

The EV-TRACK summary add-on: integration of experimental information in databases to ensure comprehensive interpretation of biological knowledge on extracellular vesicles

Introduction

Released in multiple body fluids, EVs protect their content from degradation and are increasingly considered for the development of novel clinical applications such as liquid biopsy tests [1–3]. The development of these tests is currently mainly driven by in-depth analysis of the protein and RNA cargo of EVs using omics approaches to identify biomarkers for disease. In addition, researchers aim to delineate EV functions in (patho)physiological conditions by integrating this knowledge on EV cargo [4].

Detection of EV-associated protein and RNA is highly valuable [5–7] but remains a challenge. Not all extracellular RNA and proteins are associated with EVs. Other extracellular macromolecular structures overlapping in size and/or density with EVs, such as protein aggregates, ribonucleoproteins and lipoprotein particles, contain RNA and proteins and are frequent contaminants in EV preparations [8–11]. Both the EV source and the method of choice determine the degree of specificity to which these contaminants can be separated from EVs [12]. A multitude of methods have become available to separate EVs from biofluids but each method achieves this with different specificity and efficiency, resulting in method-dependent identification of EV cargo [13,14]. To allow for the interpretation of contaminant-induced bias and to ensure reproducibility, transparent reporting of EV separation and characterization is crucial. To promote transparent reporting and reproducibility we released the open-source knowledgebase EV-TRACK that centralizes (meta) data of EV separation and characterization [13]. Currently, EV-TRACK includes experimental parameters of 2165 EV experiments from 1355 publications. For each experiment, the completeness of reporting the generic and method-specific information that facilitates interpretation and reproduction of the experiment is assessed by a checklist, summarized into the EV-METRIC (13; evtrack.org/about.php). Supported by the community, EV-TRACK has been included in the 2018 update of the MISEV guidelines (Minimal Information for Studies of Extracellular Vesicles) [15].

To enhance validation of EV-associated biomarkers and functions and, in general, to centralize knowledge on EV cargo, a multitude of databases have been created. EV-contained RNA and/or proteins are accessible on specialized databases such as EVpedia, Vesiclepedia, Exocarta, and more recently ExoRbase and EVmiRNA [16–20]. These databases are driven by community annotation and present cargo information retrieved using a variety of separation methods. Dependent upon the specificity of the method, this cargo thus associates with differential likelihood to EVs or extracellular macromolecular structures. As such, one of the main challenges of these databases is to ensure access to unbiased experimental information to interpret the EV content and thus to fit the purpose of biological knowledge discovery. By providing users the EV-METRIC linked to the EV-TRACK entry for reported studies, the 2019 update of Vesiclepedia was a first step towards integrating EV-TRACK knowledge in EV-related databases [21]. We present here the importance, development and integration of the EV-TRACK summary add-on to further integrate experimental information relevant to the interpretation of knowledge in databases and thus facilitate true EV cargo and function discovery using publicly available data.

Development of the EV-TRACK summary add-on

Currently, experimental information on EV separation and characterization provided by EV-related databases is limited and heterogeneously reported between platforms. To address this issue, we developed an EV-TRACK summary add-on (Figure 1), which can easily be integrated in EV-related databases using the following hyperlink http://www.evtrack.org/study_summary.php?PMID= completed with the PubMed ID of the specific study. The summary add-on provides an instant overview of the nine experimental parameters that form the EV-METRIC, a measure for transparent reporting of separation and characterization methods (13; evtrack.org/about.php). Doughnut charts indicate the proportion of reporting adherence to

