

# Bioinformatics Report HW1 Group 5

## Rheumatoid arthritis-related human-oral microbiome proteins

Giorgio Giannone<sup>1</sup>, Livia Lombardi<sup>1</sup>

### Abstract

The whole system of microbiome represents a prosthesis of the human organism through its several interactions. Human healthy structure is strictly related to the existence of microbes. Rheumatoid arthritis (RA) is the most common inflammatory autoimmune disorder that appears when the immune system doesn't work properly. In this scenario, multiple genes and environment factors act to cause pathologic disorders; but genetic predisposition seems to have the main correlation/susceptibility with it. So, a RA's risk has a big inheritance component. Furthermore, some recent researches have pointed out that ACPA proteins (anti-citrullinated) presence could be responsible for the trigger of the disease, since are occurring in 50% of RA's cases. Because of the abundance and uncertain evidences which cause this pathology could be interesting look at the problem from an informatics point of view and try to figure out some type of dependency in the triangle genes-interactions-disease.[1]

### I. INTRODUCTION

**R**HEUMATOID ARTHRITIS (RA) causes chronic inflammation, affecting mostly joints but also other tissues and organs. There are evidences of a link between RA and Periodontal Diseases (PD) which has similar mechanism of manifestation. Rheumatoid arthritis and Periodontal Disease both present an overproduction of a variety of proteins, protagonist of the disease activity. Unlike the RA, PD shows clearly bacterial causes which can easy activate immunological reactions. Certain genes have been identified that increases the risk of RA contraction, or even some environmental factors or smoking activity, but since the variety of joint disease is hard to assign RA diagnosis.

### II. COLLECT BASIC INFORMATION ABOUT SEED GENES

Starting from a seed genes list which includes a subset of genes involved in the rheumatoid arthritis disease, we can retrieve through different techniques some essential associated information. Using REST requests and the 'bioservice' library, we retrieved all the features and correspondences about the genes in the list. In particular we collected these information from the Uniprot and HGNC websites which are responsible for approving symbols and names for genes within a chromosome. The Table I shows all the features considered in the collection. For every main protein name we report:

- the official gene symbol approved in the HGNC database
- the 'Uniprot ID' Entry name that is a unique identifier often containing biological information
- the 'Uniprot AC' accession number that is assigned to each sequence once included in the UniprotKB database
- the Entrez ID that consists of unique integers identifiers

In the table II a brief description of genes functions is reported.

### III. COLLECT INTERACTION DATA

We processed genes interaction using a server requests in python from three different Online Interaction Repository: Apid, Biogrid, String. Under the constraint of genes belonging to the human species (characterized by the number "9606") we extracted all the binary protein interactions - or at least the level 1 interactome. The table III show the interaction information retrieved from the three different servers.

#### IV. ARRANGE INTERACTION DATA

In this section we studied the interaction between the selected genes and proteins. In particular we studied:

- the interactome: the interactions among seed genes
- the union interactome: the interaction between genes and proteins found in at least one database
- the intersection interactome: the interaction between genes and proteins in all database

The Intersection interactome information is showed in the Table IV.

#### V. ENRICHMENT ANALYSIS

In order to retrieve the functional profile of the gene list and then understand the related biological processes the enrichment analysis plays a fundamental role and since we obtained the first and second order interaction among genes and proteins, we can perform it. The first enrichment technique is called Gene Ontology (GO), this is a method based on the idea of associating a collection of genes with a functional biological term, so genes are classified given their different characteristics in the so called 'bins'. Given a genes list, through the GO technique and thanks to some statistical test, we can observe which bin has been enriched for the input gene list.[2] In the enrichment analysis scenario another used technique is called 'Pathway Analysis', it consists in the identification or the construction of a pathway starting from a set of proteins. More in detail a pathway is a process involved in a cell or in a tissue. Watching proteins' interactions, the pathways analysis let to identify the diseases most statistically related to the triggered proteins chains.[3] In particular we perform the techniques both on 'Enrichr', a comprehensive gene set enrichment analysis, that computes three types of enrichment scores:

- p-Value : the closer to zero, the more significant the particular GO term associated with the group of genes is or in the case of the pathway analysis the more probably the disease is
- Z-score : the deviation from the expected rank of the test
- Combined Score : a combined score that multiplies the log of the p-value by the z-score

