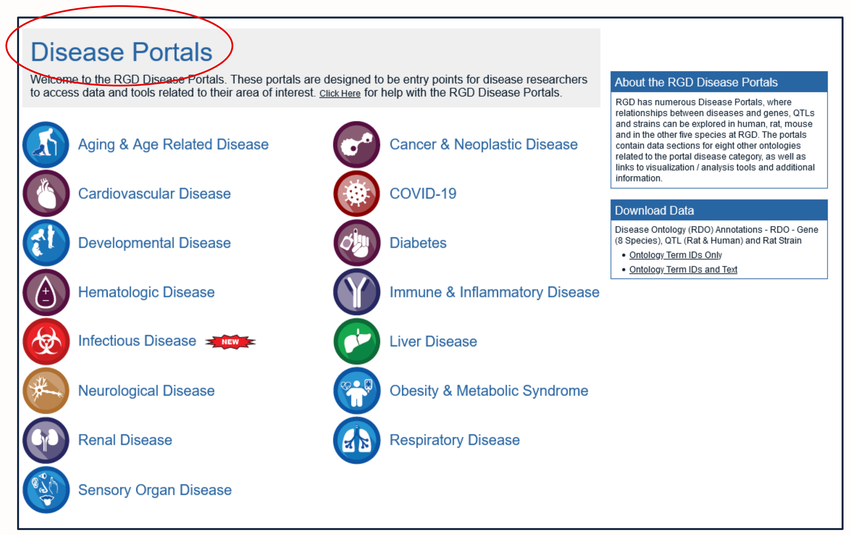
Ontology Analysis of Infectious Disease Genes at the Rat Genome Database

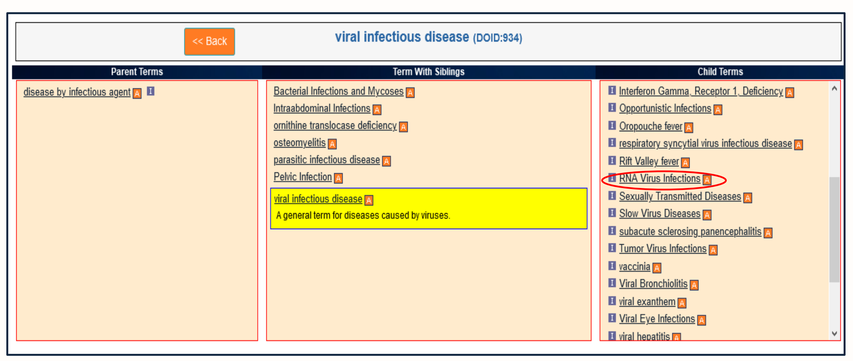
Shur-Jen Wang , Jennifer R Smith , G Thomas Hayman, Stan JF Laulederkind, Mary L Kaldunski, Mahima Vedi, Monika Tutaj, Matthew J Hoffman, Morgan Hill, Wendy Demos, Jeffrey L De Pons, Melinda R Dwinell*, Anne E Kwitek and the RGD team

Rat Genome Database, Department of Biomedical Engineering; *Department of Physiology, Medical College of Wisconsin Milwaukee, WI, 53226, USA

Abstract

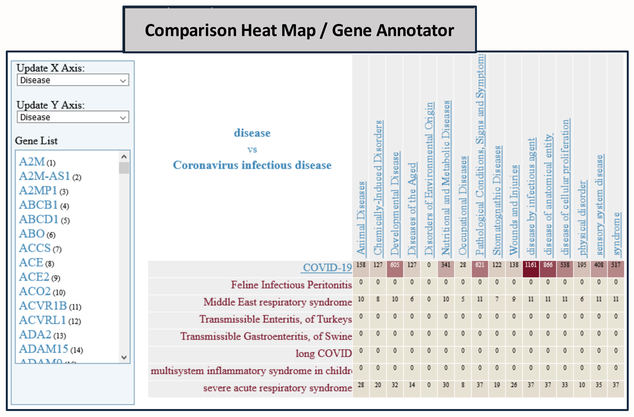
The Rat Genome Database (RGD) is the premier site for integrating rat genomic and genetic data with curated biological information. Multiple ontologies are used to annotate genes, strains, and QTLs. RGD has focused its manual curation efforts to annotating these genome objects with disease ontology terms. The RGD Disease Portals, with periodic addition of new portals and disease groups since they were launched, provide easy access to all genes, QTLs, strains, phenotypes, and biological functions associated with specific disease areas. The newly added "Infectious Diseases Portal" is the 15th portal to date. Here we present disease ontology enrichment analysis across selected disease gene sets within the Infectious Disease Portal using the Multi-Ontology Enrichment Tool (MOET). As expected, the COVID-19 disease gene set is highly enriched with coronavirus infectious disease and related diseases. Between coronavirus infectious disease and respiratory tract infections, several disease areas are also enriched among the COVID-19 genes. The enrichment profile of the COVID-19 gene set is compared to those of "Middle East respiratory syndrome" (MERS) and "severe acute respiratory syndrome" (SARS). Additional analysis of phenotypes and Gene Ontology were also performed using MOET and other tools at RGD. Using the Gene Annotator tool, we showed that close to one third of the COVID-19 genes were also annotated with "abnormality of nervous system" phenotypes. In this poster, we demonstrate the utility of RGD tools in analyzing coronavirus disease genes. We can look at not only the enrichment profile of a disease gene set, but also the gene distribution between two selected ontologies.



