations of fresh hPL-supplemented media (PL1-EV, PL2-EV or PL3-EV) served as controls (Table 1).

In the absence of any EV preparation, mdMLR assay cultures contained more than 30% (31.3–40.3%) activated CD4 T cells and more than 40% activated CD8 T cells (41.1–55.0%) (Figure 3, supplementary Figure 3). The application of 25 μg of the PL-EVs control samples did not significantly alter the amount of activated CD4 and CD8 T cells (Figure 3, supplementary Figure 3). Also, no severe impacts on the percentage of activated CD4 and CD8 T cells were recorded when MSC-EV preparations 41.5-EVb2 (36.5%; 50.6%), 16.3-EVc (32.3%; 47.5%) or 84-EVf (37.4%; 55%) were applied (Figure 3; supplementary Figure 3). In contrast, clear reductions in activated CD4 and CD8 T-cell contents were observed when MSC-EV preparations 41.5-EVc2 (12.2%; 28.3%), 41.5-EVb2 (21.3%; 31.3%) or 87-EVf (15.0%; 34%) were applied (Figure 3; supplementary Figure 3). Thus, the results of the
mdMLR analysis reproduce previous evaluations and confirm that independent MSC-EV preparations, even if they derive from independent batches of the same MSC stock, can considerably differ in their in vitro immunomodulatory capabilities.

**MSC-EV preparations with in vitro immunomodulatory capabilities modulate GVHD symptoms in vivo**

Next, it was investigated whether the in vitro immunomodulatory properties of the given MSC-EV preparations reflect their therapeutic potency in suppressing aGVHD symptoms in GVHD mice. Consequently, aliquots of five different MSC-EV preparations (MSC-EV preparations 16.3-EV, 41.5-EVα2, 41.5-EVα3 and 87-EVα) and as controls two PL-EV preparations (PLα-EV and PLβ-EV) were applied according to the experimental aGVHD protocol established here (Figure 2A). Each individual EV preparation was applied to 5 to 10 overt GVHD mice.

The results of the in vivo study qualify MSC-EV preparations 41.5-EVα2, 41.5-EVα3 and 87-EVα as effective, since they improved aGVHD clinical and gut histological scores significantly (Figure 4A). In contrast, 41.5-EVα2 preparation was largely ineffective. For unknown reasons, MSC-EV preparations 16.3-EV increased the aGVHD symptoms in four of 10 animals to a degree at which they needed to be sacrificed before the scheduled experiment’s end (Figure 4A). Analyses of intestinal specimen were performed representatively for three or four mice either being treated with PBS or the MSC-EV preparations 41.5-EVα3 or 16.3-EV at d11, respectively. Within the gut specimen from GVHD mice that had received the 41.5-EVα3 preparation, the colon tissue appearance was similar that of non-GVHD control mice (Figure 4B). In contrast, administration of 16.3-EV preparation apparently had no therapeutic effect on the gut morphology in the three mice analyzed histologically. In these animals as well as in the PBS-treated control GVHD animals, the colon crypts were almost completely destroyed, and the guts revealed a huge proportion of necrotic tissue (Figure 4B,C).

In addition, the frequencies of Tregs in intestinal biopsies and weights of spleens of these mice were determined. Upon MSC-EV application, the abundance of intestinal Tregs was comparably increased, independent of the therapeutic capabilities the individual MSC-EV preparations finally achieved on the aGVHD symptomatology (Figure 4D). However, the spleen weights of MSC-EV treated and control GVHD mice were distinguishable. After treatment with the 41.5-EVα3 preparation, mice had up to four times heavier spleen weights than GVHD mice treated with the MSC-EV 16.3-EV preparation (Figure 4E). Thus, although MScs were raised under identical cell culture conditions and EVs prepared with a standardized protocol, the obtained data confirm functional differences among independently produced MSC-EV preparations, including batch-to-batch variations.

