Scientific Foundation SPIROSKI, Skopje, Republic of North Macedonia
South East European Journal of Immunology. 2024 May 30; 7:67-80.
https://doi.org/10.3889/seejim.2024.6064
eISSN: 1857-9388
Category: *Original article*
Section: *Immunogenetics*

OPEN ACCESS

Streptococcus pneumoniae and Influenza (H1N1) Virus Genome Study against Human Genome Sequences-BLASTn and DAVID Analysis

Mark Christopher Arokiaraj[1]*, Jarad Wilson[2], Jingqiao Lu[2], RuoPan Huang[2]

[1]*Cardiology Consultant and Researcher, Pondicherry, India;* [2]*RayBiotech, Peachtree Corners, Georgia, USA*

Edited by: Slavica Hristomanova-Mitkovska
Citation: Arokiaraj MC, Wilson J, Lu J, Huang RP. Streptococcus Pneumoniae and Influenza (H1N1) Virus Genome Study Against Human Genome Sequences – BLASTn and DAVID Analysis. SEE J Immunol. 2024 May 30;7:67-80. https://doi.org/10.3889/seejim.2024.6064
Keywords: *Streptococcus pneumoniae*; Influenza (H1N1) virus; Computational biology; BLASTn; DAVID; NDEx
***Correspondence:** Mark Christopher Arokiaraj, Cardiology Consultant and Researcher, Pondicherry-605001, India.
E-mail: christomark@gmail.com
Received: 21-Dec-2023
Revised: 29-Mar-2024
Accepted: 23-Apr-2024
Published: 30-May-2024
Copyright: © 2024 Mark Christopher Arokiaraj, Jarad Wilson, Jingqiao Lu, RuoPan Huang
Funding: This research did not receive any financial support
Competing Interests: The authors have declared that no competing interests exist
Open Access: This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

Abstract

AIMS: *Streptococcus pneumoniae* and influenza H1N1 virus are common organisms associated with human infections. These infections could play a significant role in immune regulation. The study was performed to analyze the genome sequences of these organisms with human genome and study its functional significance.

MATERIALS AND METHODS: The study was performed to analyze the overlapping of genome sequences in *S. pneumoniae* and influenza (H1N1) virus against human genome sequences by BLASTn sequence analysis. The alignments are studied against the corresponding genes for their functional significance with DAVID and NDEx software.

RESULTS: Several hits or overlapping nucleotide segments were identified. Between streptococcus and Homo Sapiens 287 overlaps were identified, and among influenza and Homo sapiens 124 hits were identified. A wide range of functional significance of these genes were identified, and the results are presented in this study. The results show insights into functional pathways and biological activities associated with the respective vaccinations or infections by these microorganisms.

CONCLUSION: The common organisms such as *S. pneumoniae* and influenza H1N1 virus actively interact with the immune system and result in a wide range of immune regulations.

Introduction

Streptococcus pneumoniae and influenza (H1N1) are common vaccinations given to the elderly general population. Infections by these organisms are associated with mortality and morbidity worldwide, and the regional incidence varies [1]. Vaccinations against these organisms are associated with cardiovascular benefits and also a general reduction in mortality [2], [3], [4]. Improvement in trained immunity has been postulated for non-specific protection of these vaccinations against COVID-19 [5]. The study was performed to analyze the genome sequence overlap between these micro-organisms and the human genome and to evaluate the possible modifications and influences of these vaccinations on human gene expression.

Materials and Methods

BLASTn analysis [6], [7] was performed by comparing the *S. pneumoniae* strain Hu17 and Haemophilus influenza (H1N1) 2009 California genome sequences against the human genome (Homo sapiens Hg38). The genome sequence of Influenza H1N1 2009 California (2009 California) was downloaded from National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/641809/). The Influenza H1N1 2009 California genome sequences were compared against Homo sapiens hg38 downloaded from NCBI (https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.26/) with BLASTn tools (2.13.0+) on a Linux server, respectively. The genome sequence of Streptococcus Pneumonia strain Hu17 (2009 California) was downloaded from NCBI (https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/641809/). The *S. pneumoniae* strain Hu17 genome sequences were compared against Homo sapiens hg38 downloaded from NCBI (https://www.ncbi.nlm.nih.gov/assembly/GCF_00000 1405.26/) with BLASTn tools (2.13.0+) on a Linux server, respectively. The hits/alignments on the genome of Homo sapiens hg38 reported by the blast analysis were retrieved with customized Python scripts. The regions of hits were annotated with gene information, respectively.

The corresponding genes were studied for functional annotations to study background processes,

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	BIOGRID_INTERACTION	vir like m6A methyltransferase associated(VIRMA)	RT		33	26.8	9.7E-6	3.7E-2
	BIOGRID_INTERACTION	formyl peptide receptor 1(FPR1)	RT		5	4.1	5.9E-3	1.0E0
	BIOGRID_INTERACTION	cystatin S(CST4)	RT		4	3.3	9.9E-3	1.0E0
	BIOGRID_INTERACTION	amylase alpha 1C(AMY1C)	RT		4	3.3	9.9E-3	1.0E0
	BIOGRID_INTERACTION	sodium voltage-gated channel beta subunit 2(SCN2B)	RT		4	3.3	1.0E-2	1.0E0
	BIOGRID_INTERACTION	protein phosphatase 3 regulatory subunit B, beta(PPP3R2)	RT		3	2.4	1.1E-2	1.0E0
	BIOGRID_INTERACTION	KH RNA binding domain containing, signal transduction associated 1(KHDRBS1)	RT		6	4.9	1.1E-2	1.0E0
	BIOGRID_INTERACTION	Fas ligand(FASLG)	RT		4	3.3	1.1E-2	1.0E0
	BIOGRID_INTERACTION	capicua transcriptional repressor(CIC)	RT		9	7.3	1.3E-2	1.0E0
	BIOGRID_INTERACTION	trafficking protein particle complex subunit 2B(TRAPPC2B)	RT		3	2.4	1.3E-2	1.0E0
	BIOGRID_INTERACTION	G protein-coupled receptor 17(GPR17)	RT		6	4.9	1.4E-2	1.0E0
	BIOGRID_INTERACTION	solute carrier family 18 member A2(SLC18A2)	RT		4	3.3	1.5E-2	1.0E0
	BIOGRID_INTERACTION	heterogeneous nuclear ribonucleoprotein L(HNRNPL)	RT		17	13.8	1.9E-2	1.0E0
	BIOGRID_INTERACTION	hook microtubule tethering protein 1(HOOK1)	RT		5	4.1	2.3E-2	1.0E0
	BIOGRID_INTERACTION	Rac family small GTPase 1(RAC1)	RT		6	4.9	2.5E-2	1.0E0
	BIOGRID_INTERACTION	SLIT-ROBO Rho GTPase activating protein 2(SRGAP2)	RT		4	3.3	2.5E-2	1.0E0
	BIOGRID_INTERACTION	dynein cytoplasmic 1 light intermediate chain 1(Dync1li1)	RT		3	2.4	2.6E-2	1.0E0
	BIOGRID_INTERACTION	par-3 family cell polarity regulator(PARD3)	RT		5	4.1	3.1E-2	1.0E0
	BIOGRID_INTERACTION	trafficking protein particle complex subunit 8(TRAPPC8)	RT		3	2.4	3.2E-2	1.0E0
	BIOGRID_INTERACTION	proteasome activator subunit 4(PSME4)	RT		3	2.4	3.3E-2	1.0E0
	BIOGRID_INTERACTION	ubiquitin specific peptidase 7(USP7)	RT		7	5.7	3.5E-2	1.0E0
	BIOGRID_INTERACTION	protein inhibitor of activated STAT 1(PIAS1)	RT		4	3.3	3.5E-2	1.0E0
	BIOGRID_INTERACTION	DDB1 and CUL4 associated factor 15(DCAF15)	RT		5	4.1	3.9E-2	1.0E0
	BIOGRID_INTERACTION	lymphocyte transmembrane adaptor 1(LAX1)	RT		2	1.6	4.0E-2	1.0E0
	BIOGRID_INTERACTION	ankyrin repeat domain 28(Ankrd28)	RT		2	1.6	4.0E-2	1.0E0
	BIOGRID_INTERACTION	gap junction protein alpha 1(GJA1)	RT		7	5.7	4.0E-2	1.0E0
	BIOGRID_INTERACTION	butyrophilin like 9(BTNL9)	RT		4	3.3	4.2E-2	1.0E0
	BIOGRID_INTERACTION	RAB4A, member RAS oncogene family(RAB4A)	RT		6	4.9	4.3E-2	1.0E0
	BIOGRID_INTERACTION	protein kinase cAMP-activated catalytic subunit alpha(PRKACA)	RT		5	4.1	4.3E-2	1.0E0
	BIOGRID_INTERACTION	RAB2B, member RAS oncogene family(RAB2B)	RT		3	2.4	4.3E-2	1.0E0
	BIOGRID_INTERACTION	RAB11A, member RAS oncogene family(RAB11A)	RT		8	6.5	4.8E-2	1.0E0
	BIOGRID_INTERACTION	late endosomal/lysosomal adaptor, MAPK and MTOR activator 1(LAMTOR1)	RT		8	6.5	5.3E-2	1.0E0
	BIOGRID_INTERACTION	peroxisome proliferator activated receptor alpha(PPARA)	RT		3	2.4	5.4E-2	1.0E0
	BIOGRID_INTERACTION	WD repeat domain, phosphoinositide interacting 2(Wipi2)	RT		2	1.6	5.4E-2	1.0E0
	BIOGRID_INTERACTION	chromosome 17 open reading frame 50(C17orf50)	RT		2	1.6	5.4E-2	1.0E0
	BIOGRID_INTERACTION	pleckstrin homology and FYVE domain containing 2(PLEKHF2)	RT		4	3.3	5.8E-2	1.0E0
	BIOGRID_INTERACTION	trafficking protein particle complex subunit 3L(TRAPPC3L)	RT		2	1.6	5.9E-2	1.0E0
	BIOGRID_INTERACTION	estrogen receptor binding site associated antigen 9(EBAG9)	RT		5	4.1	6.0E-2	1.0E0
	BIOGRID_INTERACTION	DDB1 and CUL4 associated factor 6(DCAF6)	RT		3	2.4	6.1E-2	1.0E0
	BIOGRID_INTERACTION	protein phosphatase 1 catalytic subunit alpha(PPP1CA)	RT		3	2.4	6.2E-2	1.0E0
	BIOGRID_INTERACTION	TNF receptor associated factor 1(TRAF1)	RT		5	4.1	6.3E-2	1.0E0
	BIOGRID_INTERACTION	lysosomal associated membrane protein 2(LAMP2)	RT		7	5.7	6.5E-2	1.0E0
	BIOGRID_INTERACTION	G protein-coupled receptor 45(GPR45)	RT		5	4.1	6.6E-2	1.0E0
	BIOGRID_INTERACTION	kinesin family member 14(KIF14)	RT		15	12.2	6.6E-2	1.0E0
	BIOGRID_INTERACTION	RAB9A, member RAS oncogene family(RAB9A)	RT		7	5.7	6.7E-2	1.0E0
	BIOGRID_INTERACTION	LSM7 homolog, U6 small nuclear RNA and mRNA degradation associated(LSM7)	RT		3	2.4	6.9E-2	1.0E0
	BIOGRID_INTERACTION	Cbl proto-oncogene B(CBLB)	RT		3	2.4	7.2E-2	1.0E0
	BIOGRID_INTERACTION	chymotrypsin like elastase 2A(CELA2A)	RT		2	1.6	7.3E-2	1.0E0
	BIOGRID_INTERACTION	thrombospondin 1(THBS1)	RT		3	2.4	7.8E-2	1.0E0
	BIOGRID_INTERACTION	secreted phosphoprotein 1(SPP1)	RT		3	2.4	8.0E-2	1.0E0
	BIOGRID_INTERACTION	lysine demethylase 6B(KDM6B)	RT		3	2.4	8.1E-2	1.0E0
	BIOGRID_INTERACTION	ST14 transmembrane serine protease matriptase(ST14)	RT		4	3.3	8.2E-2	1.0E0
	BIOGRID_INTERACTION	tripartite motif containing 66(TRIM66)	RT		4	3.3	8.2E-2	1.0E0
	BIOGRID_INTERACTION	CD2 molecule(CD2)	RT		2	1.6	8.3E-2	1.0E0
	BIOGRID_INTERACTION	WW domain containing oxidoreductase(WWOX)	RT		7	5.7	8.5E-2	1.0E0
	BIOGRID_INTERACTION	ezrin(EZR)	RT		6	4.9	8.5E-2	1.0E0
	BIOGRID_INTERACTION	GABA type A receptor-associated protein(GABARAP)	RT		3	2.4	8.7E-2	1.0E0
	BIOGRID_INTERACTION	chemerin chemokine-like receptor 1(CMKLR1)	RT		3	2.4	8.9E-2	1.0E0
	BIOGRID_INTERACTION	syntaxin 7(STX7)	RT		7	5.7	9.0E-2	1.0E0
	BIOGRID_INTERACTION	calcium voltage-gated channel subunit alpha1 A(CACNA1A)	RT		3	2.4	9.0E-2	1.0E0
	BIOGRID_INTERACTION	FHF complex subunit HOOK interacting protein 1B(FHIP1B)	RT		2	1.6	9.2E-2	1.0E0
	BIOGRID_INTERACTION	cell division cycle associated 5(Cdca5)	RT		2	1.6	9.2E-2	1.0E0
	BIOGRID_INTERACTION	cadherin 3(CDH3)	RT		2	1.6	9.2E-2	1.0E0
	BIOGRID_INTERACTION	tripartite motif containing 21(TRIM21)	RT		4	3.3	9.4E-2	1.0E0
	BIOGRID_INTERACTION	nucleoporin 214(NUP214)	RT		3	2.4	9.6E-2	1.0E0
	BIOGRID_INTERACTION	SEC62 homolog, preprotein translocation factor(SEC62)	RT		6	4.9	9.7E-2	1.0E0
	BIOGRID_INTERACTION	trafficking protein particle complex 8(Trappc8)	RT		2	1.6	9.7E-2	1.0E0
	BIOGRID_INTERACTION	ERBB receptor feedback inhibitor 1(ERRFI1)	RT		3	2.4	9.8E-2	1.0E0

