1 Supplementary data

2

3 The Male Fertility Gene Atlas – A web tool for collecting and integrating OMICS data in

- 4 the context of male infertility
- 5
- 6 H. Krenz, J. Gromoll, T. Darde, F. Chalmel, M. Dugas, F. Tüttelmann

8 Extended Methods

9 **Requirements engineering**

The MFGA is designed to support researchers and clinicians in the field of male infertility in finding and evaluating the analysis results of relevant publications. To ensure that the MFGA offers a relevant scope and focusses on the central needs and interests of the targeted user group, an extensive requirements engineering process had to be employed. This process aimed at defining the most relevant parts of the system design; in case of the MFGA, this means answering the following five questions:

- 16 1. What kind of data should be available?
- 17 2. How should the database be maintained and kept up to date?
- 18 3. Who should be authorised to access the data on the website?
- 19 4. What interface should be provided to access the data?
- 20 5. How should the data be visualised?

Since user-centric design is proven to be a critical factor of success for software projects 21 (Maguire and Bevan, 2002), the requirements were acquired in close cooperation with a 22 23 large group of prospective users of the MFGA. These researchers and clinicians are part of the Münster-based clinical research unit (CRU326) (http://www.male-germ-cells.de) and 24 provided input on a broad range of aspects of male infertility, germ cell development as well 25 as sperm morphology, motility, and function. The development of the MFGA is based on the 26 27 software development and requirements engineering paradigm Rapid Prototyping, which 28 focusses on developing early testing versions of the product to enable informed user feedback and iterative improvements (Budde et al., 1992; Gordon and Bieman, 1995). 29

The requirements analysis was based on a set of 39 representative and relevant publications/data sets that the MFGA should be able to host, provided by members of the CRU326. The publications were highly heterogeneous and covered a broad range of methodologies, from GWA studies to targeted genotyping, microarray, bulk or single-cell RNA sequencing as well as methylation. The complete list is available in Suppl. Tab. S1 and
has already been included into the MFGA database.

36

37 Current state of requirements

The requirements regarding the kind of data that should be available in the MFGA can be 38 summarised in two points: First, the MFGA should host data representing the analysis results 39 of a comprehensive set of publications on male infertility. Generally speaking, the MFGA 40 41 needs to be able to deal with publications of arbitrary form and content. To account for the extremely high degree of heterogeneity between publications, a very flexible, modular data 42 model is required. However, to enable the MFGA's central service of offering selective data 43 retrieval operations, it is crucial to identify, standardise and tag recurrent relevant information 44 items and to curate them manually for each publication. Whenever applicable, 45 standardisation should be based on ontologies as provided by OBO Foundry (Smith et al., 46 2007). Second, the MFGA should provide complementary data from external databases to 47 increase the comprehensibility of information, e.g., to explain the background of gene names 48 and provide links to more details. However, the MFGA is not supposed to contain any raw 49 data; instead the system should link to the corresponding repositories if provided by a 50 publication. Also, the MFGA is not planned for processing raw data by itself; the system is 51 only intended to show aggregated result data provided by publications. 52

The database of the MFGA should be fully public and accessible to all researchers and 53 54 clinicians in the field of male infertility. Maintenance and curation of selected publications should be provided by the development team. In order to keep the database up to date in a 55 structured way, a recurring process for preparing new relevant publications will be 56 implemented (Suppl. Fig. S1). It comprises three main steps: selection of relevant 57 publications, manual curation according to the requirements of the MFGA database and data 58 upload. Prospectively, it should be triggered once per month and combine the proposed 59 publications of four different sources into one prioritised list. The four information sources are 60

(1) the list of not yet processed publications from the previous month, (2) newly published 61 manuscripts from the CRU326, (3) other relevant new publications identified by the MFGA 62 team, and, as an additional unbiased source, (4) recent publications queried from PubMed 63 (query: "(male OR men) AND (fertility OR infertility) AND (gene OR genetic) AND 64 ("YYY/MM/01"[Date - Publication] : "YEAR/MM/31"[Date - Publication])"). Publications will 65 be curated and uploaded by priority. Eventually, users are supposed to be enabled to act as 66 registered contributors and support selection, curation and upload of content. Thus, a user 67 management and authentication system is required. 68

Regarding the interface for accessing the information, users should be able to view a 69 complete list of publications and their key features to get a quick overview on the full content 70 of the data base. Additionally, an advanced search function should provide identifying and 71 72 showing subsets of relevant publications based on the occurrence of gene names or IDs in tables and figures as well as meta-information like data type, OMICS, processed tissues/cells 73 74 and conditions/species. Detailed information and results of each individual publication should be displayed by comprehensible, standardised overviews, including images, tables and on-75 76 demand plots. As a special feature, the MFGA should display data from different publications in an aggregated way to allow for meta-analyses and integrated analyses. All pages should 77 be enriched with appropriate additional information from external databases and with cross 78 references, allowing for quick navigation through the atlas. 79

80

81 IT architecture

The overall IT architecture of the MFGA (Suppl. Fig. S2) was specified according to the identified requirements. In general, the MFGA consists of a relational database and a *Java* web application as the user interface for accessing it. Both the database and the web application are securely hosted on a server at the University of Münster, Germany, and can be accessed freely via the URL <u>https://mfga.uni-muenster.de</u>.

The part of the web application that is actually shown in the web browser of the user's 87 machine (Suppl. Fig. S2, left) is based on the three languages HTML, CSS and Java Script. 88 The general contents of the web pages are defined by HTML, whereas CSS is employed to 89 specify their layout. For the MFGA, the CSS library Bootstrap v 4.0.0-alpha.6 (Otto et al., 90 2020a) has been added to support building the web pages in a responsive way, i.e., to 91 enable an automated adjustment of the layout to different browsers and devices (Otto et al., 92 2020b). Lastly, Java Script enables user interaction and regulates the web pages' responses, 93 94 making them dynamic on the user side. Here, the Highcharts v 6.1.4 (Highsoft AS, 2020) library has been embedded to allow for interactive charts on the MFGA web pages. 95 Additionally, the library jQuery v 1.8.2 (JS Foundation, 2020a) is used, which simplifies 96 navigation and manipulation of the HTML content compared to pure Java Script. Of major 97 importance for the MFGA are the AJAX functions of the jQuery library that enable requesting 98 and loading data from the server without reloading the web page (JS Foundation, 2020b). 99 100 These functions reduce the amount of data that simultaneously needs to be available in the 101 memory of the user's machine and increase the loading speed of the web pages.

