Databases and ontologies

EVpedia: a community web portal for extracellular vesicles research

Dae-Kyum Kim¹,†,‡, Jaewook Lee¹,†,‡, Sae Rom Kim¹,‡,
Dong-Sic Choi¹,†, Yae Jin Yoon²,†, Ji Hyun Kim¹,‡, Gyeongyun Go¹,†,
Dinh Nhung¹,†, Kahye Hong¹,†, Su Chul Jang¹,‡, Si-Hyun Kim¹,‡,
Kyong-Su Park¹,†, Oh Youn Kim¹,‡, Hyun Taek Park¹,†, Ji Hye Seo¹,‡,
Elena Aikawa³, Monika Baj-Krzyworzeka⁴, Bas W. M. van Balkom⁵,
Mattias Belting⁶, Lionel Blanc⁷, Vincent Bond⁸, Antonella Bongiovanni⁹,
Francesc E. Borràs¹⁰, Luc Buée¹¹, Edit I. Buza’s¹², Lesley Cheng¹³,
Aled Clayton¹⁴, Emanuele Cocucci¹⁵, Charles S. Dela Cruz¹⁶,
Dominic M. Desiderio¹⁷, Dolores Di Vizio¹⁸, Karin Ekström¹⁹,²⁰,
Juan M. Falcon-Perez²¹, Chris Gardiner²², Bernd Giebel²³,
David W. Greening²⁴, Julia Christina Gross²⁵, Dwijendra Gupta²⁶,
An Hendrix²⁷, Andrew F. Hill¹³, Michelle M. Hill²⁸, Esther Nolte-‘t Hoen²⁹,
Do Won Hwang³⁰, Jameel Inal³¹, Medicarla V. Jagannadham³²,
Muthuvel Jayachandran³³, Young-Koo Jee³⁴, Malene Jørgensen³⁵,
Kwang Pyo Kim³⁶, Yoon-Keun Kim³⁷, Thomas Kislinger³⁸, Cecilia
Lässer³⁹, Dong Soo Lee³⁰, Hakmo Lee⁴⁰, Johannes van Leeuwen⁴¹,
Thomas Lener⁴²,⁴³, Ming-Lin Liu⁴⁴,⁴⁵, Jan Lötvall³⁹, Antonio Marcilla⁴⁶,
Suresh Mathivanan²⁴, Andreas Möller⁴⁷, Jess Morhayim⁴¹, François
Mullier⁴⁸,⁴⁹, Irina Nazarenko⁵⁰, Rienk Nieuwland⁵¹, Diana N. Nunes⁵²,
Ken Pang⁵³,⁵⁴, Jaesung Park⁵⁵, Tushar Patel⁵⁶, Gabriella Pocsfalvi⁵⁷,
Hernando del Portillo⁵⁸, Ulrich Putz⁵⁹, Marcel I. Ramirez⁶⁰,
Marcio L. Rodrigues⁶¹,⁶², Tae-Young Roh¹,², Felix Royo²¹,
Susmita Sahoo⁶³, Raymond Schifffers⁶⁴, Shivani Sharma⁶⁵,
Pia Siljander⁶⁶, Richard J. Simpson²⁴, Carolina Soekmadji⁶⁷,
Philip Stahl⁶⁸, Allan Stensballe⁶⁹, Ewa Stepień⁷⁰, Hidetoshi Tahara⁷¹,
Arne Trummer⁷², Hadi Valadi⁷³, Laura J. Vella⁷⁴, Sun Nyunt Wai⁷⁵,
Kenneth Witwer⁷⁶, Maria Yáñez-Mó⁷⁷, Hyewon Youn³⁰,
Reinhard Zeidler⁷⁸ and Yong Song Gho¹,†,*¹

¹Department of Life Sciences, Pohang University of Science and Technology, Pohang, Republic of Korea, ²Division of Integrative Biosciences and Biotechnology, Pohang University of Science and Technology, Pohang, Republic of Korea, ³Cardiovascular Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA, ⁴Department of Clinical Immunology, Polish-American Institute of Paediatrics, Jagiellonian University Medical College, Cracow, Poland, ⁵Department of Nephrology and Hypertension, University Medical Center Utrecht, Utrecht, The Netherlands, ⁶Section of Oncology, Department of Clinical Sciences, Lund University, Lund, Sweden, ⁷The Feinstein Institute for Medical Research, Manhasset, NY, USA, ⁸Department of Microbiology, Biochemistry,
EVpedia: a portal for extracellular vesicles research