Figure 1: Influenza H1N1 virus versus Homo sapiens – functional annotation chart with biogrid interactions by DAVID analysis

which could influence physiological changes. The blast analysis results were further processed for functional analysis of the concerned genes through the database for annotation, visualization, and integrated discovery

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	REACTOME_PATHWAY	FCERI mediated Ca+2 mobilization	RT		3	2.4	1.1E-2	1.0E0
	REACTOME_PATHWAY	RHOF GTPase cycle	RT		3	2.4	1.8E-2	1.0E0
	REACTOME_PATHWAY	Fc epsilon receptor (FCERI) signaling	RT		4	3.3	2.8E-2	1.0E0
	REACTOME_PATHWAY	RHOD GTPase cycle	RT		3	2.4	2.9E-2	1.0E0
	REACTOME_PATHWAY	Signal Transduction	RT		20	16.3	3.7E-2	1.0E0
	REACTOME_PATHWAY	RHOG GTPase cycle	RT		3	2.4	5.1E-2	1.0E0
	REACTOME_PATHWAY	RAC1 GTPase cycle	RT		4	3.3	6.3E-2	1.0E0
	REACTOME_PATHWAY	RHO GTPase cycle	RT		6	4.9	6.9E-2	1.0E0
	REACTOME_PATHWAY	RAC2 GTPase cycle	RT		3	2.4	7.0E-2	1.0E0
	REACTOME_PATHWAY	MAPK6/MAPK4 signaling	RT		3	2.4	7.1E-2	1.0E0
	REACTOME_PATHWAY	Post-translational protein modification	RT		12	9.8	9.0E-2	1.0E0

Figure 2: Influenza H1N1 virus versus Homo sapiens – genes in pathway map

(DAVID) [8] to understand the biological implications of the outcomes. The functional description of the genes was obtained from NCBI website.

Results

Based on the BLASTn analysis, 124 hits or overlapping sequences were observed in the analysis of H1N1 versus Homo sapiens. Two hundred and eighty-seven hits were observed between the overlapping segments of *S. pneumoniae* and Homo sapiens. The results are published in the data repository (Supplement Data 1 and 2). DAVID analysis showed various gene enrichments related to functions with statistical significance. The results are published in the supplement data repository (Supplement Data 3). The enrichment values and the level of significance of the genes are also shown in the results. A wide range of enrichment and functional annotations are observed, which give insights into the background processes after the vaccinations or infections with these organisms. The functional annotation chart with biogrid interactions (Figure 1), genes and pathway map (Figure 2), protein domain (Figure 3), and tissue characterization (Figure 4) derived from DAVID analysis between influenza H1N1 and Homo sapiens indicate the various functional associations with associated genes. Similarly, the results of DAVID analysis with *S. pneumoniae* and Homo sapiens show the associated protein domain (Figure 5), pathways (Figure 6), genes associated with tissue expression (Figure 7), and biogrid interactions (Supplement Figure 1). Influenza genome versus human genome showed association with some of the commonly associated genes

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	INTERPRO	Src homology-3 domain	RT		6	4.9	5.9E-3	8.7E-1
	INTERPRO	Pleckstrin homology-like domain	RT		8	6.5	7.7E-3	8.7E-1
	INTERPRO	Armadillo-type fold	RT		7	5.7	9.5E-3	8.7E-1
	INTERPRO	ATPase, AAA-type, core	RT		3	2.4	2.9E-2	1.0E0
	INTERPRO	RNA recognition motif domain	RT		5	4.1	3.1E-2	1.0E0
	INTERPRO	AAA+ ATPase domain	RT		4	3.3	3.4E-2	1.0E0
	INTERPRO	Nucleotide-binding, alpha-beta plait	RT		5	4.1	4.9E-2	1.0E0
	INTERPRO	Diacylglycerol kinase, accessory domain	RT		2	1.6	4.9E-2	1.0E0
	INTERPRO	Pleckstrin homology domain	RT		5	4.1	5.0E-2	1.0E0
	INTERPRO	MAD homology 1, Dwarfin-type	RT		2	1.6	5.9E-2	1.0E0
	INTERPRO	Diacylglycerol kinase, catalytic domain	RT		2	1.6	7.3E-2	1.0E0
	INTERPRO	ATP-NAD kinase-like domain	RT		2	1.6	8.2E-2	1.0E0
	INTERPRO	Inorganic polyphosphate/ATP-NAD kinase, domain 1	RT		2	1.6	8.2E-2	1.0E0
	INTERPRO	P-loop containing nucleoside triphosphate hydrolase	RT		9	7.3	9.0E-2	1.0E0

Figure 3: Influenza H1N1 virus versus Homo sapiens – functional protein domains

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	UP_TISSUE	Cervix carcinoma	RT		40	32.5	4.4E-3	3.9E-1
	UP_TISSUE	Brain	RT		56	45.5	8.4E-3	3.9E-1
	UP_TISSUE	Cerebellum	RT		12	9.8	2.1E-2	6.1E-1
	UP_TISSUE	Leukemic T-cell	RT		21	17.1	2.6E-2	6.1E-1
	UP_TISSUE	Amygdala	RT		10	8.1	3.3E-2	6.2E-1
	UP_TISSUE	Platelet	RT		7	5.7	6.0E-2	9.3E-1
	UP_TISSUE	Spinal cord	RT		3	2.4	7.8E-2	1.0E0

Figure 4: Influenza H1N1 virus versus Homo sapiens – tissue expressions

– KAT6A, KRAS, EGFR, TPJ3, DAZ1/4, and FOFR4. Some of the pathways associated with influenza genome analysis were NFKB1, mTOR, EGR1, SMAD4, TBL1X, and YY1.

Numerically, *S. pneumoniae* has more functional interactions than influenza H1N1 virus, though the extent of biological outcomes is difficult to quantify. NDEx analysis of the overlapping genes showed the pathways associated with the genes and gene ontology analysis (Figure 8 and 9) showed the relevant associated genes (Supplement Figure 2 and 3). Heat shock-related pathways and signaling was predominantly involved with S. pneumoniae-homo sapiens overlapping genes (Figure 9). Genes related to histone modifications, myc pathways, DNA damage check point signaling, FGF/IGF, and notch signaling were associated with influenza – Homo sapiens overlap related genes (Supplement Figure 2). PRKN, SNCA, STUB1, GPR37, SEPT, TUBA, and TUBB genes related pathways concerning proteasomal degradation were associated with S. pneumonia-Homo sapiens overlapping genes (Supplement Figure 3). Some of the other seen by pathway figures similarity were KRAS, EGFR and by INDRO GO similarity ELAVL1 and MAPKAPK2 were some of the pathways identified (Supplement Figure 4) when analyzed between influenza H1N1 versus Homo sapiens. Genes such as HSF1, GCase, and HSPA9 were some of the pathways identified when analyzed between *S. pneumoniae* versus Homo sapiens (Supplement Figure 5). Table 1 gives the summary of the genes and pathways associated with influenza virus and *S. pneumoniae* genome.