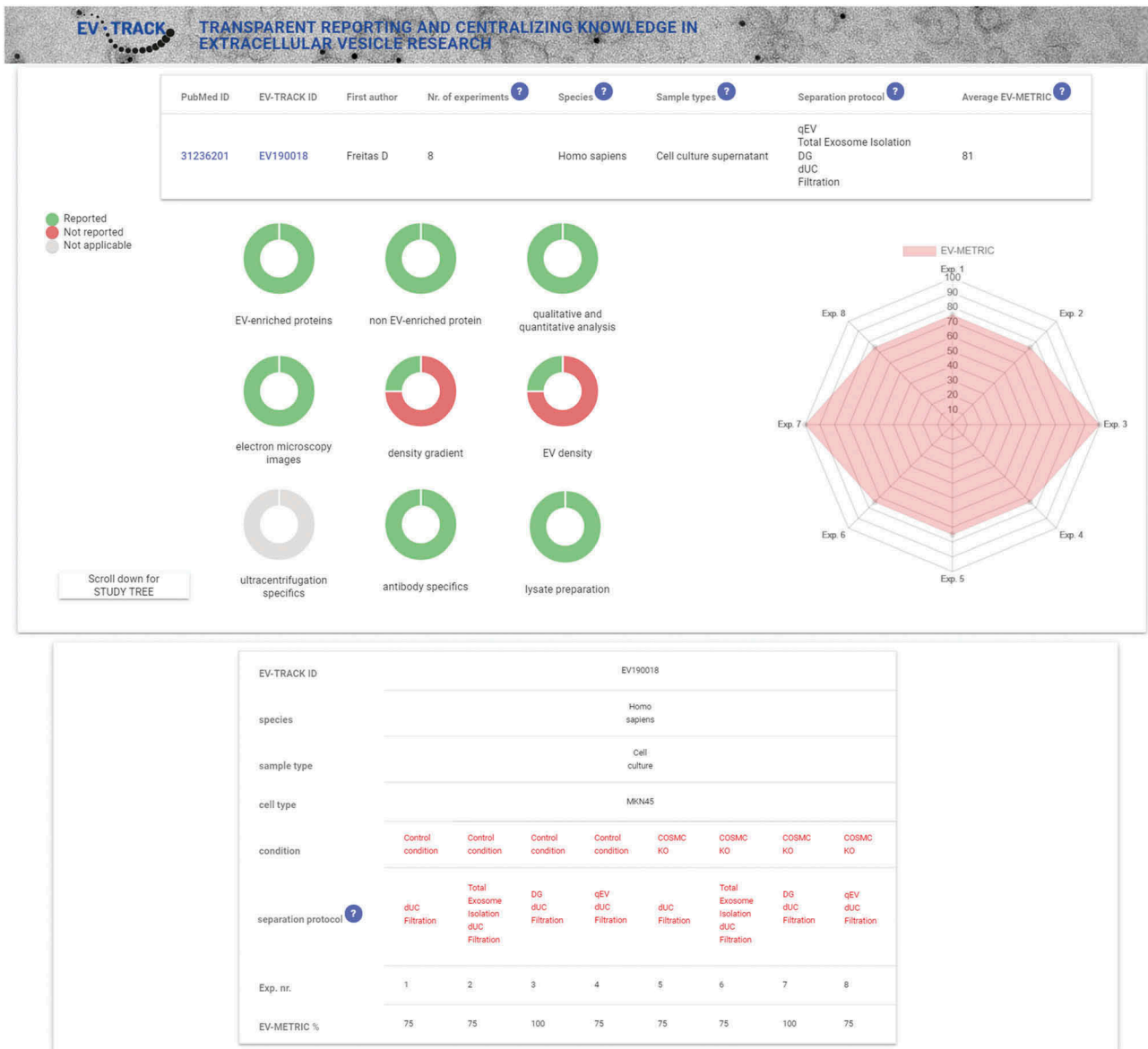


Figure 1. Presentation of the EV-TRACK study summary add-on.