102 For the interface layer on the server side (Suppl. Fig. S2, right), Java Server Pages and its 103 JSTL libraries are used to dynamically define the contents that are rendered as static HTML at the web browser on the client side (Oracle Corporation, 2019). Thus, different data, texts 104 and images can be loaded into an identical Java Server Pages frame depending on the user 105 request. Based on this and the framework Apache Tiles v 3.0.7 (Apache Software 106 Foundation, 2017b) that is employed to assemble multiple Java Server Pages into a final 107 108 web page at runtime, a high degree of reusability of common page elements (Apache Software Foundation, 2017a) and Java Server Pages code snippets is enabled. 109

The transmission of data between the database layer and the interface layer is handled in the application layer (Suppl. Fig. S2, right). For the MFGA, this is facilitated by the three frameworks *Spring v 4.3.7.RELEASE* (Pivotal Software Inc., 2020), *Hibernate v 5.2.9.Final* (Red Hat Inc., 2020) and *Spring Security 4.2.2.RELEASE* (*Pivotal Software Inc., 2020*). *Spring* provides infrastructural support (Pivotal Software Inc., 2019a) while *Hibernate* bridges

the gap between the paradigms of object-oriented programming in Java and relational 115 116 databases by mapping Java classes and objects to data tables (Red Hat Inc., 2018) and vice 117 versa. Also, the Hibernate framework provides the means to formulate queries in Java and 118 facilitates data retrieval (Red Hat Inc., 2018). User authentication and authorisation to access data, however, is controlled by Spring Security (Pivotal Software Inc., 2019b). It enables the 119 120 differentiation between administrators that have read/write access to all data in the MFGA 121 and users that have restricted read-only access to public data. All dependencies of the web 122 application are managed by the Apache Maven v 3.6.1 (Porter et al., 2020) framework, which builds the application before it is deployed on the server. 123

On the data layer (Suppl. Fig. S2, right), the MFGA server employs the open source *MariaDB* v 10.1.41-*MariaDB* relational database (MariaDB Corporation, 2020) that is regularly backed up. MariaDB is the licence-free alternative to MySQL, which is the most commonly used database for public websites (Dyer, 2015). However, both databases have similar functionality and are fast and stable (Dyer, 2015).

129

130 Data model

131 For representing the highly heterogeneous publications, a very flexible and thus modular 132 data model was required. Generally, database paradigms such as noSQL are specified on 133 accommodating unstructured, heterogeneous data (Parker et al., 2013). However, since the MFGA needs to employ many complex queries with aggregation methods to fulfil search 134 135 requests, a relational database was the better choice (Parker et al., 2013). Thus, information items that are preserved over all publications had to be identified and transformed into 136 relational data tables. The seven classes of information that were modelled are publication, 137 data set, processed tissue type, processed cell type, cohort's condition, image and table 138 139 (Tab. 1, column 1). Each of these classes has multiple relevant information items that need to be filled with text or numbers (Tab. 1, column 2 & 3). It is important to notice that each 140 141 publication must have exactly one object in the class of information for publication. Since one publication can contain information on multiple data sets, multiple data set objects can be
attached to each publication object. Similarly, each data set object can be specified by an
arbitrary number of processed tissue and cell types as well as cohorts with different
conditions (one object per tissue type, cell type and condition) and can have images and
tables attached. These relationships are depicted as an entity relationship diagram (Chen,
1976) in Suppl. Fig. S3.

148 For images and tables of publications that should be stored in the MFGA database, further 149 attention is required. In the case of images, a copy of the file (in addition to its metainformation) needs to be available on the server. Most importantly, any depicted gene names 150 and IDs need to be extracted from the image and attached to it as annotation for later 151 information retrieval operations. In the case of tables, the data needs to be extracted from 152 the file provided by the publication and stored in the database cell by cell. The cell content is 153 154 required to be annotated either as one of the sets of standard columns (gene name, artefact ID, chromosome, start position, end position, distance to related gene, p-value) or as an 155 additional text/numeric column. This is necessary to allow for efficient standard queries while 156 157 preserving the ability to store arbitrary tables, query all cell contents and represent tables in an interactive way on the website, i.e., enabling the user to apply sorting and filtering 158 operations as well as plotting. The cell values of standard columns are all stored in one 159 standard relational data table, while additional text or numeric cell values are stored in an 160 161 attribute value pair data table (Suppl. Fig. S3). For the current implementation, tables only need to fulfil two requirements to be uploaded: The first row needs to be defined as a header, 162 and cells cannot be merged. 163

165 Table legend

Suppl. Table S1. List of publications available in the MFGA database. The data model has been developed based on of 39 heterogeneous publications to represent the variety of techniques and topics to be covered for a gene atlas in the field of male infertility and germ cells. These 39 publications are marked with an asterisk (*). The database of the MFGA is constantly growing and new publications are added on a regular basis. To date (30 April 2020), it contains an additional 14 publications. Thus, in total 54 publications are available.

173 Figure legends

Suppl. Figure S1. The process of identifying, selecting and prioritising novel publications depicted based on Business Process Model and Notation (Chinosi and Trombetta, 2012). The process is triggered monthly. Potential publications are collected and merged into a list. This list is then prioritised by the MFGA team. Subsequently, curation and upload are performed in the determined order.

179

Suppl. Figure S2. IT Architecture of the MFGA. The MFGA consists of two parts: The user 180 machine (left side) on which the web application is run and the webserver (right side) on 181 182 which the application and the database are hosted. The content on the user machine is based on the languages HTML, CSS and Java Script and the libraries Bootstrap v 4.0.0-183 alpha.6 (Otto et al., 2020a), Highchart v 6.1.4 (Highsoft AS, 2020), Ajax and jQuery (JS 184 185 Foundation, 2020a). The backend employs the frameworks Apache Tiles v 3.0.7 (Apache Software Foundation, 2017b), Spring v 4.3.7.RELEASE (Pivotal Software Inc., 2020), 186 Hibernate v 5.2.9. Final (Red Hat Inc., 2020) and Apache Maven v 3.6.1 (Porter et al., 2020) 187 188 as well as the technology Java Server Pages (Oracle Corporation, 2019) and the languages Java and MariaDB v 10.1.41-MariaDB (MariaDB Corporation, 2020). The application 189 provides different interfaces and different access rights for users and admins, which are 190 depicted in two different tracks of interaction. 191

192

Suppl. Figure S3. MFGA's data model to store the content of publications in a 193 relational database. The model is depicted as Entity-Relationship Model according to the 194 Chen notation (Chen, 1976). Each publication must have exactly one data object of the type 195 publication for general metal information like author and title. Additionally, a data object of the 196 type publication can have zero or more objects of the type data set attached. Equivalently, 197 each data set can have zero or more data objects of the types processed tissue type. 198 processed cell type, cohort's condition, image and data table. The data table object stores 199 200 general meta information on a data table like title and description. For the actual data, there has to be one object of the type table row attached per actual table row and per table row 201 object as many attribute value pair objects as there are non-standard columns in the table. 202 203

Suppl. Figure S4a. Example screenshot of the MFGA search for gene *STAG3* (https://mfga.uni-muenster.de/search.html). The side bar on the left contains the search form. Here, *STAG3* is typed into the search field and search option 'all search terms' is checked. The main window presents the search results. The plots show the number of returned data sets and the frequency of *STAG3* in data tables grouped by publication. Below, the returned data sets are presented in a table (shortened) enriched with further meta information, linking to the overview of data set and publication. Retrieved 30 April 2020.

211

Suppl. Figure S4b. Example screenshot of the MFGA search for gene STAG3. It shows part of the lower table (shortened) that results from the search for the gene STAG3 in the MFGA database and contains the specific data tables and images in which the gene was found. Retrieved 30 April 2020.

216

Suppl. Figure S5a. Example screenshot of the MFGA search for condition azoospermia 217 (https://mfga.uni-muenster.de/search.html). The side bar on the left contains the search 218 form. Here, 'HP:0000027: Azoospermia' and HPO terms on lower hierarchy levels are 219 checked as well as search option 'all search terms'. HPO terms that are in a direct 220 hierarchical relation are gueried with the logical operator 'OR'. The plots show the number of 221 222 returned data sets and the most frequent genes in their data tables. Below, the returned data 223 sets are presented in a table (shortened) enriched with further meta information, linking to 224 the overview of data set and publication. Retrieved 30 April 2020. 225

Suppl. Figure S5b. Example screenshot of the MFGA search for condition
 azoospermia (<u>https://mfga.uni-muenster.de/search.html</u>). It shows part of the lower table
 (shortened) that results from the search for condition azoospermia in the MFGA database
 containing frequency and names of genes mentioned in data tables of the returned data sets.
 Retrieved 30 April 2020.