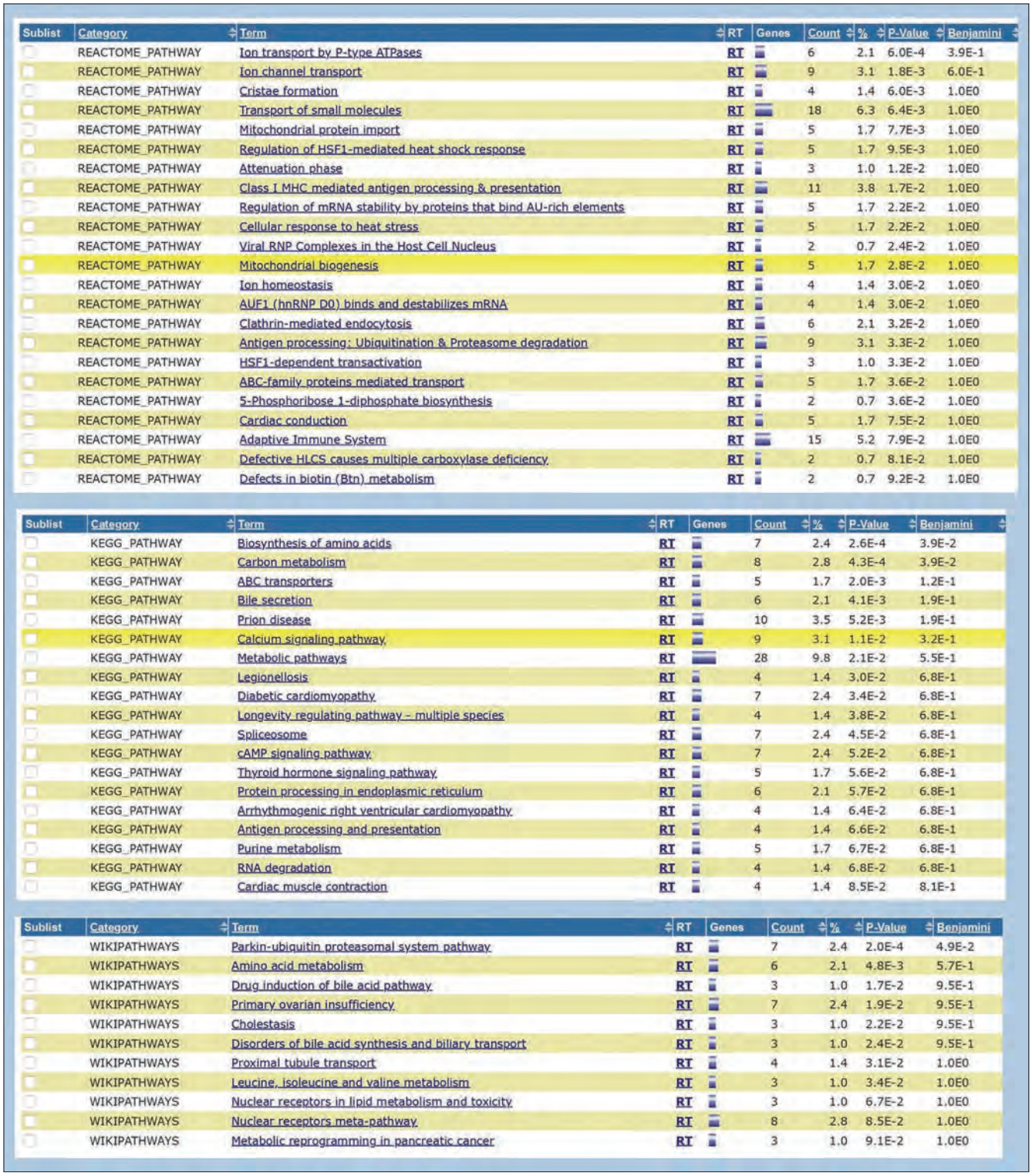
Discussion

In the comparison of the nucleotide sequences, there were many overlapping significant alignments

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	INTERPRO	Peptidase M1, alanine aminopeptidase/leukotriene A4 hydrolase	RT		5	1.7	5.7E-6	1.8E-3
	INTERPRO	Peptidase M1, membrane alanine aminopeptidase, N-terminal	RT		5	1.7	8.1E-6	1.8E-3
	INTERPRO	Heat shock protein 70, conserved site	RT		5	1.7	1.5E-5	1.8E-3
	INTERPRO	Heat shock protein 70 family	RT		5	1.7	1.5E-5	1.8E-3
	INTERPRO	AAA+ ATPase domain	RT		10	3.5	2.0E-5	1.8E-3
	INTERPRO	Cation-transporting P-type ATPase, C-terminal	RT		5	1.7	3.3E-5	1.8E-3
	INTERPRO	Cation-transporting P-type ATPase, N-terminal	RT		5	1.7	4.2E-5	1.8E-3
	INTERPRO	P-type ATPase, phosphorylation site	RT		6	2.1	4.2E-5	1.8E-3
	INTERPRO	P-type ATPase, A domain	RT		6	2.1	4.2E-5	1.8E-3
	INTERPRO	Cation-transporting P-type ATPase	RT		6	2.1	4.2E-5	1.8E-3
	INTERPRO	P-type ATPase, cytoplasmic domain N	RT		6	2.1	4.2E-5	1.8E-3
	INTERPRO	P-loop containing nucleoside triphosphate hydrolase	RT		25	8.7	4.3E-5	1.8E-3
	INTERPRO	P-type ATPase, transmembrane domain	RT		6	2.1	4.8E-5	1.9E-3
	INTERPRO	Domain of unknown function DUF3358	RT		4	1.4	6.3E-5	2.2E-3
	INTERPRO	Carbamoyl-phosphate synthetase large subunit-like, ATP-binding domain	RT		4	1.4	6.3E-5	2.2E-3
	INTERPRO	ABC transporter, conserved site	RT		6	2.1	1.2E-4	3.9E-3
	INTERPRO	ABC transporter-like	RT		6	2.1	2.0E-4	6.1E-3
	INTERPRO	ABC transporter, transmembrane domain, type 1	RT		5	1.7	2.4E-4	6.7E-3
	INTERPRO	Pre-ATP-grasp domain	RT		4	1.4	3.1E-4	8.4E-3
	INTERPRO	ATP-grasp fold	RT		4	1.4	3.9E-4	1.0E-2
	INTERPRO	ATP-grasp fold, subdomain 1	RT		4	1.4	7.2E-4	1.7E-2
	INTERPRO	ATPase, alpha/beta subunit, N-terminal	RT		3	1.0	1.1E-3	2.3E-2
	INTERPRO	ATPase, alpha/beta subunit, nucleotide-binding domain, active site	RT		3	1.0	1.1E-3	2.3E-2
	INTERPRO	ATPase, F1/V1/A1 complex, alpha/beta subunit, nucleotide-binding domain	RT		3	1.0	1.1E-3	2.3E-2
	INTERPRO	HAD-like domain	RT		6	2.1	1.5E-3	3.1E-2
	INTERPRO	Cyclophilin-like peptidyl-prolyl cis-trans isomerase domain	RT		4	1.4	2.0E-3	4.0E-2
	INTERPRO	Galactose-binding domain-like	RT		6	2.1	2.2E-3	4.2E-2
	INTERPRO	Chromo domain-like	RT		4	1.4	5.1E-3	9.3E-2
	INTERPRO	SNF2-related	RT		4	1.4	5.5E-3	9.4E-2
	INTERPRO	Chromo domain/shadow	RT		4	1.4	5.5E-3	9.4E-2
	INTERPRO	Helicase, C-terminal	RT		6	2.1	6.6E-3	1.1E-1
	INTERPRO	Helicase, superfamily 1/2, ATP-binding domain	RT		6	2.1	7.6E-3	1.2E-1
	INTERPRO	Organic anion transporter polypeptide OATP	RT		3	1.0	1.5E-2	2.3E-1
	INTERPRO	Kazal domain	RT		4	1.4	1.8E-2	2.3E-1
	INTERPRO	Cyclophilin-type peptidyl-prolyl cis-trans isomerase, conserved site	RT		3	1.0	2.1E-2	2.3E-1
	INTERPRO	S-adenosylmethionine synthetase, central domain	RT		2	0.7	2.1E-2	2.3E-1
	INTERPRO	S-adenosylmethionine synthetase superfamily	RT		2	0.7	2.1E-2	2.3E-1
	INTERPRO	S-adenosylmethionine synthetase, C-terminal	RT		2	0.7	2.1E-2	2.3E-1
	INTERPRO	S-adenosylmethionine synthetase, N-terminal	RT		2	0.7	2.1E-2	2.3E-1
	INTERPRO	S-adenosylmethionine synthetase, conserved site	RT		2	0.7	2.1E-2	2.3E-1
	INTERPRO	Phosphoglycerate kinase, conserved site	RT		2	0.7	2.1E-2	2.3E-1
	INTERPRO	Isy1-like splicing	RT		2	0.7	2.1E-2	2.3E-1
	INTERPRO	ATPase, F1 complex beta subunit/V1 complex, C-terminal	RT		2	0.7	2.1E-2	2.3E-1
	INTERPRO	Phosphoglycerate kinase	RT		2	0.7	2.1E-2	2.3E-1
	INTERPRO	S-adenosylmethionine synthetase	RT		2	0.7	2.1E-2	2.3E-1
	INTERPRO	ATPase, F1/A1 complex, alpha subunit, N-terminal	RT		2	0.7	2.1E-2	2.3E-1
	INTERPRO	Phosphoglycerate kinase, N-terminal	RT		2	0.7	2.1E-2	2.3E-1
	INTERPRO	Laminin G domain	RT		4	1.4	2.6E-2	2.8E-1
	INTERPRO	Carbamoyl-phosphate synthase large subunit, CPSase domain	RT		2	0.7	3.1E-2	2.9E-1
	INTERPRO	Carbamoyl-phosphate synthetase, large subunit oligomerisation domain	RT		2	0.7	3.1E-2	2.9E-1
	INTERPRO	Calcium-transporting P-type ATPase, subfamily IIA, SERCA-type	RT		2	0.7	3.1E-2	2.9E-1
	INTERPRO	Phosphoribosyl pyrophosphate synthetase, conserved site	RT		2	0.7	3.1E-2	2.9E-1
	INTERPRO	Carbamoyl-phosphate synthase, small subunit N-terminal domain	RT		2	0.7	3.1E-2	2.9E-1
	INTERPRO	Carbamoyl-phosphate synthase, small subunit	RT		2	0.7	3.1E-2	2.9E-1
	INTERPRO	Carbamoyl-phosphate synthase, large subunit	RT		2	0.7	3.1E-2	2.9E-1
	INTERPRO	Chromo domain	RT		3	1.0	3.3E-2	3.0E-1
	INTERPRO	Cadherin	RT		5	1.7	3.7E-2	3.3E-1
	INTERPRO	Armadillo-type fold	RT		9	3.1	3.9E-2	3.4E-1
	INTERPRO	Cadherin-like	RT		5	1.7	4.0E-2	3.4E-1
	INTERPRO	Biotin-binding site	RT		2	0.7	4.2E-2	3.4E-1
	INTERPRO	Methylglyoxal synthase-like domain	RT		2	0.7	4.2E-2	3.4E-1
	INTERPRO	Beta-galactosidase 1-like	RT		2	0.7	4.2E-2	3.4E-1
	INTERPRO	Glycoside hydrolase, family 35	RT		2	0.7	4.2E-2	3.4E-1
	INTERPRO	Ribose-phosphate diphosphokinase	RT		2	0.7	5.2E-2	3.9E-1
	INTERPRO	Biotin carboxylation domain	RT		2	0.7	5.2E-2	3.9E-1
	INTERPRO	Pyridine nucleotide-disulphide oxidoreductase, dimerisation	RT		2	0.7	5.2E-2	3.9E-1
	INTERPRO	Pyridine nucleotide-disulphide oxidoreductase, class I, active site	RT		2	0.7	5.2E-2	3.9E-1
	INTERPRO	Biotin carboxylase, C-terminal	RT		2	0.7	5.2E-2	3.9E-1
	INTERPRO	Carbamoyl-phosphate synthase, large subunit, N-terminal	RT		2	0.7	5.2E-2	3.9E-1
	INTERPRO	Rudiment single hybrid motif	RT		2	0.7	6.2E-2	4.5E-1
	INTERPRO	Glutamine amidotransferase	RT		2	0.7	6.2E-2	4.5E-1
	INTERPRO	Sodium/potassium-transporting P-type ATPase, subfamily IIC	RT		2	0.7	7.2E-2	5.1E-1
	INTERPRO	Pleckstrin homology domain	RT		7	2.4	7.5E-2	5.2E-1
	INTERPRO	FAD/NAD-linked reductase, dimerisation	RT		2	0.7	8.2E-2	5.7E-1
	INTERPRO	Kelch-like protein, gigaxonin	RT		3	1.0	8.6E-2	5.8E-1
	INTERPRO	Protein kinase, catalytic domain	RT		10	3.5	8.7E-2	5.9E-1
	INTERPRO	BEN domain	RT		2	0.7	9.2E-2	6.1E-1