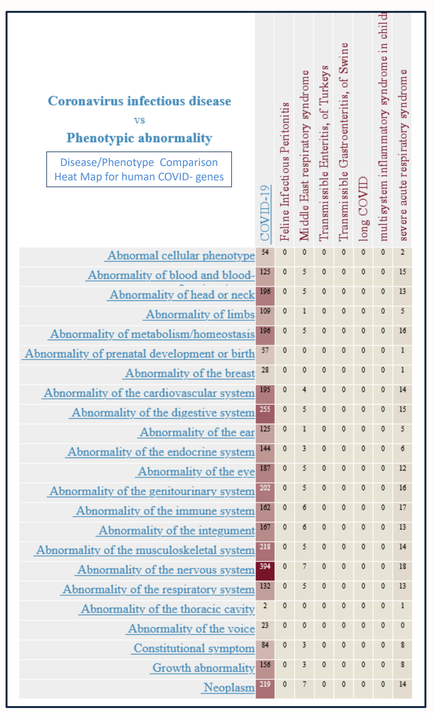


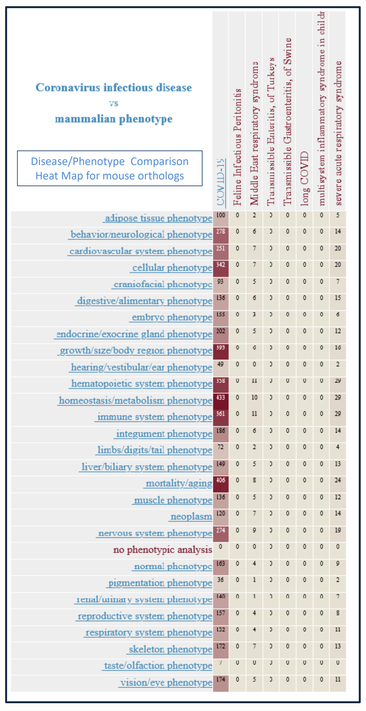
COVID-19/MOET			COVID-19 /CTD set analyzer		
count	corrected p	RGD disease	count	Corrected P	MESH disease
1164	0.00E+00	disease by infectious agent	380	1.30E-122	Neoplasms
1164	0.00E+00	Coronavirus infectious disease	286	1.50E-115	Neoplasms by Histologic Type
1164	0.00E+00	COVID-19	315	1.24E-110	Digestive System Diseases
1164	0.00E+00	Coronaviridae Infections	264	1.72E-108	Liver Diseases
1164	0.00E+00	RNA Virus Infections	308	1.42E-98	Neoplasms by Site
1164	0.00E+00	Nidovirales Infections	224	1.24E-95	Neoplasms, Glandular and Epithelial
1164	0.00E+00	viral infectious disease	187	2.99E-85	Immune System Diseases
153	6.83E-37	hepatocellular carcinoma	195	9.60E-85	Carcinoma
153	8.12E-37	liver carcinoma	305	2.32E-80	Pathological Conditions, Signs and Symptoms
153	7.37E-36	liver cancer	195	5.01E-78	Digestive System Neoplasms
125	9.88E-36	Respiratory Tract Infections	122	1.96E-76	Carcinoma, Hepatocellular
159	2.01E-35	hepatobiliary system cancer	162	1.09E-74	Adenocarcinoma
312	5.09E-32	liver disease	242	2.67E-74	Pathologic Processes
314	1.67E-29	hepatobiliary disease	136	4.89E-73	Liver Neoplasms
164	2.04E-29	Liver Neoplasms	198	3.50E-71	Skin and Connective Tissue Diseases
215	1.87E-27	adenocarcinoma	156	1.65E-65	Respiratory Tract Diseases
287	1.00E-26	primary immunodeficiency disease	245	8.49E-65	Nervous System Diseases
219	2.30E-26	autoimmune disease	181	3.69E-63	Female Urogenital Diseases and Pregnancy Complications
267	3.20E-26	Inflammation	168	1.81E-62	Skin Diseases
331	1.09E-25	immune system disease	168	1.03E-59	Female Urogenital Diseases