**MSC-EV preparations modulate GVHD symptoms in dose-dependent manners**

So far, all MSC-EV preparations were applied at doses of 1.2 × 10⁶ cell equivalents per mouse. Against the background of an average body weight of 25 g, this reflects the EV equivalents of ~4.8 × 10⁷ cells per kg body weight. In contrast to the doses that were applied here, we previously applied roughly 5 × 10⁵ cell equivalents per kg of body weight to a patient [14]. Thus, compared with the successfully treated patient with aGVHD, ~100-fold increased MSC-EV doses were applied to the aGVHD mice. To evaluate dose effects of the applied MSC-EV preparations, the active 41.5-EVα3 and the inactive 16.3-EV, preparations were applied in comparison to the original dose in 3-fold lower and 3-fold higher doses. Notably, reduction of the 41.5-EVα3 preparation dose reduced the strength of the observed therapeutic effect. In contrast, the application of the three-time doses did not reveal any further improvement (Figure 5A). This implies that threshold activities exist, which had been reached for the 41.5-EVα3 preparation with the original dose. As expected, reduction of 16.3-EV preparation had no impact on GVHD symptoms, while the triple dose slightly improved the symptoms, almost to the same extent as the 1/3 dose of 41.5-EVα3 preparation (Figure 5A,B).

**Conventional EV characteristics fail to distinguish active and inactive MSC-EV preparations**

Attempting to identify a potential surrogate marker representing an efficient MSC-EV preparation, we performed comparative analyses of all preparations in respect to EV-typical parameters, namely particle and protein concentration and the intensities of the CD9, CD63, CD81, syntenin and calnexin-specific bands on western blots (Table 1; supplementary Table 2; supplementary Figure 2). However, no coincidence was observed. For example, the 41.5-EVα3 preparation revealed a comparable protein concentration than the 16.3-EV preparation but showed a lower particle concentration (Figure 6A). In contrast, the abundance of exosomal marker proteins was higher for the 41.5-EVα3 preparation than for the 16.3-EV preparation (Figure 6B). Notably, hardly any differences in the band intensities of the exosomal proteins were found among 41.5-EVα2 and 41.5-EVα2 preparations. Although the 41.5-EVα2 preparation contained more particles with a higher purity index than the 41.5-EVα2 preparation (Figure 6A), 41.5-EVα2 preparation was identified as being capable of modulating GVHD symptoms at the applied doses, whereas 41.5-EVα2 preparation lacked this capability. Thus, none of the applied markers provides any information about the MSC-EV preparations’ therapeutic potential.

---

Table 1

<table>
<thead>
<tr>
<th>EV batch</th>
<th>Particle, no/μL</th>
<th>Particle diameters, nm</th>
<th>Protein concn., μg/μL</th>
<th>Purity, particle no/μg protein</th>
<th>Applied volume (μL) in mdMLR, corresponding to 25 μg protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.3-EVα</td>
<td>3.3 × 10⁷</td>
<td>114.5</td>
<td>4.89</td>
<td>6.7 × 10⁷</td>
<td>5.14</td>
</tr>
<tr>
<td>41.5-EVα2</td>
<td>1.2 × 10⁷</td>
<td>132.2</td>
<td>5.61</td>
<td>2.1 × 10⁷</td>
<td>4.45</td>
</tr>
<tr>
<td>41.5-EVα3</td>
<td>3.8 × 10⁷</td>
<td>125.6</td>
<td>7.84</td>
<td>4.8 × 10⁷</td>
<td>3.19</td>
</tr>
<tr>
<td>41.5-EVα3</td>
<td>1.6 × 10⁷</td>
<td>125.6</td>
<td>4.80</td>
<td>3.3 × 10⁷</td>
<td>5.21</td>
</tr>
<tr>
<td>84-EVα</td>
<td>2.3 × 10⁷</td>
<td>114.5</td>
<td>5.12</td>
<td>4.5 × 10⁷</td>
<td>4.88</td>
</tr>
<tr>
<td>87-EVα</td>
<td>2.0 × 10⁷</td>
<td>114.2</td>
<td>5.80</td>
<td>3.4 × 10⁷</td>
<td>4.31</td>
</tr>
<tr>
<td>PLα-EV</td>
<td>6.4 × 10⁶</td>
<td>103</td>
<td>8.33</td>
<td>7.7 × 10⁷</td>
<td>3.00</td>
</tr>
<tr>
<td>PLβ-EV</td>
<td>3.5 × 10⁶</td>
<td>107.7</td>
<td>9.59</td>
<td>3.7 × 10⁷</td>
<td>2.61</td>
</tr>
<tr>
<td>PLβ-EV</td>
<td>8.2 × 10⁵</td>
<td>125.2</td>
<td>7.52</td>
<td>11.0 × 10⁷</td>
<td>3.32</td>
</tr>
</tbody>
</table>

EV, extracellular vesicles; mdMLR, multi-donor mixed lymphocyte reaction; MSC, mesenchymal stromal cells.