231

236

Suppl. Figure S6. Example screenshot of the MFGA publication overviews
 (<u>https://mfga.uni-muenster.de/publication/details?publicationId=3881147</u>). The
 overview shows the publication van der Bijl *et al.* (2019), presenting title, authors, abstract,
 information on copyright and first data set. Retrieved 30 April 2020.

Suppl. Figure S7. Example screenshot of the MFGA data table overviews
 (<u>https://mfga.uni-muenster.de/publication/results?ssDetailsId=3968811</u>). The overview
 shows supplementary table 3 of the van der Bijl *et al.* (2019), containing citation, title and
 description of the data table as well as information on copyright. Retrieved 30 April 2020.

Suppl. Figure S8a. Example screenshot of the MFGA search for gene expression of
Sertoli cells of human testis tissue using single cell RNA sequencing (https://mfga.unimuenster.de/search.html). The plots show the number of returned data sets and the most
frequent genes in their data tables. Below, the returned data sets are presented in a table
(shortened) enriched with further meta information, linking to the overview of data set and
publication. Retrieved 30 April 2020.

Suppl. Figure S8b. Example screenshot of the MFGA search for gene expression of sertoli cells of human testis tissue using single cell RNA sequencing (<u>https://mfga.uni-</u> <u>muenster.de/search.html</u>). It shows part of the lower table that results from the search in the MFGA database containing frequency and names of genes mentioned in data tables of the returned data sets. Retrieved 30 April 2020.

- 1 Suppl. Table S1. List of publications available in the MFGA database. The data model
- 2 has been developed based on of 39 heterogeneous publications to represent the variety of
- 3 techniques and topics to be covered for a gene atlas in the field of male infertility and germ
- 4 cells. These 39 publications are marked with an asterisk (*). The database of the MFGA is
- 5 constantly growing and new publications are added on a regular basis. To date (30 April
- 6 2020), it contains an additional 14 publications. Thus, in total 54 publications are available.

1	Agarwal A, Panner Selvam MK, Baskaran S. Proteomic Analyses of Human Sperm Cells:
	Understanding the Role of Proteins and Molecular Pathways Affecting Male
	Reproductive Health. International journal of molecular sciences 2020;21.
2	Amaral A, Paiva C, Attardo Parrinello C, Estanyol JM, Ballescà JL, Ramalho-Santos J,
	Oliva R. Identification of proteins involved in human sperm motility using high-
	throughput differential proteomics. Journal of proteome research 2014;13:5670–5684.
3	Araujo TF, Friedrich C, Grangeiro CHP, Martelli LR, Grzesiuk JD, Emich J, Wyrwoll MJ,
	Kliesch S, Simões AL, Tüttelmann F. Sequence analysis of 37 candidate genes for
	male infertility: challenges in variant assessment and validating genes. Andrology
	2020;8:434–441.
4	Aston KI, Carrell DT. Genome-wide study of single-nucleotide polymorphisms associated
	with azoospermia and severe oligozoospermia. Journal of andrology 2009;30:711–
_	725. [°]
5	Aston KI, Krausz C, Larace I, Ruiz-Castane E, Carrell DT. Evaluation of 172 candidate
	of map of European descent. Human reproduction (Ovford, England) 2010:25:1282
	1307 *
6	Ben Khelifa M. Coutton C. Zouari R. Karaouzène T. Rendu J. Bidart M. Yassine S. Pierre
Ŭ	V. Delaroche J. Hennebicg S. et al. Mutations in DNAH1, which Encodes an Inner
	Arm Heavy Chain Dynein, Lead to Male Infertility from Multiple Morphological
	Abnormalities of the Sperm Flagella. The American Journal of Human Genetics
	2014;94:95–104.*
7	Busch AS, Tüttelmann F, Cremers J-F, Schubert M, Nordhoff V, Schüring AN, Zitzmann
	M, Gromoll J, Kliesch S. FSHB -211 GT Polymorphism as Predictor for TESE
	Success in Patients With Unexplained Azoospermia. The Journal of clinical
	endocrinology and metabolism 2019;104:2315–2324.
8	Castillo J, Jodar M, Oliva R. The contribution of human sperm proteins to the
	development and epigenome of the preimplantation embryo. Human reproduction
0	update 2018;24:535–555.
9	Chaimel F, Lardenois A, Evrard B, Mathieu R, Feig C, Demougin P, Gattiker A, Schulze
	w, Jegou B, Kirchnoll C, et al. Global numan tissue profiling and protein network
	target genes for male infertility. Human reproduction (Oxford, England)
	2012·27·3233–3248 *
10	Chen D. Sun N. Hou L. Kim R. Faith J. Aslanvan M. Tao Y. Zheng Y. Fu J. Liu W. et al.
	Human Primordial Germ Cells Are Specified from Lineage-Primed Progenitors. Cell
	Reports 2019;29:4568-4582.e5.
11	Chen Y, Zheng Y, Gao Y, Lin Z, Yang S, Wang T, Wang Q, Xie N, Hua R, Liu M, et al.
	Single-cell RNA-seq uncovers dynamic processes and critical regulators in mouse
	spermatogenesis. Cell Res 2018;28:879–896.*
12	Coutton C, Martinez G, Kherraf ZE, Amiri-Yekta A, Boguenet M, Saut A, He X, Zhang F,
	Cristou-Kent M, Escoffier J, et al. Severe Astheno-Teratozoospermia Due to Sperm
	Flagellum Malformations in Humans and Mice. American Journal of Human Genetics
4.0	
13	Dalgaard IVID, Weinhold N, Edsgard D, Silver JD, Pers TH, Nielsen JE, Jørgensen N,
	Juur A, Gerus TA, Giwercman A, et al. A genome-wide association study of men with symptoms of testicular dysgonasis syndrome and its network biology interpretation
	symptoms of testicular dysgenesis syndrome and its network biology interpretation.
	Juina of Medical Genetics 2012,43.30-03.