Figure 5: Streptococcus pneumoniae versus Homo sapiens – functional protein domain

between the influenza (H1N1) and Streptococcus pneumonia and the human genome. This could be an evolutionary or a survival process where the organisms function and induce immunological processes by molecular



Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	REACTOME_PATHWAY	Ion transport by P-type ATPases	RT		6	2.1	6.0E-4	3.9E-1
	REACTOME_PATHWAY	Ion channel transport	RT		9	3.1	1.8E-3	6.0E-1
	REACTOME_PATHWAY	Cristae formation	RT		4	1.4	6.0E-3	1.0E0
	REACTOME_PATHWAY	Transport of small molecules	RT		18	6.3	6.4E-3	1.0E0
	REACTOME_PATHWAY	Mitochondrial protein import	RT		5	1.7	7.7E-3	1.0E0
	REACTOME_PATHWAY	Regulation of HSF1-mediated heat shock response	RT		5	1.7	9.5E-3	1.0E0
	REACTOME_PATHWAY	Attenuation phase	RT		3	1.0	1.2E-2	1.0E0
	REACTOME_PATHWAY	Class I MHC mediated antigen processing & presentation	RT		11	3.8	1.7E-2	1.0E0
	REACTOME_PATHWAY	Regulation of mRNA stability by proteins that bind AU-rich elements	RT		5	1.7	2.2E-2	1.0E0
	REACTOME_PATHWAY	Cellular response to heat stress	RT		5	1.7	2.2E-2	1.0E0
	REACTOME_PATHWAY	Viral RNP Complexes in the Host Cell Nucleus	RT		2	0.7	2.4E-2	1.0E0
	REACTOME_PATHWAY	Mitochondrial biogenesis	RT		5	1.7	2.8E-2	1.0E0
	REACTOME_PATHWAY	Ion homeostasis	RT		4	1.4	3.0E-2	1.0E0
	REACTOME_PATHWAY	AUF1 (hnRNP D0) binds and destabilizes mRNA	RT		4	1.4	3.0E-2	1.0E0
	REACTOME_PATHWAY	Clathrin-mediated endocytosis	RT		6	2.1	3.2E-2	1.0E0
	REACTOME_PATHWAY	Antigen processing: Ubiquitination & Proteasome degradation	RT		9	3.1	3.3E-2	1.0E0
	REACTOME_PATHWAY	HSF1-dependent transactivation	RT		3	1.0	3.3E-2	1.0E0
	REACTOME_PATHWAY	ABC-family proteins mediated transport	RT		5	1.7	3.6E-2	1.0E0
	REACTOME_PATHWAY	5-Phosphoribose 1-diphosphate biosynthesis	RT		2	0.7	3.6E-2	1.0E0
	REACTOME_PATHWAY	Cardiac conduction	RT		5	1.7	7.5E-2	1.0E0
	REACTOME_PATHWAY	Adaptive Immune System	RT		15	5.2	7.9E-2	1.0E0
	REACTOME_PATHWAY	Defective HLCS causes multiple carboxylase deficiency	RT		2	0.7	8.1E-2	1.0E0
	REACTOME_PATHWAY	Defects in biotin (Btn) metabolism	RT		2	0.7	9.2E-2	1.0E0

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	KEGG_PATHWAY	Biosynthesis of amino acids	RT		7	2.4	2.6E-4	3.9E-2
	KEGG_PATHWAY	Carbon metabolism	RT		8	2.8	4.3E-4	3.9E-2
	KEGG_PATHWAY	ABC transporters	RT		5	1.7	2.0E-3	1.2E-1
	KEGG_PATHWAY	Bile secretion	RT		6	2.1	4.1E-3	1.9E-1
	KEGG_PATHWAY	Prion disease	RT		10	3.5	5.2E-3	1.9E-1
	KEGG_PATHWAY	Calcium signaling pathway	RT		9	3.1	1.1E-2	3.2E-1
	KEGG_PATHWAY	Metabolic pathways	RT		28	9.8	2.1E-2	5.5E-1
	KEGG_PATHWAY	Legionellosis	RT		4	1.4	3.0E-2	6.8E-1
	KEGG_PATHWAY	Diabetic cardiomyopathy	RT		7	2.4	3.4E-2	6.8E-1
	KEGG_PATHWAY	Longevity regulating pathway - multiple species	RT		4	1.4	3.8E-2	6.8E-1
	KEGG_PATHWAY	Spliceosome	RT		7	2.4	4.5E-2	6.8E-1
	KEGG_PATHWAY	cAMP signaling pathway	RT		7	2.4	5.2E-2	6.8E-1
	KEGG_PATHWAY	Thyroid hormone signaling pathway	RT		5	1.7	5.6E-2	6.8E-1
	KEGG_PATHWAY	Protein processing in endoplasmic reticulum	RT		6	2.1	5.7E-2	6.8E-1
	KEGG_PATHWAY	Arrhythmogenic right ventricular cardiomyopathy	RT		4	1.4	6.4E-2	6.8E-1
	KEGG_PATHWAY	Antigen processing and presentation	RT		4	1.4	6.6E-2	6.8E-1
	KEGG_PATHWAY	Purine metabolism	RT		5	1.7	6.7E-2	6.8E-1
	KEGG_PATHWAY	RNA degradation	RT		4	1.4	6.8E-2	6.8E-1
	KEGG_PATHWAY	Cardiac muscle contraction	RT		4	1.4	8.5E-2	8.1E-1

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	WIKIPATHWAYS	Parkin-ubiquitin proteasomal system pathway	RT		7	2.4	2.0E-4	4.9E-2
	WIKIPATHWAYS	Amino acid metabolism	RT		6	2.1	4.8E-3	5.7E-1
	WIKIPATHWAYS	Drug induction of bile acid pathway	RT		3	1.0	1.7E-2	9.5E-1
	WIKIPATHWAYS	Primary ovarian insufficiency	RT		7	2.4	1.9E-2	9.5E-1
	WIKIPATHWAYS	Cholestasis	RT		3	1.0	2.2E-2	9.5E-1
	WIKIPATHWAYS	Disorders of bile acid synthesis and biliary transport	RT		3	1.0	2.4E-2	9.5E-1
	WIKIPATHWAYS	Proximal tubule transport	RT		4	1.4	3.1E-2	1.0E0
	WIKIPATHWAYS	Leucine, isoleucine and valine metabolism	RT		3	1.0	3.4E-2	1.0E0
	WIKIPATHWAYS	Nuclear receptors in lipid metabolism and toxicity	RT		3	1.0	6.7E-2	1.0E0
	WIKIPATHWAYS	Nuclear receptors meta-pathway	RT		8	2.8	8.5E-2	1.0E0
	WIKIPATHWAYS	Metabolic reprogramming in pancreatic cancer	RT		3	1.0	9.1E-2	1.0E0

Figure 6: Streptococcus pneumoniae versus Homo sapiens – genes in pathway map

mimicry. This can lead to the induction of cellular processes – cytoplasmic, nuclear, and various levels as seen through functional analysis with DAVID. These vaccinations are associated with various benefits, as reported by previous studies. These vaccinations are protective against diseases like COVID-19 through changes in adaptive immunity and are also reported to have cardioprotective effects, though the exact mechanisms are not understood [9]. The countries which had high influenza or lower respiratory tract infections burden illnesses and countries which had higher influenza vaccinations, like South Korea, had significantly lower COVID-19 mortality during the pandemic [10]. Various pathways and genes are associated with the overlapping gene. Predominantly, these are inflammation, death signaling and pathways related to epigenetic modifications. Heat shock proteins are involved in normal and other cellular activities