---

R.J. Madel et al. / Cytotherapy 00 (2023) 1–16

ARTICLE IN PRESS
Despite huge similarities and the failure in identification of potential surrogate markers, proteomic analyses discriminate active and inactive MSC-EV preparations.

Aiming to identify functionally relevant proteins and potential surrogate markers discriminating functional active and inactive MSC-EV preparations, multiplexed quantitative mass spectrometry analyses of the MSC-EV and PL-EV samples were performed that had been applied to the GVHD mice (16.3-EV, 41.5-EV, 41.5-EV2, 41.5-EV3, 84-EV, and 87-EV) were tested for their immunomodulatory capabilities in the mdMLR assay. Specifically, approximately 6x10⁵ cells of a mononuclear cell mixture derived from peripheral blood samples of 12 healthy donors were cultured for 5 days, either in the presence or absence of 25 µg of given EV preparations. Thereafter, cells were harvested and labeled with antibody cocktails containing anti-CD4, anti-CD25 and anti-CD54 antibodies. CD4 T cells were gated according to their light scatter features as 7AAD⁻, CD4⁺SSC⁻ cells (blue, A). The proportions of activated CD4 cells (CD25⁺CD54⁺) were determined by flow cytometric analyses (B). *: controls are identical (respective EV preparations were analysed in parallel at the same day).
Fig. 4. Independent MSC-EV preparations differ in their ability to inhibit GVHD symptoms in vivo. (A) Comparative analysis of ionizing irradiated (IIR) mice that after being co-transplanted (BM or BM+T) after onset of GVHD were treated with five independent MSC-EV preparations (blue, therapeutically effective MSC-EV preparations; red, therapeutically ineffective MSC-EV preparations; gray, hPL-EV control; green, PBS control; ↓, death or scarification at time point before intended experiment end as indicated). Clinical scoring based on five criteria (weight change, posture, activity, fur texture and skin integrity); weights are also illustrated separately. (B-E) Comparison of therapeutic effects of a specific effective (41.5-EVα) and a specific ineffective (16.3-EVβ) MSC-EV preparation on aGVHD symptoms: (B) Representative posture of irradiated BM-transplanted mouse (healthy appearance) or irradiated mice co-transplanted with BM and T cells either being treated with PBS or with 41.5-EVα or 16.3-EVβ as well as representative hematoxylin and eosin-stained colon sections of a representative mouse of each group. Arrows point on colon crypts or areas with infiltrating cells (scale bar: 50 μm). (C) Histology score of gut sections of n = 3 mice per group. (D) Proportion of Tregs among the population of colon residing CD4 T cells and (E) spleen weights of PBS- and MSC-EV–treated GVHD mice. (C-E) Representation of mean ± standard deviation for n = 3 mice per group. IIR, ionizing irradiation; Tn, numbers of co-transplanted allogenic T cells; BM, allogenic T cell-depleted bone marrow cell transplantation; BM+T, bone marrow and T cell co-transplantation; ns, nonsignificant. Analysis of variance ****P < 0.0001, ***P < 0.001, **P < 0.01, *P < 0.05).
samples of the three active and the three inactive MSC-EV preparations as well as of the three PL-EV preparations were also analyzed. A simple principal component analysis based on the 991 identified proteins revealed high similarities among the different MSC-EV preparations, which clearly differed from the similar-appearing PL-EV preparations (Figure 7A). Strikingly, this analysis revealed higher grades of similarity among the active or the inactive MSC-EV preparations, respectively, than between individual active and inactive MSC-EV preparations (Figure 7A). Thus, the functional discrimination of active versus inactive MSC-EV preparations is reflected by their proteome.

Upon comparing the proteome profiles for differential abundance using a moderated t-test (limma) [39], 69 proteins were identified whose abundance in MSC-EV preparations was clearly higher in every of the MSC-EV preparations than in the PL-EV preparations (fold-change ≥2 and fdr <0.05) (Figure 7B). Proteins with clearly different abundances among MSC-EVα and MSC-EVβ preparations, however, were not identified (Figure 7B, supplementary Table 3).