14	Gkountela S, Zhang KX, Shafiq TA, Liao W-W, Hargan-Calvopiña J, Chen P-Y, Clark AT. DNA Demethylation Dynamics in the Human Prenatal Germline. <i>Cell</i> 2015; 161 :1425– 1436.*
15	Gomes Fernandes M, Bialecka M, Salvatori DCF, Chuva de Sousa Lopes SM. Characterization of migratory primordial germ cells in the aorta-gonad-mesonephros of a 4.5-week-old human embryo: a toolbox to evaluate in vitro early gametogenesis. Molecular human reproduction 2018;24:233–243.*
16	Green CD, Ma Q, Manske GL, Shami AN, Zheng X, Marini S, Moritz L, Sultan C, Gurczynski SJ, Moore BB, et al. A Comprehensive Roadmap of Murine Spermatogenesis Defined by Single-Cell RNA-Seq. Developmental Cell 2018;46:651- 667.e10.*
17	Guo F, Yan L, Guo H, Li L, Hu B, Zhao Y, Yong J, Hu Y, Wang X, Wei Y, et al. The Transcriptome and DNA Methylome Landscapes of Human Primordial Germ Cells. Cell 2015;161:1437–1452.*
18	Guo J, Grow EJ, Mlcochova H, Maher GJ, Lindskog C, Nie X, Guo Y, Takei Y, Yun J, Cai L, et al. The adult human testis transcriptional cell atlas. Cell Res 2018;28:1141– 1157.*
19	Guo J, Grow EJ, Yi C, Mlcochova H, Maher GJ, Lindskog C, Murphy PJ, Wike CL, Carrell DT, Goriely A, et al. Chromatin and Single-Cell RNA-Seq Profiling Reveal Dynamic Signaling and Metabolic Transitions during Human Spermatogonial Stem Cell Development. Cell Stem Cell 2017;21:533-546.e6.*
20	Hammoud SS, Low DHP, Yi C, Lee CL, Oatley JM, Payne CJ, Carrell DT, Guccione E, Cairns BR. Transcription and imprinting dynamics in developing postnatal male germline stem cells. Genes & development 2015;29:2312–2324.*
21	He X, Li W, Wu H, Lv M, Liu W, Liu C, Zhu F, Li C, Fang Y, Yang C, et al. Novel homozygous CFAP69 mutations in humans and mice cause severe asthenoteratospermia with multiple morphological abnormalities of the sperm flagella. Journal of Medical Genetics 2019;56:96–103.
22	Hermann BP, Cheng K, Singh A, La Roa-De Cruz L, Mutoji KN, Chen I-C, Gildersleeve H, Lehle JD, Mayo M, Westernströer B, et al. The Mammalian Spermatogenesis Single-Cell Transcriptome, from Spermatogonial Stem Cells to Spermatids. Cell Reports 2018;25:1650-1667.e8.*
23	 Höben IM, Hjeij R, Olbrich H, Dougherty GW, Nöthe-Menchen T, Aprea I, Frank D, Pennekamp P, Dworniczak B, Wallmeier J, et al. Mutations in C11orf70 Cause Primary Ciliary Dyskinesia with Randomization of Left/Right Body Asymmetry Due to Defects of Outer and Inner Dynein Arms. American journal of human genetics 2018;102:973–984.*
24	Hu Z, Li Z, Yu J, Tong C, Lin Y, Guo X, Lu F, Dong J, Xia Y, Wen Y, et al. Association analysis identifies new risk loci for non-obstructive azoospermia in Chinese men. Nat Commun;5:1–7.*
25	Irie N, Weinberger L, Tang WWC, Kobayashi T, Viukov S, Manor YS, Dietmann S, Hanna JH, Surani MA. SOX17 Is a Critical Specifier of Human Primordial Germ Cell Fate. Cell 2015;160:253–268.*
26	Kherraf Z-E, Amiri-Yekta A, Dacheux D, Karaouzène T, Coutton C, Christou-Kent M, Martinez G, Landrein N, Le Tanno P, Fourati Ben Mustapha S, et al. A Homozygous Ancestral SVA-Insertion-Mediated Deletion in WDR66 Induces Multiple Morphological Abnormalities of the Sperm Flagellum and Male Infertility. American journal of human genetics 2018;103:400–412.
27	Kosova G, Scott NM, Niederberger C, Prins GS, Ober C. Genome-wide Association Study Identifies Candidate Genes for Male Fertility Traits in Humans. The American Journal of Human Genetics 2012;90:950–961.*
28	Laurentino S, Heckmann L, Di Persio S, Li X, Meyer Zu Hörste G, Wistuba J, Cremers J- F, Gromoll J, Kliesch S, Schlatt S, et al. High-resolution analysis of germ cells from men with sex chromosomal aneuploidies reveals normal transcriptome but impaired imprinting. Clinical epigenetics 2019;11:127.

29	Li L, Dong J, Yan L, Yong J, Liu X, Hu Y, Fan X, Wu X, Guo H, Wang X, et al. Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. Cell Stem Cell 2017;20:858-873.e4.*
30	Lima AC, Carvalho F, Gonçalves J, Fernandes S, Marques PI, Sousa M, Barros A,
	Seixas S, Amorim A, Conrad DF, et al. Rare double sex and mab-3-related
	transcription factor 1 regulatory variants in severe spermatogenic failure. Andrology 2015;3:825–833.*
31	Lopes AM, Aston KI, Thompson E, Carvalho F, Gonçalves J, Huang N, Matthiesen R,
	Noordam MJ, Quintela I, Ramu A, et al. Human spermatogenic failure purges
	the gene DMRT1. PLoS genetics 2013:0:e1003349 *
32	Lukassen S. Bosch F. Ekici AB. Winterpacht A. Characterization of germ cell
02	differentiation in the male mouse through single-cell RNA sequencing. Sci Rep:8:1–
	7.*
33	Martinez G, Kherraf Z-E, Zouari R, Fourati Ben Mustapha S, Saut A, Pernet-Gallay K,
	Bertrand A, Bidart M, Hograindleur JP, Amiri-Yekta A, et al. Whole-exome
	sequencing identifies mutations in FSIP2 as a recurrent cause of multiple
	morphological abnormalities of the sperm flagella. Human reproduction (Oxford,
3/	Eliyidilu) 2010,00.1970-1904. Mitsunada S. Odajima I. Yawata S. Shioda K. Owa C. Isselbacher K.I. Hanna, IH. Shioda
54	T. Relevance of iPSC-derived human PGC-like cells at the surface of embryoid
	bodies to prechemotaxis migrating PGCs. Proceedings of the National Academy of
	Sciences of the United States of America 2017;114:E9913-E9922.*
35	Nakamura S, Miyado M, Saito K, Katsumi M, Nakamura A, Kobori Y, Tanaka Y, Ishikawa
	H, Yoshida A, Okada H, et al. Next-generation sequencing for patients with non-
	obstructive azoospermia: implications for significant roles of monogenic/oligogenic
36	Mutations. Androiogy 2017;5:824–831.
50	infertility: snapshot on morphological abnormalities of the sperm flagellum. Basic and
	clinical andrology 2019;29:2.
37	Olbrich H, Schmidts M, Werner C, Onoufriadis A, Loges NT, Raidt J, Banki NF,
	Shoemark A, Burgoyne T, Al Turki S, et al. Recessive HYDIN Mutations Cause
	Primary Ciliary Dyskinesia without Randomization of Left-Right Body Asymmetry. The
20	American Journal of Human Genetics 2012;91:672–684."
30	Leblond G. O'Toole F. Hara C. et al. Ktu/PF13 is required for cytoplasmic pre-
	assembly of axonemal dyneins. Nature 2008;456:611–616.*
39	Paff T, Loges NT, Aprea I, Wu K, Bakey Z, Haarman EG, Daniels JMA, Sistermans EA,
	Bogunovic N, Dougherty GW, et al. Mutations in PIH1D3 Cause X-Linked Primary
	Ciliary Dyskinesia with Outer and Inner Dynein Arm Defects. The American Journal of
40	Human Genetics 2017;100:160–168.*
40	Ruf-Zamojski F, Fribourg M, Ge Y, Nair V, Pincas H, Zaslavsky E, Nudelman G, Tuminelle S I, Wetensbe H, Turgeon JL, et al. Regulatory Architecture of the L&T2
	Gonadotrope Cell Underlying the Response to Gonadotropin-Releasing Hormone
	Front. Endocrinol. 2018;9:34.*
41	Sasaki K, Yokobayashi S, Nakamura T, Okamoto I, Yabuta Y, Kurimoto K, Ohta H,
	Moritoki Y, Iwatani C, Tsuchiya H, et al. Robust In Vitro Induction of Human Germ
	Cell Fate from Pluripotent Stem Cells. Cell Stem Cell 2015;17:178–194.*
42	Sato Y, Jinam T, Iwamoto T, Yamauchi A, Imoto I, Inoue I, Tajima A. Replication study
	and meta-analysis of numan nonopstructive azoospermia in Japanese populations. Biology of reproduction 2013:88:87 *
43	Sohni A. Tan K. Song H-W. Burow D. Rooii DG de Laurent L. Heigh T-C. Rahah R.
75	Hammoud SS. Vicini E. et al. The Neonatal and Adult Human Testis Defined at the
	Single-Cell Level. Cell Reports 2019;26:1501-1517.e4.*
44	Sugawa F, Araúzo-Bravo MJ, Yoon J, Kim K-P, Aramaki S, Wu G, Stehling M, Psathaki