The main page of the summary add-on (upper panel) contains general study information including the number of EV experiments performed, the origin of EVs (species and sample type) and a summary of the separation protocol reported in the study. The average EV-METRIC value given in the upper table is detailed in doughnut charts with visualization of the nine components of the EV-METRIC for the individual study experiments. Depending on the separation and characterization methods implemented, ultracentrifugation specifics, antibody specifics and lysate preparation are applicable or not. For studies with more than two experiments, the EV-METRIC of individual experiments is represented on a radar chart. The study tree (lower panel) shows a summary of the separation protocol of study experiments and their individual EV-METRIC. Extended information on the separation protocols and characterization methods employed for each experiment, can be accessed by clicking on the EV-TRACK ID. For more information on the composition, calculation and use of the EV-METRIC we refer to [13] and evtrack.org/about.php. Abbreviations: DG: density gradient; dUC: differential ultracentrifugation.

each of the nine experimental parameters (Figure 1). Where applicable, the study tree provides a schematic overview of different EV-related experiments and indicates the EV-METRIC for each individual study experiment. Additional experimental information can be viewed by clicking the EV-TRACK ID hyperlink, which redirects to the full entry in the EV-TRACK knowledgebase. For

studies that have not yet been recorded in the EV-TRACK knowledgebase, users following this hyperlink will be invited to connect to the My EV-TRACK page and submit the publication for annotation. Once curation has been completed by EV-TRACK administrators, the EV-TRACK study summary will be automatically generated and become available on all EV-related databases

providing the hyperlink to the EV-TRACK summary add-on. This summary add-on will assist data interpretation and as such enable end-users of EV-related databases to reliably search data for biological knowledge discovery.

EV-TRACK analysis of ExoRbase and EVmiRNA-related publications

To illustrate the added value of the EV-TRACK summary add-on in EV-related databases, we performed an in-depth analysis of publications that were key to the development of the recently released ExoRbase and EVmiRNA databases [17,18].

ExoRbase collects and allows to query publicly available RNA seq data on EV-associated circRNA, lncRNA and mRNA identified in blood from healthy donors or patients. At the time of release, ExoRbase contained RNA seq information from six experiments (87 samples) that were retrieved from the NCBI Gene Expression Omnibus. One published experiment (GSE93078; 2 samples) separated EVs from serum by differential ultracentrifugation prior to RNA seq analysis. Five experiments (GSE100206, GSE99985, GSE100063, GSE100207, GSE100232) separated EV-associated RNA from plasma using the exoRNeasy Serum/Plasma Maxi kit. For these experiments, no characterization of studied EVs was available. To complete annotation of identified RNA datasets, Li et al. performed a PubMed search, which resulted in an additional selection of 77 EV-associated RNA from 18 published studies. To interpret the transparency and reproducibility of these studies, we submitted them to the EV-TRACK knowledgebase and retrieved the EV-METRIC (Figure 2(a)). One study was not included since the manuscript was not available in English.

Calculation of the average EV-METRIC per study revealed that all studies have an EV-METRIC lower than 15%. Almost half of the studies (8/17 or 47%) did not complete the reporting of a single EV-METRIC parameter, corresponding to a metric of 0%. EVs were separated from serum, plasma and cell culture supernatant in respectively 65% (11/17), 23% (4/17) and 12% (2/17) of studies (Figure 2(b)). Since serum contains activated platelet EVs released after coagulation, EV-associated RNA identified from serum samples may warrant validation to allow unbiased consideration [15,22]. Commercially available kits (53%, 9/17) and differential ultracentrifugation (35%, 6/17) were most commonly implemented to separate EVs from biofluids (Figure 2(c)). Previous studies have demonstrated that the choice of separation method impacts the identification of EV-associated RNA since each separation method retrieves EVs with a different specificity [14]. Some of the implemented commercial kits are known to separate EVs with high efficiency but low specificity,