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	HPA_NORMAL_TISSUE_CELLTYPE	bronchus; ciliated cells (ciliary rootlets)	RT		14	4.9	1.4E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	liver; hepatocytes	RT		86	30.1	1.5E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	soft tissue 1; chondrocytes	RT		39	13.6	1.8E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	liver; cholangiocytes	RT		63	22.0	1.8E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	hippocampus; glial cells	RT		65	22.7	2.1E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	hippocampus; neuronal cells	RT		91	31.8	2.2E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	nasopharynx; ciliated cells (ciliary rootlets)	RT		11	3.8	2.6E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	duodenum; goblet cells	RT		11	3.8	3.1E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	bronchus; ciliated cells (cell body)	RT		23	8.0	3.1E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	cerebral cortex; neuropil	RT		75	26.2	3.2E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	endometrium 1; glandular cells	RT		97	33.9	3.6E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	rectum; endothelial cells	RT		8	2.8	4.1E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	rectum; peripheral nerve/ganglion	RT		8	2.8	4.6E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	cerebellum; cells in molecular layer	RT		67	23.4	4.6E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	thyroid gland; glandular cells	RT		100	35.0	4.7E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	cerebral cortex; neuronal cells	RT		95	33.2	4.7E-2	6.0E-1
	HPA_NORMAL_TISSUE_CELLTYPE	skin 1; melanocytes	RT		77	26.9	5.9E-2	6.9E-1
	HPA_NORMAL_TISSUE_CELLTYPE	nasopharynx; ciliated cells (cell body)	RT		21	7.3	6.4E-2	6.9E-1
	HPA_NORMAL_TISSUE_CELLTYPE	adipose tissue; adipocytes	RT		59	20.6	7.3E-2	6.9E-1
	HPA_NORMAL_TISSUE_CELLTYPE	colon; endothelial cells	RT		81	28.3	7.4E-2	6.9E-1
	HPA_NORMAL_TISSUE_CELLTYPE	testis; sertoli cells	RT		37	12.9	7.7E-2	6.9E-1
	HPA_NORMAL_TISSUE_CELLTYPE	fallopian tube; glandular cells	RT		82	28.7	7.9E-2	6.9E-1
	HPA_NORMAL_TISSUE_CELLTYPE	pancreas; pancreatic endocrine cells	RT		80	28.0	8.2E-2	6.9E-1
	HPA_NORMAL_TISSUE_CELLTYPE	ovary; ovarian stroma cells	RT		59	20.6	8.9E-2	6.9E-1
	HPA_NORMAL_TISSUE_CELLTYPE	soft tissue 1; fibroblasts	RT		62	21.7	9.4E-2	6.9E-1
	HPA_NORMAL_TISSUE_CELLTYPE	duodenum; endocrine cells	RT		11	3.8	9.6E-2	6.9E-1
	HPA_NORMAL_TISSUE_CELLTYPE	cerebral cortex; endothelial cells	RT		67	23.4	9.8E-2	6.9E-1

Figure 7: Streptococcus pneumoniae versus Homo sapiens – tissue expressions

Table 1: Summary of the genes and pathways associated with overlap of the influenza virus (H1N1) and *Streptococcus pneumoniae* genome

Influenza virus (H1N1)		*Streptococcus pneumoniae*	
Genes	Pathways	Genes	Pathways
HIST1H16	NFKB1	DNAJB5	sHSP/HSPB
FOFR4	mTOR	HSPA5	Hsp70/HSPB
ERBB2	EGR1	HSPA1A	Hsp40/DNAJ
KRAS	SMAD4	HSPA8	Hsp90/HSDC
EGFR	TBL1X	HSPA9	Chaperonin
TPJ3	YY1	HSPD1	CCT-Hsp60
KAT6A	AKT	HSPA1B	
CTIF		PRKN	
DAZ1		SNCA	
DAZ4		TUBB/TUBA	

during stress, inflammation, etc. They are targets for cancer therapy, inflammation, myocardial ischemia, transplantation, and neurodegenerative diseases' activity modulation [11].

The TBL1X gene is associated with central hypothyroidism and hearing loss in children. NFKB1 is actively involved in inflammatory and immune responses. SMAD4 is involved in activation of normal cells and apoptosis of abnormal cells. mTOR is involved in cellular metabolism, immune responses, survival, and proliferation. EGR1 is involved in tissue repair after and the process of fibrosis. DAZ1/4 is deleted in azoospermia. CTIF is involved in active perinuclear transcription.

The Parkin protein, which aggressively breaks down extraneous proteins by labeling the damaged and surplus proteins with ubiquitin, is made by the PRKN gene. Early-onset parkinsonism is known to be caused by mutations in the PRKN gene [12] Microtubule-related genes TUBB and TUBA and genes SEPT are associated with particular polymerization during mitosis. The chaperone proteins HSC70 and HSP70's adenosine triphosphatase activity is bound and inhibited by the CHIP protein, which is encoded by the STUB1 gene, preventing their forward processes [13] Defects in the protein alpha-synuclein encoded by SNCA are linked to parkinsonism disorder. It has been demonstrated that GPR37 [14], which encodes the G-protein coupled receptor protein, interacts with Parkin (ligase) and HSPA1A. GPR37 is glial and neuroprotective factor receptors [15]. Activation of GPR37 has been observed to be neuroprotective in stroke.

As shown in Figure 8, some of the actively involved cellular pathways related genes associated with influenza-Homo sapiens overlap include extracellular structure organization ITGBF, LAMA2, LAMB2, ITG A3 and ITG B5, positive regulation of cell motility through Akt1, PLCG1, regulatory region DNA binding through NFKB1, mTOR, SMAD4, and EGR1 pathways, and PRPF, SRSF 1/11, and CSTF pathways associated with RNA processing. Oncogenes and metabolic pathways both interact significantly with the AKT pathway [16]. AKT pathway stimulation has been found to be increased in the gene expression in pancreas, breast and prostate tumors. Numerous target molecules are being synthesized to target the AKT pathway as a potential treatment for cancer [17]. Poor survival and increased tumor growth are linked to higher PLCG1 [18]. PLCG1 mutations are associated with skin malignancies [19].

The biological processes of stress and inflammation, nuclear export, and gene expression regulation all require MAPKAPK2. It is crucial for the control of inflammation and tumors [20], [21]. Some

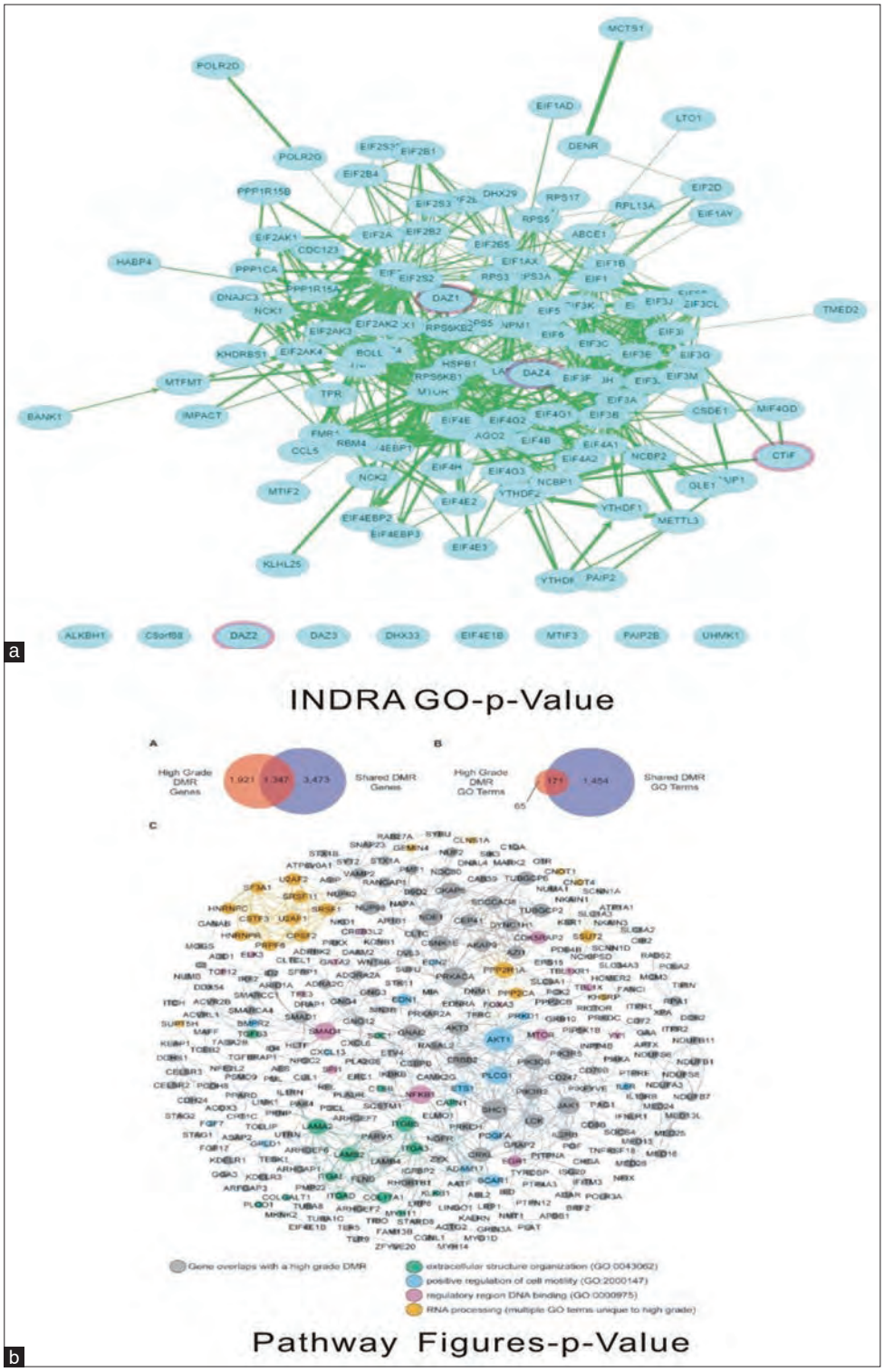


Figure 8: (a and b) Gene ontology and pathway analysis through NDEx – influenza virus versus Homo sapiens

of the MAPKAPK2-related tumors are skin, prostrate, colorectal, bladder cancers, and DSS-induced ischemic colitis [22]. ELAVL1 largely links the 3' UTSs of interferon-stimulated genes to mRNA stability [23]. One of the most crucial pathways in mammalian cells, the EGFR signaling pathway controls a number of crucial processes, including intercellular communication during development as well as proliferation, migration, differentiation, and death. A key gene in the etiology of lung cancer is EGFR [24] EGFR is also associated as a growth marker with carcinoma of stomach [25].