![Fig. 5.](image)

**Fig. 5.** MSC-EV preparations modulate GVHD symptoms in dose dependent manners. (A, B) Clinical score upon triple, single and third dose administration of (A) 41.5-EVα and (B) 16.3-EVα preparations to experimental aGVHD mice according to the regimen represented in Figure 2A (mean values ± standard deviation of n = 8–10; graphs depict cumulative results of two independent experiments. BM+T: bone marrow and T cell co-transplantation; IIR: ionizing irradiation; ns, non-significant; Analysis of variance. *** P < 0.001, * P < 0.05.

![Fig. 6.](image)

**Fig. 6.** Recorded molecular and physical features of different MSC-EV preparations do not correlate with their functional properties: (A) Particle and protein concentrations of the applied MSC-EV preparations were compared; resulting purity indices were calculated as particles per protein amount. (B) Intensities of specific western blot bands for selected proteins were evaluated semi-quantitatively applying ImageJ analyses on documented western blot images (supplementary Figure 3, supplementary Table 2). Blue, 41.5-EVα preparation exemplifying a confirmed active example; red, 16.3-EVβ, exemplifying a confirmed inactive preparation.
Previously, profiles of published MSC-EV proteome analysis were compared, resulting in the identification of 134 proteins that were recovered in at least seven of 10 given proteome profiles of human MSC-EV preparations, with 75 of them considered as MSC-EV specific. Out of those, 15 proteins were also discovered in other species MSC-EV preparations [50]. The remaining 59 proteins are more broadly expressed EV markers [50]. In the current MSC-EV profiles, 10 of the 15 MSC-EV proteins that were discovered in MSC-EV preparations of all species were recovered, whereof six were present only in MSC-EV samples and four of them also in PL-EV samples. Of the 60 human MSC-EV specific proteins, 10 were recovered exclusively in the given MSC-EV preparations and 36 also in the PL-EV samples. Of

![Proteomic analyses discriminate between active and inactive samples but fail to identify surrogate markers.](image-url)
the 59 more common EV markers, two appeared MSC-EV specific and 49 were also detected in the PL-EV samples. Thus, in total 107 of the previously reported 134 proteins were detected in the given MSC-EV samples (Figure 7C, supplementary Table 4).

MicroRNA (miRNA) cargo does not distinguish active from inactive MSC-EVs

Next, we screened for miRNAs with different abundances in active compared with inactive MSC-EV preparations. Applying a multiplex RT-PCR approach, we screened for the presence of 372 different miRNAs in the MSC-EV and PL-EV preparations that had been administered to the GVHD mice. Twelve spiked in synthetic miRNAs served as control. Recorded data were normalized to the concentration of spike in controls.

Of the 372 miRNAs 228 were detected in all MSC-EV and PL-EV preparations (n = 9) above background, and 314 miRNAs were detected in at least three of nine samples. Comparable as for the proteome analysis, a simple principal component analysis based on the 314 detected microRNAs revealed high similarities among the different MSC-EV preparations and clearly differentiated them from the PL-EV preparations which appeared very similar to each other (Figure 7D). In contrast to the proteome analysis, however, obtained miRNA profiles did not reveal higher grades of similarity among the active or the inactive MSC-EV preparations, respectively. In summary, we failed to identify any specific miRNA being significantly (fdr <0.05) more abundant in active than inactive MSC-EV preparations or vice versa.

Discussion

Even though we set up a standardized procedure for the manufacturing of MSC-EV products [13–15], obtained MSC preparations differed in their immunomodulatory capabilities in vitro [51]. Here, after establishing an optimized aGVHD mouse model in which defined numbers of allogeneic T-cell–depleted BM cells and defined numbers of allogenic T cells were co-transplanted into myeloablated mice for the very reproducible aGVHD induction, we demonstrated that only a proportion of their MSC-EV products provided activities being potent to suppress the induced aGVHD symptoms effectively. Remarkably, the potency to suppress the aGVHD symptoms in vivo correlates with their ability to suppress CD4 and CD8 T-cell activation in a novel MLR assay, the mdMLR assay [27]: only MSC-EV preparations being able to reduce the content of activated T cells revealed GVHD improving activities and vice versa. Notably, although some MSC stocks have a higher and others a lower incidence for obtaining MSC-EV preparations counteracting T-cell activation in this assay, all MSC stocks that were repetitively used for the MSC-EV production resulted in the manufacturing of both immunomodulatory active as well as inactive MSC-EV products, indicating batch-to-batch variation issues. Remarkably, despite the critical differences in the activity, we failed to identify any surrogate marker in a proteome profiling approach, which otherwise revealed the presence of a huge proportion of typical MSC-EV marker proteins that had been identified in a proteome comparison approach recently [50]. We also failed to identify any miRNA whose presence correlates with the MSC-EV preparations’ capabilities in suppressing T-cell activation in the mdMLR assay. Thus, despite the fact that independent MSC-EV preparations can appear very similar, they can differ significantly in their functional properties. Thus, as discussed in a recent position paper [24], the presented results underscore the need for appropriate potency testing of clinical MSC-EV products.