whereas differential ultracentrifugation separates EVs with low efficiency and low specificity [3,14]. Especially in blood, in addition to protein-associated RNA, notable RNA contamination sources include lipoprotein particles such as HDL and LDL. These are known to transport miRNA [23] and are prone to co-separation with EVs due to overlapping size or density. Thus, association of specific RNA to EVs requires further validation. Electron microscopy (EM), western blot (WB) and nanoparticle tracking analysis (NTA) were the most implemented characterization methods, being performed in at least one experiment in respectively 65% (11/17), 47% (8/17) and 35% (6/17) of studies. Almost 30% (5/17) of studies did not characterize EVs used for downstream RNA analysis (Figure 2(d)). Combined RNase and protease treatment prior to RNA isolation to demonstrate the protection of identified RNA by a lipid membrane [24] was not performed in at least one experiment in the analysed studies, while 18% (3/17) of studies performed a RNase only treatment in at least one experiment (Figure 2(f)).

The EVmiRNA database allows to query EV-associated miRNA expression profiles obtained from multiple human biofluids and cell culture supernatant from normal or disease-related samples (Figure 2(b)) [17]. At the time of release, the database included EV miRNA datasets publicly available on NCBI Sequence Read Archive from 24 published studies. EV-TRACK analysis of these studies revealed that 54% (13/24) of studies did not complete the reporting of a single EV-METRIC parameter, while 20% of studies had an average EV-METRIC between 35 and 60% (Figure 2(a)). Methods used to separate EVs from biofluids were diverse, with 42% (10/24) of studies using the commercially available ExoQuick kit and 21% (5/24) using density-based separation. Pelleting of EVs by differential ultracentrifugation was the most represented strategy to separate EVs (46% or 11/24 studies) (Figure 2(c)). EM, WB and NTA were the main methods used to characterize EVs and were performed for at least one experiment in respectively 58% (14/24), 46% (11/24) and 25% (6/24) of studies. Strikingly, in almost 40% of studies (9/24) no characterization was performed of the EVs prior to RNA isolation (Figure 2(d)). Less than 10% of studies (2/24) controlled for RNA association to EVs by using a combined protease/RNase treatment in at least one experiment (Figure 2(f)).

Altogether, only 6% of study experiments (7/117) available in ExoRbase and EVmiRNA upon their release obtained an EV-METRIC above 50%. A significant proportion (58% or 68/117) failed to adhere to a single component of the EV-METRIC and no characterization of EV was performed in 48% (56/117) of study experiments (Figure 2(e)). This result is consistent with the previous EV-TRACK analysis of EV-related publications,