Chemistry and Hematology, University Medical Center Utrecht, Utrecht, The Netherlands, 65California Nanosystems Institute, University of California Los Angeles, Los Angeles, CA, USA, 66Division of Biochemistry and Biotechnology, Department of Biosciences, University of Helsinki, Helsinki, Finland, 67Australian Prostate Cancer Research Centre-Queensland, Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, QLD, Australia, 68Department of Cell Biology and Physiology, Washington University School of Medicine, Saint Louis, MO, USA, 69Department of Health Science and Technology, Aalborg University, Aalborg, Denmark, 70Department of Medical Physics, Jagiellonian University, Cracow, Poland, 71Department of Cellular and Molecular Biology, Graduate School of Biomedical Science, Hiroshima University, Hiroshima, Japan, 72Department of Hematology, Hemostasis, Oncology and Stem Cell Transplantation, Hannover Medical School, Hannover, Germany, 73Department of Rheumatology and Inflammation Research, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden, 74Ludwig Institute for Cancer Research Melbourne, Austin Hospital, Heidelberg, VIC, Australia, 75Department of Molecular Biology, Umeå University, Umeå, Sweden, 76Department of Molecular and Comparative Pathobiology, The Johns Hopkins University School of Medicine, Baltimore, MD, USA, 77Unidad de Investigación, Hospital Santa Cristina, Instituto de Investigación Sanitaria Princesa, Madrid, Spain, 78Department of Otolaryngology and Research Group Gene Vectors, University of Muenchen, Helmholtz-Zentrum München, Munich, Germany

*To whom correspondence should be addressed.
†The authors wish to be known that, in their opinion, the first two authors should be regarded as joint First Authors.
Associate Editor: Janet Kelso
*Only these authors are not listed alphabetically by their last name.

Received on July 9, 2014; revised on October 25, 2014; accepted on November 5, 2014

Abstract

Motivation: Extracellular vesicles (EVs) are spherical bilayered proteolipids, harboring various bio-active molecules. Due to the complexity of the vesicular nomenclatures and components, online searches for EV-related publications and vesicular components are currently challenging.

Results: We present an improved version of EVpedia, a public database for EVs research. This community web portal contains a database of publications and vesicular components, identification of orthologous vesicular components, bioinformatic tools and a personalized function. EVpedia includes 6879 publications, 172 080 vesicular components from 263 high-throughput datasets, and has been accessed more than 65 000 times from more than 750 cities. In addition, about 350 members from 73 international research groups have participated in developing EVpedia. This free web-based database might serve as a useful resource to stimulate the emerging field of EV research.

Availability and implementation: The web site was implemented in PHP, Java, MySQL and Apache, and is freely available at http://evpedia.info.

Contact: ysgho@postech.ac.kr

1 Introduction

Almost all living organisms on earth shed extracellular vesicles (EVs) into their microenvironment. EVs are spherical bilayered proteolipids with an average diameter of 20–1000 nm (Ellen et al., 2009; Lee et al., 2008; Théry et al., 2009). Cells release EVs either constitutively or in a regulated manner and these vesicles harbor a specific subset of proteins, mRNAs, miRNAs, lipids and metabolites reflecting their originating cell types and conditions (Bellingham et al., 2012; Choi et al., 2014; de Jong et al., 2012; Duperthuy et al., 2013; Mayr et al., 2009; Raimondo et al., 2011; Simpson et al., 2008; Subra et al., 2007; Wai et al., 2003). EVs are also found in various biological fluids, such as amniotic fluid, ascites, breast milk, plasma, saliva, semen, serum and urine (Asea et al., 2008; Caby et al., 2005; Cheng et al., 2014a, b; Dai et al., 2008; George et al., 1982; Lasser et al., 2011; Poljakov et al., 2009; Raj et al., 2012; Witwer et al., 2013). Recent advances in this fast growing field (Fig. 1) have facilitated several insights: (1) EVs play multifaceted functions in intercellular communication (Cocucci et al., 2009; Lee et al., 2008; Simons and Raposo, 2009; Théry et al., 2009); (2) EV-mediated intercellular communication is an evolutionarily conserved phenomenon (Deaner and Cookson, 2012; Lee et al., 2008, 2009); (3) EVs are rich sources of biomarkers for non-invasive diagnosis and prognosis of various human diseases (Chaput et al., 2003; Choi et al., 2013a; D’Souza-Schorey and Clancy, 2012; Muller et al., 2013; Sarlon-Bartoli et al., 2013; Shedden et al., 2003; Simpson et al., 2009) and (4) Diverse therapeutic approaches have been pursued to utilize EVs and their mimetics for vaccine, chemotherapeutic drug and siRNA delivery (Alvarez-Erviti et al., 2011; Chaput et al., 2005; Jang et al., 2013; Kordelas et al., 2014; Lai et al., 2010; Lee et al., 2012; Simpson et al., 2009; Sun et al., 2010).