After reporting the success of MSC-EV treatment of a patient with treatment-refractory GVHD [14], a couple of other groups investigated the impact of MSC-EV treatment on murine aGVHD models [52]. Despite the fact that the GVHD models differ in some details, all groups reported improvement after MSC-EV administration and demonstrated some in vitro activities, e.g., impacts on the cytokine profile of stimulated peripheral blood mononuclear cells, their impact on T-cell activation or their Treg inductive potentials, respectively [53–58]. Although three of the studies compared the in vitro activities with fibroblast-derived EV preparations and one with that of the parental MSCs, only one study compared in vivo and in vitro effects of two different MSC-EV products. Here, EVs were prepared from CM of MSCs either raised in FBS- or hPL-supplemented media [53]. For reducing inter-donor variabilities, five independent EV preparations of each type were manufactured and pooled for detailed functional analyses; activities of individual MSC-EV preparations were not compared [53]. Thus, the current study apparently comprehensively compares the functional activity of independent MSC-EV preparations that had been produced with the same standardized method for the first time.

We do not have any concrete mechanistic explanation yet as to why active and inactive MSC-EV preparations confer different activities. Since the MSC-EV production strategy has been highly standardized in our laboratory and most of the MSC-EV preparations were conducted by two laboratory members only, handling derived variations in the EV product qualities appear unlikely. Assumedly, variations likely depend on the MSC stocks being used and on stochastic events. MSC propagations are regularly polyclonal and have been reported to undergo clonal selection procedures [59]. Apparently, potentially in donor-specific manners, MSC stocks contain variable proportions of MSCs promoting or suppressing immunomodulatory activities of resulting EV preparations. Assumedly, the initial proportion of such MSC subtypes, combined with stochastic events during clonal selection procedures, are decisive for the immunomodulatory in vitro capability of resulting MSC-EV products. Thus, with a certain probability, MSC stocks with a high prevalence for MSC-EV products with immunomodulatory capabilities can also result in products lacking these capabilities and vice versa. In this context, it is worth mentioning that phenotypic and functional heterogeneity of MSCs has been reported for more than two decades now [16–19]. However, discussions about the impact of MSC heterogeneities, their origin and underlying culture conditions on their therapeutic potentials are just emerging [20–22]. Congruently, to the unsuccessful identification of any surrogate marker that reflects the immunomodulatory activity of tested MSC-EV samples, no generally accepted surrogate marker has been qualified in the MSC field that predicts the potency of cellular MSC products [21,22].

The issue in the EV field is further complicated by the fact that, for now, no generally accepted MSC-EV production strategies exist. As summarized recently, almost all groups follow their own production strategy, in using variable starting material, different media including supplements, different expansion and conditioned media harvesting strategies, different EV preparation technologies etc. [60]. As discussed in a couple of white papers in more detail, the MSC-EV production process is multiparametric, and many parameters seem to affect the functional and molecular properties of the obtained product, its metrics and its potency [24,49,61,62]. The issue is best summarized in the phrase “the process defines the product.” It is not known yet whether the proportions of active to inactive MSC-EV preparations would differ with other MSC-EV production protocols; however, using a standardized production strategy, here, we clearly demonstrate the existence of functional differences among independent MSC-EV preparations.

The issue of heterogeneity and the impact of production strategies on the potency of MSC products is highlighted by the outcomes of two phase 3 clinical trials on MSC application in patients with aGVHD and the subsequent application for market authorization and decision by the Food and Drug Administration. While a first phase 3 clinical trial failed to show efficacy of a MSC product in adult patients with aGVHD, assumedly due to MSC over-expansion issues [8], the
second phase 3 clinical trial with a MSC product, named remestemcel-L, on pediatric patients with aGVHD revealed efficacy [63,64]. Despite the success of the second trial, the Food and Drug Administration declined in October 2020 approval of remestemcel-L for pediatric aGVHD treatment in the United States. They argued that as primary MSC product of changing donors remestemcel-L needs to be considered to vary in its potency and that potency testing was not addressed appropriately (https://www.fda.gov/media/140988/download). Thus, it is evident that appropriate potency testing is mandatory for the successful translation of cellular as well as EV products into the clinics.