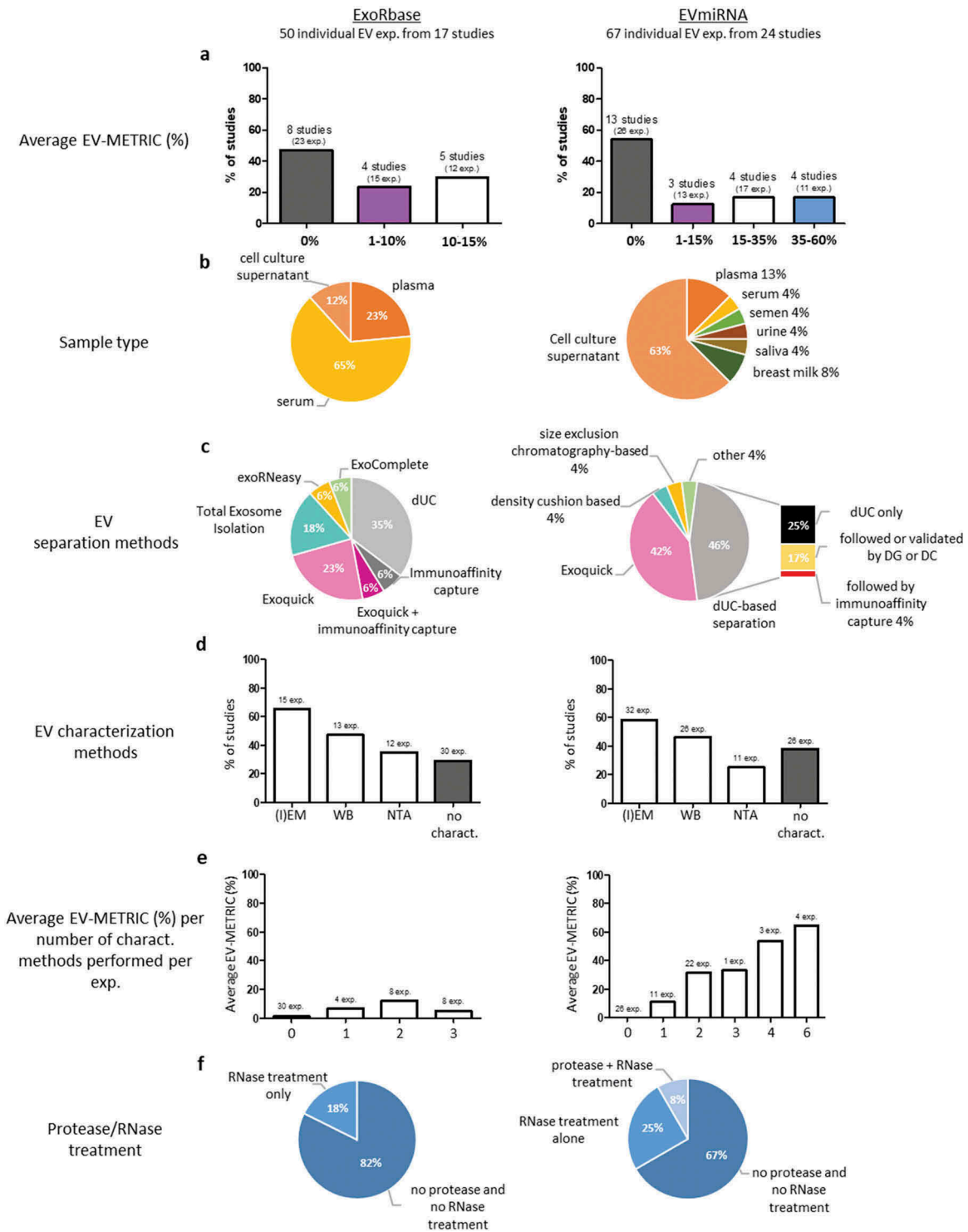


Figure 2. EV-TRACK analysis of ExoRbase and EVmiRNA publications.

Studies considered in ExoRbase as experimental validation studies for blood-EV RNA (50 individual EV experiments, 17 studies) and studies contributing to EVmiRNA datasets (67 individual EV experiment, 24 studies) were submitted to EV-TRACK and analysed. For each study, the average EV-METRIC was calculated based on the reporting of the separation and characterization components represented in Figure 1 [13]. The number of experiments in studies achieving a certain EV-METRIC is indicated in brackets (a). The percentage of studies separating EVs from a given sample type (b), using a selected EV separation method (c), and implementing a RNase and/or protease treatment prior to RNA isolation in at least one experiment (f) are represented in pie charts. The most frequently performed EV characterization methods are represented as the percentage of studies using a given method to characterize EVs in at least one experiment (d). The total number of experiments characterized with a given method are indicated on top of the bars. The percentage of studies and the total number of experiments performing no EV characterization are also included (d). The average EV-METRIC achieved by experiments in function of the number of methods employed to characterize EVs is represented in (e). The number of experiments concerned is indicated. Characterization methods considered for this analysis include western blot (WB), nanoparticle tracking analysis (NTA), dynamic light scattering (DLS), enzyme-linked immunosorbent assay (ELISA), electron microscopy (EM), immunoelectron microscopy (IEM), density gradient (DG), density cushion (DC), proteomics and flow cytometry after specific tetraspanin-mediated association of EVs with beads. Abbreviations: dUC: differential ultracentrifugation.

where the EV-METRIC was above 50% for less than 6% of study experiments and 0% for approximately 30% of study experiments [13]. EV-TRACK analysis of studies included in ExoRbase and EVmiRNA reveals that the EV-associated nature of identified RNA species should be interpreted with care and shows that integration of the EV-TRACK summary add-on can facilitate this interpretation.