TP53 gene is associated with influenza vaccination genome overlap. TP53 gene encodes p53 protein, and mutation in TP53 is associated with mutant p53 protein accumulation, which highly associated with malignancies including cervical, ovarian, and endometrial malignancies [26]. Hence, a negative regulation with TP53 can be associated reduced

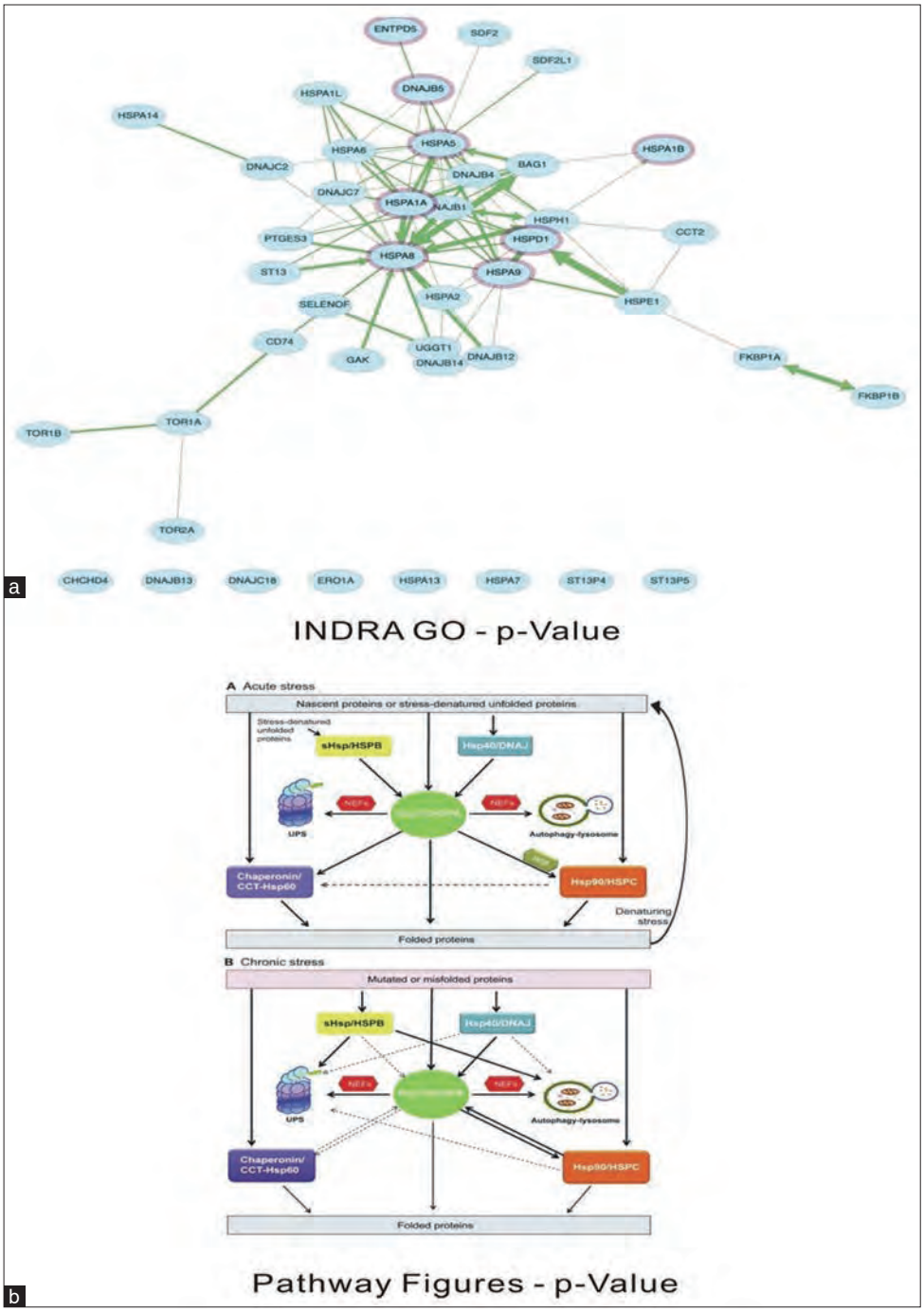


Figure 9: (a and b) Gene ontology and pathway analysis through NDEx – Streptococcus pneumoniae versus Homo sapiens

malignancies. TP53 mutations are commonly observed in ovarian cancers and acute myeloid leukaemia [27].

DNAJB5 is a marker of cervical malignancy [28]. ITAG3 is involved in immune cell infiltration [29]. ITGA3 or very common antigen 2 interacts with many extracellular proteins. ITGB5 is associated with intercellular communication during tumor progression and metastasis.

The main transcriptional regulator of heat shock proteins is HSF1. HSPA9 is a member of the HSP70 gene family. This contributes to cellular growth, stress reactivity, and mitochondrial maintenance. In addition, HSPA9 controls apoptosis. *In vivo*, it was demonstrated that this kinase's substrate is HSP [30].

HSP A9 has an inverse association with myeloid malignancies and myelodysplastic syndromes since its expression is reduced by deletion in 5q [31]. HSP 70 chaperone network system is associated with a wide range of protein folding, synthesis, and degradation and thereby has wide biological functions [32], [33]. GCase catalyzes the breakdown of the minor lipid glucosphingosine into sphingosine and water as well as the major glycolipid glucosylceramide-1 into glucose and ceramide [34]. GCase enhancement is used as a method of parkinsonism therapy [35]. GCase mutations can lead to lysosomal and mitochondrial problems.

Genes related to KRAS are involved in cellular division, growth, survival, and degeneration. Active research is being done to identify the regulatory molecules of KRAS [36]. KRAS mutations are associated with colonic cancer. The presence of KRAS mutation is associated with poor survival and aggressiveness of the tumor.

Further studies involving the protein encoded by these overlapping nucleotide sequences and the disease association in large population data would provide more information about the changes or induction of the cellular functions. Similarly, other closely associated microorganisms like commensals can also be studied to better understand the cause-effect/associations of various disease pathology in humans, such as autoimmune disorders. Incidence of non-communicable diseases, such as coronary artery disease and diabetes, is also interesting to study if there is any influence by the microbiological changes.

Conclusion

The study shows overlapping nucleotide sequences between the human genome and *S. pneumoniae* and influenza (H1N1) virus genome sequences. These overlapping sequences are also associated with various functional annotations. Further, evaluations can help to understand the modifications of the encoded proteins and metabolism during these infections.

Authors' Contributions

MCA: Conceived the idea, performed the analysis in DAVID and NDEx and wrote the paper. JW, JL, and RH: Performed the BLASTn analysis and derived the results in BLASTn.

Supplement Data Files

1. Arokiaraj, Mark Christopher; Wilson, Jarad; Lu, Jingqiao; Huang, Ruo-Pan (2023). BLAST Streptococcus Pneumoniae strain Hu17 against Homo sapiens hg38. figshare. Dataset. https://doi.org/10.6084/m9.figshare.22639783.v1
2. Arokiaraj, Mark Christopher; Wilson, Jarad; Lu, Jingqiao; Huang, Ruo-Pan (2023). Influenza H1N1 2009 California against Homo sapiens hg38 - Blast analysis. figshare. Dataset. https://doi.org/10.6084/m9.figshare.22651180.v1
3. Arokiaraj, Mark Christopher; Wilson, Jarad; Lu, Jingqiao; Huang, Ruo-Pan (2023). Streptococcus Pneumonia Hu17 and Influenza (H1N1) Virus Genome Study against Human Genome Sequences – DAVID analysis. figshare. Dataset. https://doi.org/10.6084/m9.figshare.23559300.v1

References

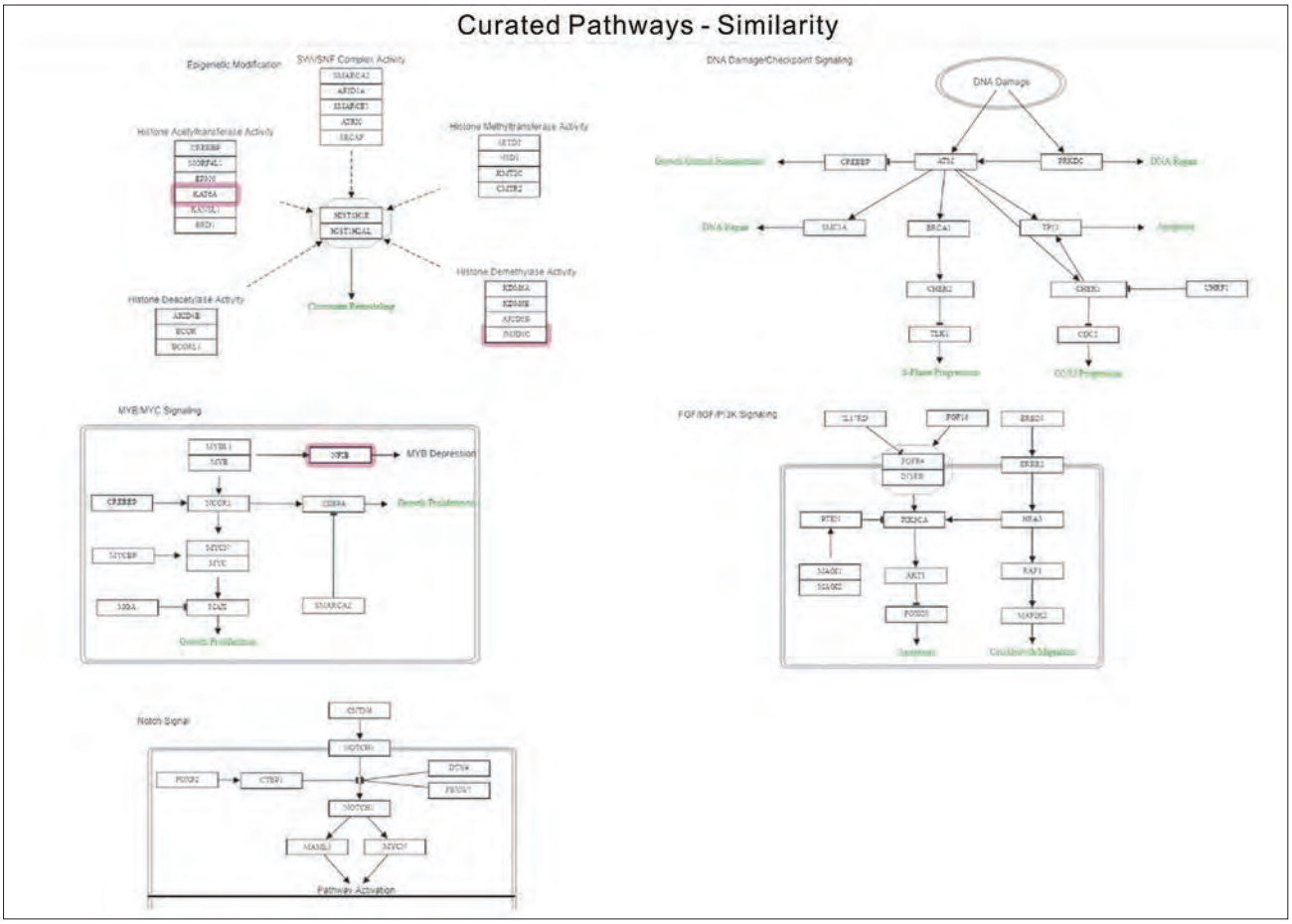
1. Arokiaraj MC. Correlation of Influenza Vaccination and Influenza Incidence on COVID-19 Severity and Other Perspectives. Available from: https://ssrn 572814 [Last accessed on 2020 Apr 10].
2. Behrouzi B, Bhatt DL, Cannon CP, Vardeny O, Lee DS, Solomon SD, *et al*. Association of influenza vaccination with cardiovascular risk: A meta-analysis. JAMA Netw Open. 2022;5(4):e228873. https://doi.org/10.1001/jamanetworkopen.2022.8873
   PMid:35486404
3. Jaiswal V, Ang SP, Lnu K, Ishak A, Pokhrel NB, Chia JE, *et al*. Effect of pneumococcal vaccine on mortality and cardiovascular outcomes: A systematic review and meta-analysis. J Clin Med. 2022;11(13):3799. https://doi.org/10.3390/jcm11133799
   PMid:35807082
4. Chung H, Buchan SA, Campigotto A, Campitelli MA, Crowcroft NS, Dubey V, *et al*. Influenza vaccine effectiveness against all-cause mortality following laboratory-confirmed influenza in older adults, 2010-2011 to 2015-2016 seasons in Ontario, Canada. Clin Infect Dis. 2020;73(5):e1191-9. https://doi.org/10.1093/cid/ciaa1862
   PMid:33354709
5. Debisarun PA, Gössling KL, Bulut O, Kilic G, Zoodsma M, Liu Z, *et al*. Induction of trained immunity by influenza vaccination-impact on COVID-19. PLoS Pathog. 2021;17(10):e1009928. https://doi.org/10.1371/journal.ppat.1009928
   PMid:34695164
6. Talamantes-Becerra B, Carling J, Georges A. omicR: A tool to facilitate BLASTn alignments for sequence data. SoftwareX. 2021;14:100702. https://doi.org/10.1016/j.softx.2021.100702
7. Lobo I. Basic local alignment search tool (BLAST). Nat Educ. 2008;1(1):21.
8. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, *et al*. DAVID: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update. Nucleic Acids Res. 2022;50(W1):W216-21. https://doi.org/10.1093/nar/gkac194
   PMid:35325185
9. Wilcox CR, Islam N, Dambha-Miller H. Association between influenza vaccination and hospitalisation or all-cause mortality in people with covid-19: A retrospective cohort study. BMJ Open Respir Res. 2021;8(1):e000857. https://doi.org/10.1136/bmjresp-2020-000857
   PMid:33664123
10. Arokiaraj MC. Considering interim interventions to control COVID-19 associated morbidity and mortality-perspectives. Front Public Health. 2020;8:444. https://doi.org/10.3389/fpubh.2020.00444
    PMid:33072682