However, potency testing is challenging itself. MSC-EVs apparently act in a multinoval manner, e.g., they can exert pro-angiogenic, pro-regenerative and anti-apoptotic functions [60,65], and the importance of each of the modalities seems to depend on the target disease [24]. Suggestively, the successful treatment of any given disease requires different MoA modalities, and as soon as one of them is missing, the therapeutic function might become lost. Previously, we have demonstrated that active as well as inactive MSC-EV preparations that had been independently manufactured from the same MSC stocks revealed comparable activities in cleaving pro-inflammatory extracellular adenosine monophosphate into anti-inflammatory adenosine and free phosphate via the ecto-5′-nucleotidase activity of CD73 [26]. Here, it was observed that MSC-EV treatment, independent of the therapeutic potency of the applied MSC-EV preparations, comparably increased the abundance of Tregs in guts of mice with GVHD. Thus, even inactive and active MSC-EV preparations contain comparable activities in both of these modalities, another MoA modality is required for the therapeutic effect. Very likely, the same activity is also critical for suppressing T-cell activation in vitro. For now, we do not know of how many critical activities are required for mediating the observed therapeutic effect. However, the access to potent and non-potent MSC-EV preparations provides an excellent starting position for dissecting MoA attributes in the near future. To learn whether CD73 ecto-5′-nucleotidase activity is essential or dispensable for the therapeutic effect, it can be investigated whether suppression of this activity affects the therapeutic potency of otherwise active MSC-EV preparations.

In principle, a collection of potent and non-potent MSC-EV preparations also is an ideal starting point for identifying surrogategs, some of which might be causally linked to the critical MoA modality. In this study, we were not able to successfully identify any concrete surrogate marker, neither by protein nor by miRNA profiling. Eventually, the abundance of HPL-derived EVs in the given MSC-EV preparations is too high to such identify differences on MSC-derived EVs. In the future, repetition of such analyses on selected EV subfractions, e.g., obtained by bead capturing, might increase the resolution for the successful surrogate marker identification. More efforts are certainly required to identify the critical differences.

Despite the limitations in having identified any molecular surrogates, the overall results of the current study should help to increase the awareness of product heterogeneity in the EV field and contribute to the avoidance of failed clinical trials as they occurred in the MSC field [8].

Author Contributions

Conception and design of the study: CK, BG. Acquisition of data: RM, VB, RD, MB, TT, NP, FS, SS, MH. Analysis and interpretation of data: all authors. Drafting and revising the manuscript: RM, VB, CK, BG. All authors have approved the final article.

Declaration of Interest Statement

BG is a scientific advisory board member of InnoveX Therapeutics SL, Mursla Ltd., PL BioScience and ReNeuron, a consultant of Fujifilm and a founding director of Exosla Ltd. All other authors have no commercial, proprietary or financial interest in the products or companies described in this article.

Funding

This study was supported by the LeitmarktAgentur.NRW and the European Union (European Regional Development Fund 2014-2020, EFRE-0800396), the Stem Cell Network North Rhine Westphalia and (ERA-NET EuroTransBio 11: EVTrust [031B0332B]; EU COST program ME-HaD [BM1202]) as well as from the European Union’s Horizon 2020 research and innovation program EVPRO [H2020-NMBP-TR-IND-2018]. The materials presented and views expressed here are the responsibility of the authors(s) only. The EU Commission takes no responsibility for any use made of the information set out.

Acknowledgments

The authors thank the Westdeutsche Spender Zentrale (WSZE) for providing bone marrow samples of healthy donors. They are also grateful to all healthy blood donors whose cells were used in the mdMLR assay. The authors thank the Imaging Center Essen (IMCES; Institute for Experimental Immunology and Imaging), especially Bernd Walkenfort and Dr. Mike Hasenberg, for method development and acquisition of the transmission electron microscopy images. They also thank Thomas Scholtyssik for technical support in the GVHD experiments.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.jcyt.2023.03.008.

References