Concluding remarks and future perspectives

EVpedia (<http://evpedia.info/> [16]), EVmiRNA (<http://bioinfo.life.hust.edu.cn/EVmiRNA> [17]), ExoRbase (www.exoRbase.org/ [18]), Exocarta (www.exocarta.org/ [19]) and Vesiclepedia (www.microvesicles.org/ [20]), are EV-related databases that aim to centralize knowledge and enable queries on EV cargo including proteins, nucleic acids and lipids for biomarker and functional discovery. These databases combine EV cargo information that 1) has been collected using different separation methods to extract EVs from biofluids, and 2) has not been well characterized and validated, as demonstrated by the absence of EV characterization in 1/3 of studies analysed (34%), and the lack of a combined protease/RNase treatment to validate RNA encapsulation in EVs (5% or 2/41 studies). Efficient separation of EVs from biofluids and identification of their cargo are challenging due to the high complexity of biofluids and the relatively low numbers of EVs, respectively [11]. As such, the choice of separation method influences the identification of EV cargo [14]. In addition, a set of quality control experiments in compliance with MISEV2018 guidelines needs to be completed prior to associating cargo with EVs [15]. The EV-TRACK knowledgebase was developed with the aim to coach researchers in transparent reporting by implementing the EV-METRIC, a minimal set of experimental parameters that improve the interpretation and reproducibility of EV experiments. By developing the EV-TRACK summary add-on, we aim to extend this coaching to other EV-related databases [23]. The summary add-on will assist researchers in the interpretation of the data, and will steer reliable data selection to increase the discovery of biological knowledge.

Data availability statement

All data that were collected during the course of this study, and that support its findings are available online <http://evtrack.org>.

Funding


This work was supported by European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No [722148], Concerted Research Actions from Ghent University and the Fund for Scientific Research-Flanders.

References



- [1] Van Niel G, D'Angelo G, Raposo G. Shedding light on the cell biology of extracellular vesicles. *Nat Rev Mol Cell Biol.* 2018;19:213–228.
- [2] Xu R, Rai A, Chen M, et al. Extracellular vesicles in cancer — implications for future improvements in cancer care. *Nat Rev Clin Oncol.* 2018;15:617.
- [3] Geeurickx E, Tulkens J, Dhondt B, et al. The generation and use of recombinant extracellular vesicles as biological reference material. *Nat Commun.* 2019;10:1–12.
- [4] De Wever O, Hendrix A. A supporting ecosystem to mature extracellular vesicles into clinical application. *Embo J.* 2019;38:e101412.
- [5] Yang KS, Im H, Hong S, et al. Multiparametric plasma EV profiling facilitates diagnosis of pancreatic malignancy. *Sci Transl Med.* 2017;9:eaal3226.
- [6] Melo SA, Luecke LB, Kahlert C, et al. Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. *Nature.* 2015;523:177–182.
- [7] Hu J, Sheng Y, Kwak KJ, et al. A signal-amplifiable biochip quantifies extracellular vesicle-associated RNAs for early cancer detection. *Nat Commun.* 2017;8:1683.
- [8] Jeppesen DK, Fenix AM, Franklin JL, et al. Reassessment of Exosome Composition. *Cell.* 2019;177:428–445.e18.
- [9] Simonsen Jens B. What are we looking at? Extracellular vesicles, lipoproteins, or both? *Circ Res.* 2017;121:920–922.
- [10] Arroyo JD, Chevillet JR, Kroh EM, et al. Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *PNAS.* 2011;108:5003–5008.
- [11] Tulkens J, Vergauwen G, Deun JV, et al. Increased levels of systemic LPS-positive bacterial extracellular vesicles in patients with intestinal barrier dysfunction. *Gut.* 2018. DOI:10.1136/gutjnl-2018-317726.
- [12] Coumans Frank AW, Brisson Alain R, Buzas Edit I, et al. Methodological guidelines to study extracellular vesicles. *Circ Res.* 2017;120:1632–1648.
- [13] EV-TRACK Consortium, Van Deun J, Mestdagh P, et al. EV-TRACK: transparent reporting and centralizing knowledge in extracellular vesicle research. *Nat Methods.* 2017;14:228–232.
- [14] Van Deun J, Mestdagh P, Sormunen R, et al. The impact of disparate isolation methods for extracellular vesicles on downstream RNA profiling. *J Extracell Vesicles.* 2014;3:24858.
- [15] Théry C, Witwer KW, Aikawa E, et al. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the international society for extracellular vesicles and update of the