Publications on EVs have grown rapidly during the last several years, indicating that the field of EVs is expanding intensively (Fig. 1). The identification of vesicle-specific cargos could help us to unravel the molecular mechanisms underlying the cargo sorting and biogenesis of EVs. In addition, this will lead to better comprehension of the pathophysiological functions of EVs, and discovery of EV-based potential biomarkers of human diseases. Therefore, many
investigators have focused on categorizing these complex vesicular components by various high-throughput technologies, such as mass spectrometry-based proteomics and lipidomics as well as micro-array- and next-generation sequencing-based transcriptomics (Barry et al., 1997; Choi et al., 2013b; Koh et al., 2010; Utleg et al., 2003; Valadi et al., 2007). Together with conventional biological approaches, these multiomics-based analyses of EVs derived from various cell types and body fluids have identified several thousands of different vesicular components.

EV secretion and EV-mediated intercellular communication are evolutionarily conserved (Biller et al., 2014; Deatherage and Cookson, 2012; Lee et al., 2008, 2009). Researchers in this field have coined dozens of different names for EVs, especially for more complex eukaryotic cell-derived EVs as listed in Box ('Extracellular Vesicles: Diverse Nomenclature') (Choi et al., 2014; Gould and Raposo, 2013; Kim et al., 2013). Nevertheless, there is progress toward a single nomenclature, since the different names in use make it difficult to follow the progress in the field. In addition, most vesicular components identified by multiomics-based high throughput analyses are presented in the supplementary information of published articles. Taken together, online searches for EV-related publications and vesicular components are currently challenging, especially for start-up researchers in this field. Therefore, a comprehensive public repository of EV-related publications and vesicular components will help the community of EV research to understand various aspects of these complex extracellular organelles.

2 Databases that store EV data
The explosion of EV data has justified the need for databases that catalog proteins, nucleic acids and lipids associated with EVs. Currently, three databases exist for EV research including ExoCarta (Simpson et al., 2012), EVpedia (Kim et al., 2013) and Vesiclepedia (Kalra et al., 2012). These existing databases have made large-scale bioinformatics analyses feasible and provide an ideal platform for EV-based biomarker studies. EVpedia provides additional benefits compared with ExoCarta and Vesiclepedia. EVpedia is the only source that contains data on both prokaryotic and eukaryotic EVs. In addition, EVpedia allows for Gene Ontology enrichment analysis, network analysis of vesicular proteins and mRNAs, and set analysis of vesicular datasets by ortholog identification. Other databases do not have any such embedded analysis tools.

3 Launch of EVpedia
EVpedia (http://evpedia.info) was first launched in January 2012 (Kim et al., 2013). To construct this public web-based database, we first collected publications on prokaryotic and eukaryotic EVs through a combination of NCBI PubMed searches (http://www.ncbi.nlm.nih.gov/pubmed) for text-mining solutions and manual curation using all nomenclatures assigned to EVs described in Box (see also Kim et al., 2013). Based on these EV-related publications, we constructed the comprehensive and integrated database of proteins, mRNAs, miRNAs, lipids and metabolites for systematic analyses of prokaryotic and eukaryotic EVs.

4 Overview of current EVpedia
Since the launch of EVpedia, we have improved this database by continually collecting additional EV-related publications and datasets, by adding more tools for systematic analyses of EVs, and by supplementing the menu bars for ‘Top 100+ EV markers’, ‘User forum’ as well as ‘My EVpedia’. Through closed and open beta tests, we built an ‘EVpedia Community’ (about 350 worldwide EV researchers) and updated EVpedia most recently in May 2014. The updated EVpedia has five functional modules for systematic analyses of EVs derived from prokaryotic and eukaryotic cells (Fig. 2): (i) a database of publications and principal investigators, (ii) a database of vesicular proteins, mRNAs, miRNAs, lipids and...
metabolites, (iii) identification of orthologous vesicular components, (iv) an array of tools for bioinformatic analyses including sequence search, set analysis, Gene Ontology enrichment analysis and network analysis, as well as (v) ‘My EVpedia’, a personalized function of EVpedia. Using ‘My EVpedia’, users privately store their own datasets, analysis results and publications of interest by creating their own accounts. New functions of this updated EVpedia are indicated in red texts in Figure 2.