	OE, Hübner K, Schöler HR. Human primordial germ cell commitment in vitro associates with a unique PRDM14 expression profile. The EMBO journal 2015;34:1009–1024.*
45	Tang S, Wang X, Li W, Yang X, Li Z, Liu W, Li C, Zhu Z, Wang L, Wang J, et al. Biallelic Mutations in CFAP43 and CFAP44 Cause Male Infertility with Multiple Morphological Abnormalities of the Sperm Flagella. The American Journal of Human Genetics 2017;100:854–864.*
46	Tang WWC, Dietmann S, Irie N, Leitch HG, Floros VI, Bradshaw CR, Hackett JA, Chinnery PF, Surani MA. A Unique Gene Regulatory Network Resets the Human Germline Epigenome for Development. Cell 2015;161:1453–1467.*
47	Ta-Shma A, Hjeij R, Perles Z, Dougherty GW, Abu Zahira I, Letteboer SJF, Antony D, Darwish A, Mans DA, Spittler S, et al. Homozygous loss-of-function mutations in MNS1 cause laterality defects and likely male infertility. PLoS genetics 2018;14:e1007602.*
48	van der Bijl N, Röpke A, Biswas U, Wöste M, Jessberger R, Kliesch S, Friedrich C, Tüttelmann F. Mutations in the stromal antigen 3 (STAG3) gene cause male infertility due to meiotic arrest. Human reproduction (Oxford, England) 2019;34:2112–2119.
49	Wang G, Guo Y, Zhou T, Shi X, Yu J, Yang Y, Wu Y, Wang J, Liu M, Chen X, et al. In- depth proteomic analysis of the human sperm reveals complex protein compositions. Journal of Proteomics 2013;79:114–122.
50	Wang M, Liu X, Chang G, Chen Y, An G, Yan L, Gao S, Xu Y, Cui Y, Dong J, et al. Single-Cell RNA Sequencing Analysis Reveals Sequential Cell Fate Transition during Human Spermatogenesis. Cell Stem Cell 2018;23:599-614.e4.*
51	Wyrwoll MJ, Temel ŞG, Nagirnaja L, Oud MS, Lopes AM, van der Heijden GW, Rotte N, Wistuba J, Wöste M, Ledig S, et al. Biallelic mutations in M1AP are a frequent cause of meiotic arrest leading to male infertility, 2019.
52	Yang F, Silber S, Leu NA, Oates RD, Marszalek JD, Skaletsky H, Brown LG, Rozen S, Page DC, Wang PJ. TEX11 is mutated in infertile men with azoospermia and regulates genome-wide recombination rates in mouse. EMBO molecular medicine 2015;7:1198–1210.*
53	Zhao H, Xu J, Zhang H, Sun J, Sun Y, Wang Z, Liu J, Ding Q, Lu S, Shi R, et al. A Genome-wide Association Study Reveals that Variants within the HLA Region Are Associated with Risk for Nonobstructive Azoospermia. The American Journal of Human Genetics 2012;90:900–906.*
54	Zimmermann C, Stévant I, Borel C, Conne B, Pitetti J-L, Calvel P, Kaessmann H, Jégou B, Chalmel F, Nef S. Research resource: the dynamic transcriptional profile of sertoli cells during the progression of spermatogenesis. Molecular endocrinology (Baltimore, Md.) 2015;29:627–642*









Walk Through **Publications** About Search



Page content

Datasets containing all of the search terms Search parameters Plots Datasets Detailed results for gene or Id search Search for data sets

Datasets containing all of the search terms

Search parameters

• Genes, Ensemble Ids, rs numbers, ... o stag3

Plots

Search

Return all data sets containing O at least one search term O at least one search term per filter group Ill search terms

Genes, proteins, Ensembl id, rs number, ...

Input comma separated list

stag3,

Gene symbol: include alias / previous symbols Include ensembl gene / transcript ids <-> gene symbol

OMICS

- Exome (9) Genome (28) Methylome (10)
- more

Analysis type

Data selection All
Filtered table

Number Of Data Sets Found Per Publication

Hover over / click on slices or legend for more information



Occurrences Of Genes Per Publication

Hover over / click on slices or legend for more information



Analysis type Bulk RNA-Sequencing (12) Functional Study (10) Genome-wide Association Study (9) Immore					 WANG et al. 2018: 1 SASAKI et al. 2015: 1 CHALMEL et al. 2012: 1 CHEN et al. 2018: 1 WANG et al. 2013: 1 CASTILLO et al. 2018: 1 				GKOUNTELA et al. 2015: 1 HAMMOUD et al. 2015: 1 TANG et al. 2015: 1 SASAKI et al. 2015: 1 CASTILLO et al. 2018: 1 CHALMEL et al. 2012: 1 Highcharts.com
Conditions									
	Datase	ts							
□	Line 1 - 25 o	of 25							
(39) <u>HP:0012265:</u> Ciliary dyskinesia (6)									
more	<							\geq	
Species	<								>
Chlamydomonas (1) Drosophila (1) Human (94)			Access Data	Publication ≟↓	Data Set ≟↓	Important Genes Or Proteins 🛓	Species =]	Omics ≟↓	Analysis Type 🛓
	Link	Access	In Rgv	Search	Search	Search	Search	Filter	✓ Filter ✓
Tissue types		A				COV17 DUM AD1		T	

Publications

HERMANN et al. 2018: 5

IRIE et al. 2017: 2

GUO et al. 2015: 2

GUO et al. 2018: 1

TANG et al. 2015: 1

Ll et al. 1017: 1

• VAN DER BIJL et al. 2019: 4

GKOUNTELA et al. 2015: 1

HAMMOUD et al. 2015: 1

LUKASSEN et al. 2018: 1

Issue types adrenal (1) attached fraction (2)		RGV	IRIE et al. 2017, Cell.	RNA-Seq results for human PGCLCs, PGCs, and TCam-2 Seminoma cells	SOX17 BLIMP1	Human	Transcriptome	Single-cell RNA- Sequencing
 BTO_0001099: blastocyst (1) more Cell types 	6	THE	GKOUNTELA et al. 2015, Cell.	Transcriptional landmarks of human prenatal germline development		Human	Transcriptome	Bulk RNA- Sequencing
advanced germline cell (2) CL_0000081: blood cell (4) CL_0000711: cumulus cell (1)	6		GUO et al. 2018, Cell Research.	Cell partitioning through the analysis of single cell transcriptomes		Human	Transcriptome	Single-cell RNA- Sequencing
more Search Sear	6	RGV	IRIE et al. 2017, Cell.	Comparison of conventional and naive (4i) hESCs	SOX17 BLIMP1	Human	Transcriptome	Single-cell RNA- Sequencing
	6	THERE	LI et al. 1017, Cell Stem Cell.	The Gene Expression Landscapes of Human FGCs and their Niche Cells in the Gonads		Human	Transcriptome	Single-cell RNA- Sequencing
	6	tic	HAMMOUD et al. 2015, Genes & Development.	Genomic profiling of developing SSCs - strand specific RNA - seq	THY1 KIT OCT4 ID4 GFRa1	Mouse Vasa-GFP Oct4-GFP⁺ Id4-GFP⁺ 	Transcriptome	Bulk RNA- Sequencing