- MISEV2014 guidelines. *J Extracell Vesicles*. 2018;7:1535750.
- [16] Kim D-K, Kang B, Kim OY, et al. EVpedia: an integrated database of high-throughput data for systemic analyses of extracellular vesicles. *J Extracell Vesicles*. 2013;2:20384.
- [17] Liu T, Zhang Q, Zhang J, et al. EVmiRNA: a database of miRNA profiling in extracellular vesicles. *Nucleic Acids Res*. 2018. DOI:10.1093/nar/gky985.
- [18] Li S, Li Y, Chen B, et al. exoRBase: a database of circRNA, lncRNA and mRNA in human blood exosomes. *Nucleic Acids Res*. 2018;46:D106–D112.
- [19] Simpson RJ, Kalra H, Mathivanan S. ExoCarta as a resource for exosomal research. *J Extracell Vesicles*. 2012;1:18374.
- [20] Kalra H, Simpson RJ, Ji H, et al. Vesiclepedia: a compendium for extracellular vesicles with continuous community annotation. *PLoS Biol*. 2012;10:e1001450.
- [21] Pathan M, Fonseka P, Chitti SV, et al. Vesiclepedia 2019: a compendium of RNA, proteins, lipids and metabolites in extracellular vesicles. *Nucleic Acids Res*. 2018. DOI:10.1093/nar/gky1029.
- [22] Antwi-Baffour S, Adjei J, Aryeh C, et al. Understanding the biosynthesis of platelets-derived extracellular vesicles. *Immun Inflamm Dis*. 2015;3:133–140.
- [23] Vickers KC, Palmisano BT, Shoucri BM, et al. MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. *Nat Cell Biol*. 2011;13:423–433.
- [24] Mateescu B, Kowal EJK, van Balkom BWM, et al. Obstacles and opportunities in the functional analysis of extracellular vesicle RNA – an ISEV position paper. *J Extracell Vesicles*. 2017;6:1286095.

Quentin Roux
*Laboratory of Experimental Cancer Research,
Department of Human Structure and Repair, Ghent
University, Ghent, Belgium
Cancer Research Institute Ghent, Ghent, Belgium*

Jan Van Deun
*Center for Systems Biology, Massachusetts General
Hospital, Boston, MA, USA
Department of Radiology, Massachusetts General
Hospital, Boston, MA, USA*
 <http://orcid.org/0000-0003-1707-6266>

Sandor Dedeyne
*Laboratory of Experimental Cancer Research,
Department of Human Structure and Repair, Ghent
University, Ghent, Belgium
Cancer Research Institute Ghent, Ghent, Belgium*

An Hendrix
 an.hendrix@ugent.be
*Laboratory of Experimental Cancer Research,
Department of Human Structure and Repair, Ghent
University, Ghent, Belgium
Cancer Research Institute Ghent, Ghent, Belgium*
 <http://orcid.org/0000-0002-1883-8309>