11. Dubey A, Prajapati KS, Swamy M, Pachauri V. Heat shock proteins: A therapeutic target worth to consider. Vet World. 2015;8(1):46-51. https://doi.org/10.14202/vetworld.2015.46-51
PMid:27046995
12. Castelo Rueda MP, Raftopoulou A, Gögele M, Borsche M, Emmert D, Fuchsberger C, *et al*. Frequency of heterozygous parkin (PRKN) variants and penetrance of parkinson's disease risk markers in the population-based CHRIS Cohort. Front Neurol. 2021;12:706145. https://doi.org/10.3389/fneur.2021.706145
PMid:34434164
13. Zhang S, Hu ZW, Mao CY, Shi CH, Xu YM. CHIP as a therapeutic target for neurological diseases. Cell Death Dis. 2020;11(9):727. https://doi.org/10.1038/s41419-020-02953-5
PMid:32908122
14. Guo Y, Sun Y, Song Z, Zheng W, Xiong W, Yang Y, *et al*. Genetic analysis and literature review of SNCA variants in Parkinson's disease. Front Aging Neurosci. 2021;13:648151. https://doi.org/10.3389/fnagi.2021.648151.
PMid:34456707
15. Kunihiro J, Nabeka H, Wakisaka H, Unuma K, Khan MS, Shimokawa T, *et al*. Prosaposin and its receptors GRP37 and GPR37L1 show increased immunoreactivity in the facial nucleus following facial nerve transection. PLoS One. 2020;15(12):e0241315. https://doi.org/10.1371/journal.pone.0241315
PMid:33259479
16. Nitulescu GM, Van De Venter M, Nitulescu G, Ungurianu A, Juzenas P, Peng Q, *et al*. The AKT pathway in oncology therapy and beyond (review). Int J Oncol. 2018;53(6):2319-31. https://doi.org/10.3892/ijo.2018.4597
PMid:30334567
17. He Y, Sun MM, Zhang GG, Yang J, Chen KS, Xu WW, *et al*. Targeting PI3K/Akt signal transduction for cancer therapy. Signal Transduct Target Ther. 2021;6(1):425. https://doi.org/10.1038/s41392-021-00828-5
PMid:34916492
18. Vaqué JP, Gómez-López G, Monsálvez V, Varela I, Martínez N, Pérez C, *et al*. PLCG1 mutations in cutaneous T-cell lymphomas. Blood. 2014;123(13):2034-43. https://doi.org/10.1182/blood-2013-05-504308
PMid:24497536
19. Li T, Yang Z, Li H, Zhu J, Wang Y, Tang Q, *et al*. Phospholipase Cγ1 (PLCG1) overexpression is associated with tumor growth and poor survival in IDH wild-type lower-grade gliomas in adult patients. Lab Invest. 2022;102(2):143-53. https://doi.org/10.1038/s41374-021-00682-7
PMid:34697421
20. Fiore M, Forli S, Manetti F. Targeting mitogen-activated protein kinase-activated protein kinase 2 (MAPKAPK2, MK2): Medicinal chemistry efforts to lead small molecule inhibitors to clinical trials. J Med Chem. 2015;59(8):3609-34. https://doi.org/10.1021/acs.jmedchem.5b01457
PMid:26502061
21. Soni S, Anand P, Padwad YS. MAPKAPK2: The master regulator of RNA-binding proteins modulates transcript stability and tumor progression. J Exp Clin Cancer Res. 2019;38(1):121. https://doi.org/10.1186/s13046-019-1115-1
PMid:30850014
22. Berggren KL, Restrepo Cruz S, Hixon MD, Cowan AT, Keysar SB, Craig S, *et al*. MAPKAPK2 (MK2) inhibition mediates radiation-induced inflammatory cytokine production and tumor growth in head and neck squamous cell carcinoma. Oncogene. 2019;38(48):7329-41. https://doi.org/10.1038/s41388-019-0945-9
PMid:31417185
23. Rothamel K, Arcos S, Kim B, Reasoner C, Lisy S, Mukherjee N, *et al*. ELAVL1 primarily couples mRNA stability with the 3' utrs of interferon-stimulated genes. Cell Rep. 2021;35(8):109178. https://doi.org/10.1016/j.celrep.2021.109178
PMid:34038724
24. Carcereny E, Morán T, Capdevila L, Cros S, Vilà L, De Los Llanos Gil M, *et al*. The epidermal growth factor receptor (EGRF) in lung cancer. Transl Respir Med. 2015;3:1. https://doi.org/10.1186/s40247-015-0013-z
PMid: 25810955
25. Wang D, Wang B, Wang R, Zhang Z, Lin Y, Huang G, *et al*. High expression of EGFR predicts poor survival in patients with resected T3 stage gastric adenocarcinoma and promotes cancer cell survival. Oncol Lett. 2017;13(5):3003-13. https://doi.org/10.3892/ol.2017.5827
PMid:28521408
26. Hu J, Cao J, Topatana W, Juengpanich S, Li S, Zhang B, *et al*. Targeting mutant p53 for cancer therapy: Direct and indirect strategies. J Hematol Amp Oncol. 2021;14(1):157. https://doi.org/10.1186/s13045-021-01169-0
27. Kennedy MC, Lowe SW. Mutant p53: It's not all one and the same. Cell Death Differ. 2022;29(5):983-7. https://doi.org/10.1038/s41418-022-00989-y
PMid:35361963
28. Yu H, Chen Y, Zhang G, Wu H, Wang K. DNAJB5 (HSC40) gene as a novel biomarker for cervical cancer. Eur J Gynaecol Oncol. 2022;43(5):18-23. https://doi.org/10.22514/ejgo.2022.023
29. Li Y, Li F, Bai X, Li Y, Ni C, Zhao X, *et al*. ITGA3 is associated with immune cell infiltration and serves as a favorable prognostic biomarker for breast cancer. Front Oncol. 2021;11:658547. https://doi.org/10.3389/fonc.2021.658547
PMid:34094951
30. Krysiak K, Tibbitts JF, Shao J, Liu T, Ndonwi M, Walter MJ. Reduced levels of Hspa9 attenuate Stat5 activation in mouse B cells. Exp Hematol. 2015;43(4):319-30. https://doi.org/10.1016/j.exphem.2014.12.005
PMid:25550197
31. Liu T, Krysiak K, Shirai CL, Kim S, Shao J, Ndonwi M, *et al*. Knockdown of Hspa9 induces TP53-dependent apoptosis in human hematopoietic progenitor cells. PLoS One. 2017;12(2):e0170470. https://doi.org/10.1371/journal.pone.0170470
PMid:28178280
32. Rosenzweig R, Nillegoda NB, Mayer MP, Bukau B. The HSP70 chaperone network. Nat Rev Mole Cell Biol. 2019;20(11):665-80. https://doi:10.1038/s41580-019-0133-3
PMid:31253954
33. Wang Z, Li Y, Yang X, Zhao J, Cheng Y, Wang J. Mechanism and complex roles of HSC70 in viral infections. Front Microbiol. 2020;11:1577. https://doi.org/10.3389/fmicb.2020.01577
PMid:32849328
34. Menozzi E, Schapira AH. Enhancing the activity of glucocerebrosidase as a treatment for Parkinson disease. CNS Drugs. 2020;34(9):915-23. https://doi.org/10.1007/s40263-020-00746-0
PMid:32607746.
35. Klein AD, Outeiro TF. Glucocerebrosidase mutations Disrupt the lysosome and now the mitochondria. Nat Commun. 2023;14(1):6383. https://doi.org/10.1038/s41467-023-42107-7
PMid:37821433
36. Huang L, Guo Z, Wang F, Fu L. KRAS mutation: From undruggable to druggable in cancer. Signal Transduct Target Ther. 2021;6(1):386. https://doi.org/10.1038/s41392-021-00780-4
PMid:34776511