				(Log2(normalized read counts))				
6	RGV	GKOUNTELA et al. 2015, Cell.	Transcriptional landmarks of human prenatal germline development	Table S4. RNA-Seq Correlation Scores, Related to Figure 1. Expressed Transcripts by Module	Human	<u>STAG3</u>	ENSG0000066923	1
6	RGV	HAMMOUD et al. 2015, Genes & Development.	Genomic profiling of developing SSCs - strand specific RNA - seq	Table S2A: FPKM values for datasets generated using Cuffdiff 2.1.1	Mouse	<u>Stag3</u>		1
	RGV	LUKASSEN et al. 2018, Scientific Reports.	Use of scRNA-seq to establish expression profiles of 2,550 germ cells from the adult mouse testis	Supplementary table S2	Mouse	<u>Stag3</u>		2
6	RGV	GUO et al. 2015, Cell.	Transcriptional Profiles of Human PGCs	Table S2.1. Z-score of FPKM in 3 cell types	Human	<u>STAG3</u>		1
	R	TANG et al. 2015, Cell.	RNA-Seq Reveals Unique Transcriptional States of hPGCs	Table S1. Gene expression levels in replicates of indicated samples by RNA- Seq (Log2(normalised read counts)), Related to Figure 2	Human	<u>STAG3</u>		1
ð	TIG	LI et al. 1017, Cell Stem Cell.	The Gene Expression Landscapes of Human FGCs and their Niche Cells in the Gonads	Table S3: Differentially Expressed Genes of Female and Male FGCs	Human	<u>STAG3</u>		2
6	RGV	GUO et al. 2015, Cell.	Transcriptional Profiles of Human PGCs	Table S3.1. Z-Score of FPKM in male PGCs.	Human	STAG3		1
6	RGV	GUO et al. 2015, Cell.	Transcriptional Profiles of Human PGCs	Table S3.2. Z-Score of FPKM in female PGCs.	Human	<u>STAG3</u>		1
		GUO et al. 2015, Cell.	Dynamic Methylome Analysis of Human PGCs Revealed a Global DNA Hypomethylation Pattern	Table S6.2. Meiosis Genes (107)	Human	<u>STAG3</u>		1
	T	WANG et al. 2018, Cell Stem Cell.	Global Transcriptional Profiling of Adult Human Testicular Cells and Cell Type Identification	Table S1.6: DEG power>0.8	Human	<u>STAG3</u>		1
		VAN DER BIJL et al. 2019, Human Reproduction.	Whole exome sequencing of patient M870 with STAG3 variant and meiotic arrest	Table S3: Top 10 list of the population sampling probability (PSAP) results. The two variants in the stromal antigen 3 (STAG3) gene are highlighted	Human	<u>STAG3</u>		2
		CHEN et al. 2018, Cell Research.	Alternative splicing genes and differentially expressed splicing factors during mouse spermatogenesis	Table S7.1: Alternative splicing genes in single- cell, Related to Figure 6	Mouse	<u>Stag3</u>		210
		IRIE et al. 2017, Cell.	Comparison of conventional and naive (4i) hESCs	Table S2.1: Transcripts expression levels in replicates of 4i WIS2 hESC and Conventional WIS2 hESC (log2(normalized read counts))	Human	<u>STAG3</u>		
		GUO et al. 2018, Cell Research.	Cell partitioning through the analysis of single cell transcriptomes	Table S2: K-means Clustering of Genes that Show Differential Expression along Germ Cell Development.	Human	<u>STAG3</u>		1



About Publications Search Walk Through

Contact Login

Page content

Datasets containing all of the search terms Search parameters Plots Datasets Sound in tables of the returned data sets (count > = 2)

Search

Return all data sets containing O at least one search term O at least one search term per filter

Search for data sets

Datasets containing all of the search terms

Search parameters

Conditions

HP:0011961: Non-obstructive azoospermia
 HP:0011962: Obstructive azoospermia
 HP:0000027: Azoospermia

Plots

group all search terms

Genes, proteins, Ensembl id, rs number, ...

Input comma separated list

Gene symbol: include alias / previous symbols Include ensembl gene / transcript ids <-> gene symbol

OMICS

Analysis type Bulk RNA-Sequencing (12)

____Functional Study (8) Genome-wide Association Study Data selection All • Filtered table •

Number Of Data Sets Found Per Publication

Hover over / click on slices or legend for more information

Publications

HU et al. 2014, Nature Communication.: 4
ZHAO et al. 2012, The American Journal of Human Genetics.: 4
VAN DER BIJL et al. 2019, Human Reproduction.: 4
SATO et al. 2013, Biology of Reproduction.: 3
LOPES et al. 2013, PLoS Genetics.: 3
WANG et al. 2018, Cell Stem Cell.: 2
ARAUJO et al. 2019, Andrology.: 2
ASTON and CARELL 2009, Journal of Andrology.: 1
ASTON et al. 2015, Andrology.: 1
NAKAMURA et al. 2017, Andrology.: 1
YANG et al. 2015, EMBO Molecular Medicine.: 1

Top20 Gene Occurrences In Distinct Data Sets

Hover over / click on bars for more information



(9)				 WYRWOLL et al. 2019, BUSCH et al. 2019, The 	bioRxiv (Preprint).: 1 e Journal of Clinical Endocrin	ology & Metabolism.: 1 Highcharts.com			TWENT'S MINIS	18-1°	oche	Highcharts.com	1
Conditions													
□	Datas	sets											
abnormality of male internal genitalia (39)	Line 1 - 2	29 of 29											
HP:0012206: Abnormal sperm motility (2)													
HP:0008669: Abnormal	<											>	
spermatogenesis (37)	<											>	
✓						Important							
Azoospermia (31) HP:0011961: Non- obstructive azoospermia (28)	Link	Access	Access Data In Rgv	PublicationŢ↓ Search	Data Set∓↓ Search	Genes Or Proteins ≟↓ Search	Search	Omics≟↓ Filter	Analysis Type≟↓ Filter ✓	HP:0011961 Non- obstructive azoospermia	HP:0011962 Obstructive azoospermia	HP:0000027 Azoospermia	
Oligography (F)				ARAUJO et al. 2019, Andrology.	Data set	DMRT1 REC8 TEX15 KLHL10 DNMT3B TEX14 SYCE1L	Human	Exome	Whole Exome Sequencing				
	B	a		ADALLIO at al. 2010	Validation	DMDT1 DECO	Luman	Conomo	Cangor				

Spermatogenesis maturation arrest (1) HP:0012265: Ciliary dyskinesia (6)		Andrology.	validation	TEX15 KLHL10 DNMT3B TEX14 SYCE1L	Turnan	Genome	Sequencing			
more Species Chlamydomonas (1) Human (91)		ASTON and CARELL 2009, Journal of Andrology.	A genome-wide association study of men with azoospermia and oligozoospermia		Human	Genome	Genome- wide Association Study	✓		
 Medaka (1) 		ASTON et al. 2010, Human Reproduction.	Candidate polymorphisms for association with oligozoospermia or azoospermia		Human	Genome	Targeted Genotyping	✓		
adrenal (1) attached fraction (2) BTO_0001099: blastocyst (1)	6	BUSCH et al. 2019, The Journal of Clinical	Data set	<u>FSHR</u>	Human	Genome	Targeted Genotyping	✓	✓	
E more Coll type Coll Type Coll Type Coll Type Coll Type Coll Type Coll Type Coll Type Coll Type Coll Type Coll Type Coll Type Coll Type Coll Type Coll Type Type Coll Type Type		Endocrinology & Metabolism.								