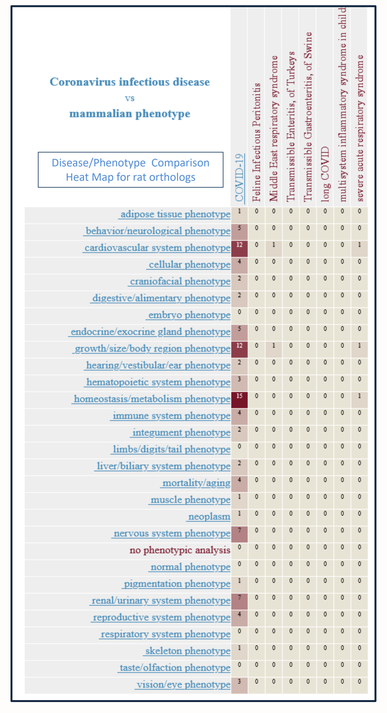
SARS/MOET			MERS/MOET		
count	corrected p	RGD disease	count	corrected p	RGD disease
90	9.38E-243	severe acute respiratory syndrome	19	2.10E-61	Middle East respiratory syndrome
90	3.17E-139	Respiratory Tract Infections	15	4.43E-29	severe acute respiratory syndrome
90	8.59E-102	Coronavirus infectious disease	11	9.86E-20	Parasitic Skin Diseases
90	9.23E-102	Coronaviridae Infections	19	2.48E-19	Coronavirus infectious disease
90	9.89E-102	Nidovirales Infections	19	2.53E-19	Coronaviridae Infections
90	1.32E-91	RNA Virus Infections	19	2.56E-19	Nidovirales Infections
90	1.90E-87	viral infectious disease	16	2.91E-19	Respiratory Tract Infections
90	3.09E-81	disease by infectious agent	11	4.97E-19	Euglenozoa Infections
90	2.65E-70	respiratory system disease	13	6.51E-19	Infectious Skin Diseases
58	6.80E-60	bacterial infectious disease	12	7.66E-18	parasitic protozoa infectious disease
59	1.11E-57	Bacterial Infections and Mycoses	14	1.30E-17	asthma
71	5.28E-46	Inflammation	10	1.41E-17	leishmaniasis
40	5.28E-46	pneumonia	19	3.11E-17	RNA Virus Infections
63	4.81E-43	autoimmune disease	13	3.47E-17	parasitic infectious disease
70	1.12E-41	primary immunodeficiency disease	12	4.19E-17	primary bacterial infectious disease
35	1.29E-40	Gram-Positive Bacterial Infections	14	4.79E-17	respiratory allergy
62	9.43E-39	lung disease	13	6.55E-17	pneumonia
30	2.04E-38	bacterial pneumonia	11	1.16E-16	tuberculosis
30	2.04E-38	Bacterial Lung Diseases	11	1.52E-16	uveitis
47	2.20E-38	Postoperative Complications	11	1.98E-16	Herpesviridae Infections



COVID-19/human disease genes			COVID-19 /mouse orthologs		
count	corrected p	Human phenotype	count	Corrected p	Mammalian phenotype
562	0.00E+00	Phenotypic abnormality	209	1.12E-12	abnormal adaptive immunity
395	4.79E-193	Abnormality of the nervous system	205	3.72E-12	abnormal cell-mediated immunity
345	1.34E-149	Abnormal nervous system physiology	203	6.11E-12	abnormal immune cell physiology
180	1.42E-73	Behavioral abnormality	198	2.25E-11	abnormal leukocyte physiology
93	1.28E-55	Schizophrenia	219	1.04E-10	abnormal hematopoietic system physiology
204	2.33E-37	Neurodevelopmental abnormality	165	2.20E-10	abnormal immune serum protein physiology
218	6.96E-29	Abnormality of higher mental function	293	3.64E-10	abnormal immune system physiology
168	1.29E-25	Intellectual disability	148	7.96E-10	abnormal professional antigen presenting cell physiology
255	7.10E-24	Abnormality of the digestive system	203	2.21E-09	abnormal blood cell physiology
220	1.64E-22	Abnormal nervous system morphology	151	4.21E-09	abnormal lymphocyte physiology
205	1.03E-21	Morphological central nervous system abnormality	79	5.18E-09	abnormal IgG level
216	3.74E-21	Neoplasm by anatomical site	104	1.00E-08	abnormal response to infection
219	4.73E-21	Neoplasm	105	1.03E-08	abnormal response to antigen
218	1.28E-20	Abnormality of the musculoskeletal system	79	1.41E-08	autoimmune response
276	2.23E-20	Mode of inheritance	93	1.57E-08	abnormal immunoglobulin level
190	9.57E-18	Abnormality of the gastrointestinal tract	79	1.75E-08	abnormal self tolerance
156	3.23E-15	Growth abnormality	99	1.88E-08	abnormal susceptibility to infection
156	4.93E-15	Abnormality of the musculature	79	4.11E-08	abnormal immune tolerance
109	6.57E-15	Abnormality of limbs	96	8.34E-08	abnormal humoral immune response
121	8.02E-15	Seizure	107	9.23E-08	abnormal cytokine secretion







RGD is funded by grant HL64541 from the National Heart, Lung, and Blood Institute