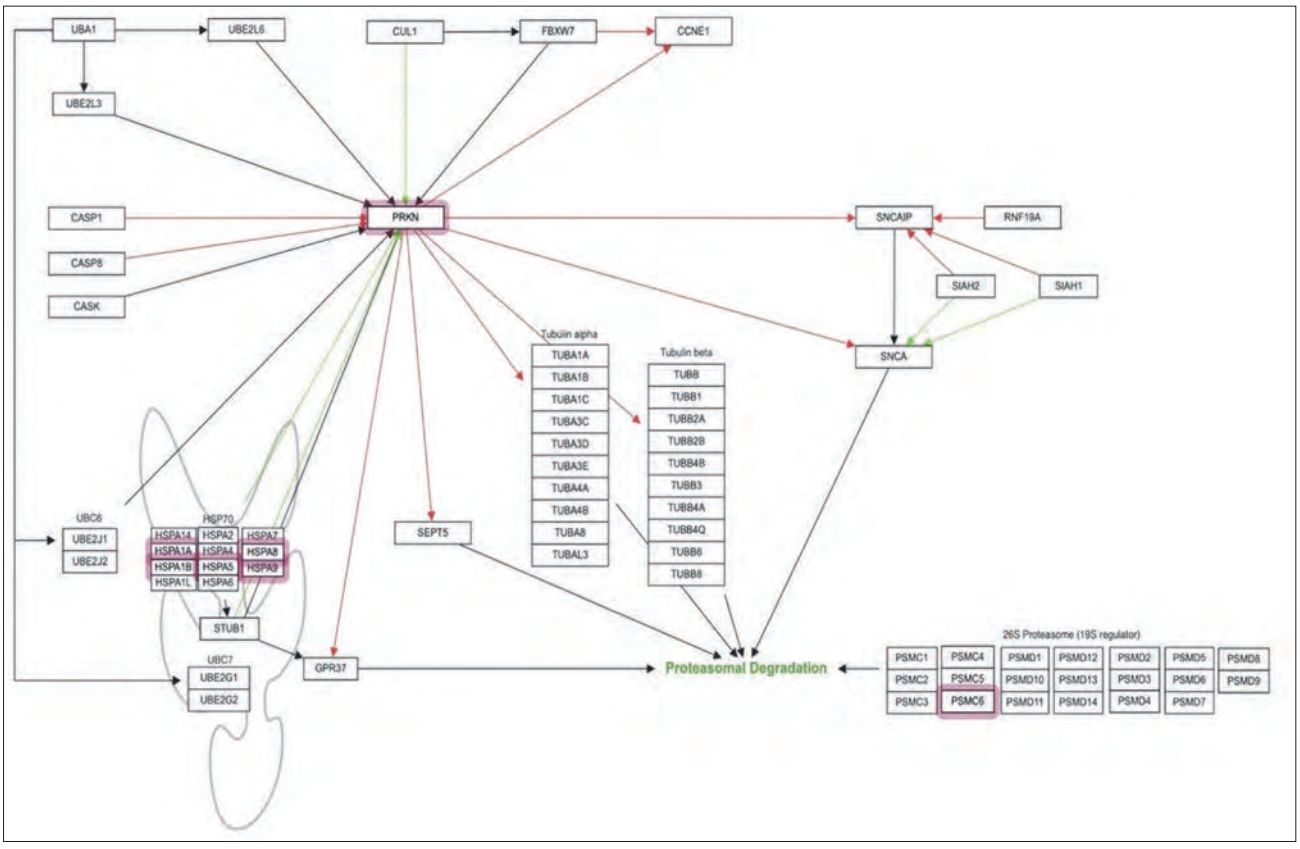
Supplement Figures



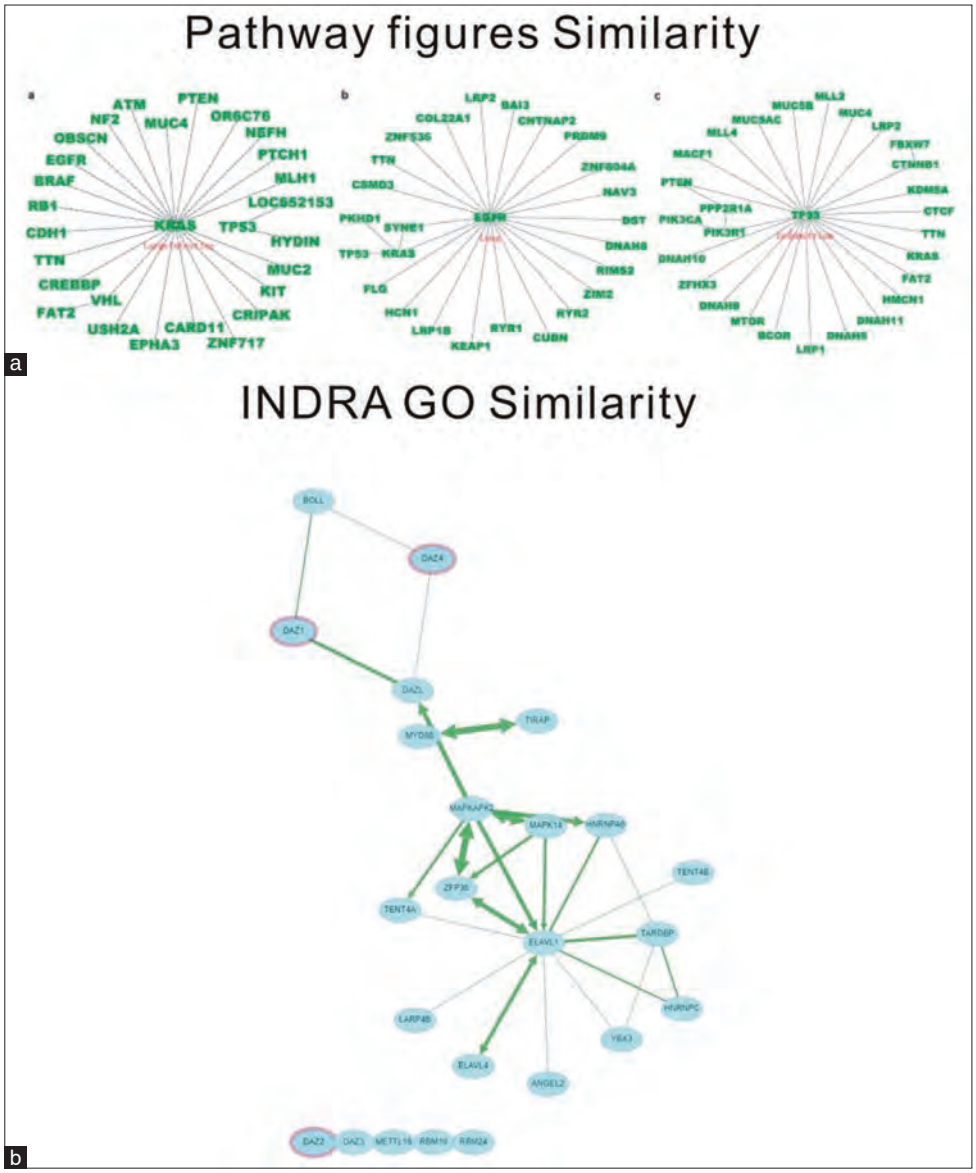
Supplement Figure 1: Streptococcus pneumoniae versus Homo sapiens overlapping genes functional analysis, BIOGRID interactions by DAVID



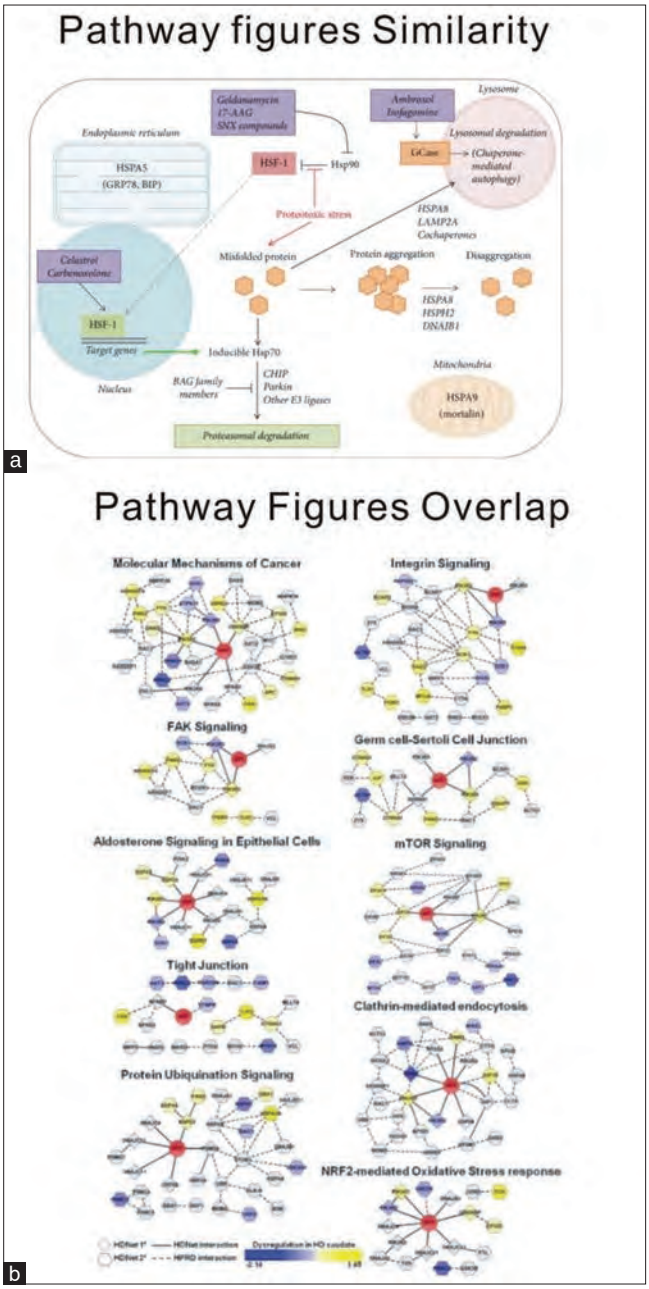
Supplement Figure 2: Curated pathways by NDEx of genes concerning influenza virus genome versus Homo sapiens



Supplement Figure 3: Curated pathways by NDEx of genes concerning Streptococcus pneumoniae versus Homo sapiens



Supplement Figure 4: (a and b) Influenza H1N1 versus. Homo sapiens, overlapping gene analysis by NDEx. Panel A shows pathways and similarity and Panel B INDRO GO similarity



Supplement Figure 5: (a and b) Streptococcus pneumoniae versus Homo sapiens, overlapping gene analysis by NDEx. Panel A shows pathways and similarity and Panel B shows pathway figures overlap