Genes found in tables of the returned data sets (count $> = 2$)										
Line 1 - 100 of 242 🗹 🔽 💌		~								
		\sim								
		> >								
Gene ≟↓	Data Sets Containing Gene In Any Table ≟↓									
Search	Search									
HLA-DRA	8									
CDC42BPA	5									
ABLIM1	5									
<u>TEX11</u>	4									
MEI1	4									
TSBP1	4									
<u>TEX15</u>	4									
<u>REC8</u>	4									
TMEM132E-CCT6B	4									
RGS9	4									
<u>MIR133B-IL17A</u>	4									
SOX5	4									
CRBN	3									
PDHA2	3									
LOC100289473	3									
	3									
CDATA 17	3									
DNAH10	3									
HADH	3									
NOTCH4	3									
DMRT1	3									
C14orf118-ESRRB	3									
<u>SYCP1</u>	3									
<u>TLE4</u>	3									
FANCM	3									
FOXN3	3									
COL5A1	3									
<u>LSM5</u>	3									
PRRC2A	3									
<u>TEX14</u>	3									
RBMXL2	3									
<u>SYCP2</u>	3									
<u>EPB41L3</u>	3									
<u>SMOX</u>	3									
<u>NFIA</u>	3									
PRDM9	3									
CADM1	3									
ESR2	3									
MSH4	3									
SYCP3	3									
ARL6	3									
KLHL10	3									



About Publications Search Walk Through

Contact Login

Page content

Description Disclaimer Histograms Table



Table S3: Top 10 list of the population sampling probability (PSAP) results. The two variants in the stromal antigen 3 (STAG3) gene are highlighted

VAN DER BIJL, N., et al. Mutations in the stromal antigen 3 (STAG3) gene cause male infertility due to meiotic arrest. Human Reproduction, 2019. https://doi.org/10.1093/humrep/dez204.

Description

Chr: Chromosome, Start: Genomic position, Ref: Reference allele, Gene.wgEncodeGencodeBasicV19: Gene name (GENCODE Release 19), Func.wgEncodeGencodeBasicV19: Regions (e.g. exonic, intronic) affected by variant (GENCODE Release 19), ExonicFunc.wgEncodeGencodeBasicV19: Exonic consequence (e.g. stopgain, frameshift insertion) of variant (GENCODE Release 19), AAChange.wgEncodeGencodeBasicV19: Amino acid changes caused by variant (GENCODE Release 19), mac63kFreq_ALL: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 database, 1000g2014sep_all: Variant allele frequency in Exome Aggregation Consortium (ExAC) 0.2 databasepart (Example frequency in Exome Aggregation Consortium (ExAC) 0

1000 Genomes Project phase 3 database, esp6500si_all: Variant allele frequency in NHLBI Exome Sequencing Project 6500 (ESP6500) database, Alt: Alternative allele, CADD_Phred: Variant combined annotation-dependent depletion (CADD) score, Geno: Genotype (hom = homozygous, het = heterozygous), Dz.Model: Disease model assumed

Disclaimer

The publication **Mutations in the stromal antigen 3 (STAG3) gene cause male infertility due to meiotic arrest.** by *Nils van der Bijl, Albrecht Röpke, Uddipta Biswas, Marius Wöste, Rolf Jessberger, Sabine Kliesch, Corinna Friedrich, Frank Tüttelmann* is published under an **open access license**: http://creativecommons.org/licenses/by-nc/4.0/. Permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Curation by the MFGA team

This data tables (Table S3: Top 10 list of the population sampling probability (PSAP) results. The two variants in the stromal antigen 3 (STAG3) gene are highlighted) has been extracted from publication or supplemental material. If not stated otherwise, title and description have been adopted. However, the format of the table has been adjusted to the MFGA requirements, allowing for filter and sort operations on the data in the web browser.



7	99796115	Ţ	STAG3	exonic	nonsynonymous SNV	STAG3:ENST00000317296.5:exon13:c.T1262G:p.L421R;STAG3:ENST00000426455.1:exon13:c.T120
1	54605318		CDCP2	exonic	frameshift insertion	CDCP2:ENST00000371330.1:exon4:c.1224dupC:p.M409fs
5	175811094		<u>NOP16</u>	exonic;UTR3;downstream	frameshift insertion	NOP16:ENST00000510123.1:exon5:c.583_584insAC:p.R195fs
12	974308		WNK1	exonic;intronic	frameshift insertion	WNK1:ENST00000530271.2:exon9:c.2173dupC:p.V724fs
Χ	122336600		<u>GRIA3</u>	exonic;intronic	frameshift insertion	GRIA3:ENST00000371266.1:exon4:c.382dupG:p.G127fs;GRIA3:ENST00000371264.3:exon3:c.382c
16	70896016	Α	HYDIN	exonic	frameshift deletion	HYDIN:ENST00000393567.2:exon69:c.11712delT:p.I3904fs
7	99796165	C	STAG3	exonic	stopgain	STAG3:ENST00000317296.5:exon13:c.C1312T:p.R438X;STAG3:ENST00000426455.1:exon13:c.C13
17	80022791	G	<u>DUS1L</u>	exonic	nonsynonymous SNV	DUS1L:ENST00000306796.5:exon2:c.C145G:p.L49V;DUS1L:ENST00000354321.7:exon1:c.C145G:p
19	41114171	C	LTBP4	exonic;UTR5	nonsynonymous SNV	LTBP4:ENST00000396819.3:exon8:c.C1202T:p.S401L;LTBP4:ENST00000204005.9:exon11:c.C1292
<						

Line 1 - 10 of 10



Publications Walk Through Search About

Login Contact

Page content

Datasets containing all of the search terms Search parameters Datasets Genes found in tables of the returned

data sets (count > = 2)

Search

Search for data sets

Datasets containing all of the search terms

Search parameters

 Analysis Type Single-cell RNA-Sequencing

 Species o Human

Tissue

• **BTO_0001363**: testis

Return all data sets containing O at least one search term O at least one search term per filter group Il search terms

Genes, proteins, Ensembl id, rs number, ...