We invite the research community to submit EV-related multiomics data and publications to EVpedia.

As of May 2014, a total of 6879 EV-related publications with 3336 principal investigators have been cataloged in EVpedia. In addition, a total of 172 080 vesicular components from 263 high-throughput datasets are listed (Table 1). All of these vesicular components could be searched by their sequences and browsed. Furthermore, in the ‘Top 100+ EV markers’ menu, the current vesicular components are sorted in the descending order of their identification counts, which are the numbers of datasets identifying those vesicular components or their orthologs.

5 Community participation and annotation in EVpedia

After the initial launch in January 2012, EVpedia has been globally accessed more than 65 000 times from more than 750 cities (Table 1). For community annotation of EVpedia, we built ‘EVpedia Community’. About 350 members from 73 international EV research groups have joined this community, in which they can exchange EV-related information and submit their multiomics data via the ‘User forum’ and ‘Upload’ menu bars in EVpedia, respectively. In addition, non-members can easily join the ‘EVpedia Community’ by adding their information via clicking the ‘Sign In’ menu. Moreover, EVpedia has been cross-linked with the website of the ‘International Society of Extracellular Vesicles’ (http://www.isiev.org).

6 Concluding remarks

EVpedia is a comprehensive database of EVs derived from prokaryotes and eukaryotes. Currently, a total of 6879 EV-related publications and 172 080 vesicular components (proteins, mRNAs, miRNAs, lipids and metabolites) are deposited in this public repository. For the systematic analysis of EVs, EVpedia also provides integrated systems biology research tools such as ‘Experiment’, ‘Browse’, ‘Analysis’, ‘Top 100+ EV markers’ and ‘My EVpedia’ menu bars. In the future, additional multiomics datasets and publications will be deposited, and we expect more researchers to join the ‘EVpedia Community’ and to share their research data. The community database is scheduled to be updated every 3 months. EVpedia, a community web portal for EV research, should serve as a useful resource to stimulate the emerging field of EV biology research and to help us to elucidate the fundamental roles of these complex extracellular organelles.
Table 1. EVpedia statistics

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Eukaryotes</th>
<th>Prokaryotes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Publications</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Articles</td>
<td>6879</td>
<td>6021</td>
<td>858</td>
</tr>
<tr>
<td>Principal investigators</td>
<td>3336</td>
<td>2886</td>
<td>483</td>
</tr>
<tr>
<td><strong>Proteomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies</td>
<td>117</td>
<td>97</td>
<td>20</td>
</tr>
<tr>
<td>Datasets</td>
<td>176</td>
<td>148</td>
<td>28</td>
</tr>
<tr>
<td>Proteins</td>
<td>78971</td>
<td>74966</td>
<td>4275</td>
</tr>
<tr>
<td><strong>Transcriptsomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mRNA Studies</td>
<td>17</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>Datasets</td>
<td>28</td>
<td>28</td>
<td>0</td>
</tr>
<tr>
<td>mRNAs</td>
<td>74430</td>
<td>74430</td>
<td>0</td>
</tr>
<tr>
<td>miRNA Studies</td>
<td>11</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Datasets</td>
<td>29</td>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>miRNAs</td>
<td>18119</td>
<td>18119</td>
<td>0</td>
</tr>
<tr>
<td><strong>Lipidomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies</td>
<td>22</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>Datasets</td>
<td>29</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td><strong>Metabolomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Datasets</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Participating</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laboratories (countries)</td>
<td>73 (20)</td>
<td>66617 (73)</td>
<td></td>
</tr>
<tr>
<td>Accesses (countries)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Funding**

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) [No. 2014-023004], the Ministry of Health and Welfare grant funded by the Korea government [No. A120273] and a grant from KRIBB Research Initiative Program. D.-K. K is supported by a National Junior Research Fellowship [No. 2014-048579] the Ministry of Health and Welfare grant funded by the Korea government (NRF) grant funded by the Korea government (MSIP) [No. 2014-023004], This work was supported by the National Research Foundation of Korea.

**Conflict of Interest:** none declared.

**References**