Input comma separated list

Gene symbol: include alias / previous symbols Include ensembl gene / transcript ids <-> gene symbol

OMICS

Exome (7) Genome (28) Methylome (10) Cell OCL_0000216: Sertoli cell

Plots

Data selection All • Filtered table •

Number Of Data Sets Found Per Publication

Hover over / click on slices or legend for more information



Top20 Gene Occurrences In Distinct Data Sets

Hover over / click on bars for more information

3.5



more

- Analysis type Bulk RNA-Sequencing (12) ____Functional Study (8) Genome-wide Association Study (9)
- Single-cell RNA-Sequencing (23) more
- Conditions
- morphology (6) **HP:000025:** Functional abnormality of male internal genitalia (39) HP:0012265: Ciliary dyskinesia (6)
- more
- Species
- Chlamydomonas (1) Human (91)



Datasets

Line 1 - 4 of 4

Publications

- GUO et al. 2018, Cell Research.: 1
- WANG et al. 2018, Cell Stem Cell.: 1
- LAURENTINO et al. 2019, Clinical Epigenetics.: 1
- SOHNI et al. 2019, Cell Reports.: 1

Highcharts.com

Highcharts.com

 \wedge

Medaka (1)

more

Tissue types adrenal (1) Lattached fraction (2) BTO_0001099: blastocyst (1) BTO_0001363: testis (36) more

Cell types advanced germline cell (2) CL_0000081: blood cell (4) LCL_0000711: cumulus cell (1) CL_0000216: Sertoli cell (10) more

Search

Link	Access	Access Data In Rgv	Publication ≟↓ Search	Data Set≞↓ Search	Important Genes Or Proteins L Search	Search	Omics≟↓ Filter	Analysis Type ≟↓ Filter	Testis	Sertoli cell
	ð		GUO et al. 2018, Cell Research.	Cell partitioning through the analysis of single cell transcriptomes		Human	Transcriptome	Single-cell RNA- Sequencing	✓	
		THERE	WANG et al. 2018, Cell Stem Cell.	Global Transcriptional Profiling of Adult Human Testicular Cells and Cell Type Identification	HMGA1 PIWIL4 TEX29 SCML1 CCDC112	Human	Transcriptome	Single-cell RNA- Sequencing		
			LAURENTINO et al. 2019, Clinical Epigenetics.	Single-cell RNA-Sequencing from Klinefelter testicular tissue		Human	Transcriptome	Single-cell RNA- Sequencing		
			SOHNI et al. 2019, Cell Reports.	Germ Cell Subsets in Neonatal Human Testes and Somatic Cell Gene Expression		Human	Transcriptome	Single-cell RNA- Sequencing		



PSIP1

<u>PSIPI</u>	
PXDN	3
<u>CYSTM1</u>	3
IGFBP2	3
TUBA1C	3
<u>UXT</u>	3
NAB1	3
<u>PIP5K1B</u>	3
<u>CST3</u>	3
PDIA3	3
PLN	3
ASAH1	3
<u>CBLN1</u>	3
FAM60A	3
TACC1	3
<u>TMEM97</u>	3
<u>UBE2A</u>	3
LMO4	3
LRCH4	3
MYL12A	3
NSA2	3
REXO2	3
<u>CBX1</u>	3
<u>CRIP1</u>	3
DYNLT3	3
<u>SLCO3A1</u>	3
<u>TMEM98</u>	3
<u>CRIP2</u>	3
<u>SEC61B</u>	3
CAP1	3
<u>EFHD1</u>	3
<u>GNG5</u>	3
<u>RRBP1</u>	3
<u>ZFP36L1</u>	3
HSPA2	3
ATP5B	3
<u>AES</u>	3
<u>SCN3A</u>	3
DNPH1	3
<u>FAM162A</u>	3
<u>HLA-B</u>	3
LINC00657	3

255 Supplemental References

- Apache Software Foundation. Apache Tiles Framework The Composite View Pattern.
 Retrieved 17 October 2019. from https://tiles.apache.org/framework/tutorial/pattern.html,
 2017a.
- Apache Software Foundation. *Apache Tiles Home*. Retrieved 09 January 2020. from https://tiles.apache.org/, 2017b.
- Budde R, Kautz K, Kuhlenkamp K, Züllighoven H. What is prototyping? *Info Technology & People* 1992;6:89–95.
- Chen PP-S. The entity-relationship model---toward a unified view of data. ACM Trans.
 Database Syst. 1976;1:9–36.
- Chinosi M, Trombetta A. BPMN: An introduction to the standard. *Computer Standards & Interfaces* 2012;**34**:124–134.
- 267 Dyer RJT. *Learning MySQL and MariaDB: Heading in the right direction with MySQL and* 268 *MariaDB*. Sebastopol, CA: O'Reilly, 2015.
- Gordon VS, Bieman JM. Rapid prototyping: lessons learned. *IEEE Softw.* 1995;**12**:85–95.
- Highsoft AS. Interactive JavaScript charts for your webpage | Highcharts. Retrieved 09
 January 2020. from https://www.highcharts.com/, 2020.
- JS Foundation. *jQuery*. Retrieved 09 January 2020. from https://jquery.com/, 2020a.
- JS Foundation. *jQuery.ajax* | *jQuery API Documentation*. Retrieved 09 January 2020. from https://api.jquery.com/jquery.ajax/, 2020b.
- Maguire M, Bevan N. User Requirements Analysis. In: Hammond J, Gross T, Wesson J
 (eds). Usability. Boston, MA: Springer US, 2002.
- MariaDB Corporation. Open Source Database (RDBMS) for the Enterprise | MariaDB.
 Retrieved 09 January 2020. from https://mariadb.com/de/, 2020.
- Oracle Corporation. JavaServer Pages Technology. Retrieved 09 December 2019. from
 https://www.oracle.com/technetwork/java/jsp-138432.html, 2019.
- 281 Otto M, Thornton J, contributors aB. *Bootstrap*. Retrieved 09 January 2020. from 282 https://getbootstrap.com/, 2020a.
- 283 Otto M, Thornton J, contributors aB. *Introduction*. Retrieved 09 January 2020. from 284 https://getbootstrap.com/docs/4.4/getting-started/introduction/, 2020b.
- Parker, Zachary; Poe, Scott; Vrbsky, Susan V. (2013): Comparing NoSQL MongoDB to an
 SQL DB. In: Ashraf Saad (Hg.): Proceedings of the 51st ACM Southeast Conference on ACMSE '13. the 51st ACM Southeast Conference. Savannah, Georgia, 04.04.2013 06.04.2013. New York, New York, USA: ACM Press, S. 1.
- Pivotal Software Inc. Spring Framework. Retrieved 10 December 2019. from
 https://spring.io/projects/spring-framework, 2019a.
- Pivotal Software Inc. Spring Security. Retrieved 10 December 2019. from
 https://spring.io/projects/spring-security, 2019b.

- 293 Pivotal Software Inc. Spring. Retrieved 09 January 2020. from https://spring.io/, 2020.
- Porter B, van Zyl J, Lamy O. *Maven Welcome to Apache Maven*. Retrieved 09 January
 2020. from https://maven.apache.org/, 2020.
- 296 Red Hat Inc. *Hibernate Getting Started Guide*. Retrieved 17 October 2019. from 297 https://docs.jboss.org/hibernate/orm/5.4/quickstart/html_single/, 2018.
- Red Hat Inc. *Hibernate. Everything data. Hibernate.* Retrieved 09 January 2020. from https://hibernate.org/, 2020.
- Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, Goldberg LJ, Eilbeck K,
 Ireland A, Mungall CJ, et al. The OBO Foundry: coordinated evolution of ontologies to
 support biomedical data integration. *Nat Biotechnol* 2007;**25**:1251–1255.
- van der Bijl N, Röpke A, Biswas U, Wöste M, Jessberger R, Kliesch S, Friedrich C,
 Tüttelmann F. Mutations in the stromal antigen 3 (STAG3) gene cause male infertility due
 to meiotic arrest. *Human reproduction (Oxford, England)* 2019;**34**:2112–2119.