In the tables V,V, V we show the first 10 over-represented GO (Gene Ontology) categories and in the tables V,V, V the first 10 over-represented pathways for:

- the seed genes
- the union interactome
- the intersection interactome

#### VI. CONCLUSION

Enrichment analysis results show some relevant information connected to the AR disease. We can see the presence of the autophosphorylation protein (Table V-V): in literature we can find the existence of a strictly relationship among the rheumatoid arthritis and the increase of this protein[4]. Another example is represented by the massive presence of the RNA polymerase II: his pause and elongation is controlled by some transcription factors and when their expression is perturbed can produce forms of autoimmunity or even tumor cells. Even the Tumor Necrosis Factor (TNF $\alpha$ ) has been linked to joint destruction in rheumatoid arthritis. Specifically RNA Polymerase II initiates transcription in resting macrophages, but stalls near the promoter until lipopolysaccharide (LPS) triggers rapid release of the Negative Elongation Factor (NELF) productive elongation; polymerase stalling is connected to the inflammatory gene activation.[5][6]

## REFERENCES

- [1] de la Rica, Lorenzo, et al. "Identification of novel markers in rheumatoid arthritis through integrated analysis of DNA methylation and microRNA expression." *Journal of autoimmunity* 41 (2013): 6-16.
- [2] <http://geneontology.org/page/go-enrichment-analysis>
- [3] Chen, Edward Y., et al. "Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool." *BMC bioinformatics* 14.1 (2013): 128.
- [4] Luukkonen, Jani, et al. "Increased amount of phosphorylated proinflammatory osteopontin in rheumatoid arthritis synovia is associated to decreased tartrate-resistant acid phosphatase 5B/5A ratio." *PloS one* 12.8 (2017): e0182904.
- [5] Adelman, Karen, et al. "Immediate mediators of the inflammatory response are poised for gene activation through RNA polymerase II stalling." *Proceedings of the National Academy of Sciences* 106.43 (2009): 18207-18212.
- [6] Lee, Tong Ihn, and Richard A. Young. "Transcriptional regulation and its misregulation in disease." *Cell* 152.6 (2013): 1237-1251

Table I  
TABLE SEED GENES INFORMATIONS

	Protein name	Symbol HGNC	Uniprot ID	Uniprot AC	Entrez ID
1	Arf-GAP with coiled-coil ANK repeat and PH domain-containing protein 2	ACAP2	ACAP2_HUMAN	Q15057	23527
2	Arf-GAP with dual PH domain-containing protein 1	ADAP1	ADAP1_HUMAN	O75689	11033
3	Arf-GAP domain and FG repeat-containing protein 1	AGFG1	AGFG1_HUMAN	P52594	3267
4	Annexin A4	ANXA4	ANXA4_HUMAN	P09525	307
5	Aurora kinase C	AURKC	AURKC_HUMAN	Q9UQB9	6795
6	BRCA2 and CDKN1A-interacting protein	BCCIP	BCCIP_HUMAN	Q9P287	56647
7	Beclin-1	BECN1	BECN1_HUMAN	Q14457	8678
8	B2 bradykinin receptor	BDKRB2	BKRB2_HUMAN	P30411	624
9	Breast cancer type 2 susceptibility protein	BRCA2	BRCA2_HUMAN	P51587	675
10	Calreticulin	CALR	CALR_HUMAN	P27797	811
11	Caveolin-1	CAV1	CAV1_HUMAN	Q03135	857
12	F-actin-capping protein subunit alpha-2	CAPZA2	CAZA2_HUMAN	P47755	830
13	Cell division control protein 42 homolog	CDC42	CDC42_HUMAN	P60953	998
14	Cyclin-dependent kinase 1	CDK1	CDK1_HUMAN	P06493	983
15	Clathrin heavy chain 1	CLTC	CLH1_HUMAN	Q00610	1213
16	Cullin-2	CUL2	CUL2_HUMAN	Q13617	8453
17	Serine/threonine-protein kinase DCLK1	DCLK1	DCLK1_HUMAN	O15075	9201
18	Myotonin-protein kinase	DMPK	DMPK_HUMAN	Q09013	1760
19	DNA polymerase delta catalytic subunit	POLD1	DPOD1_HUMAN	P28340	5424
20	Interferon-induced double-stranded RNA-activated protein kinase	EIF2AK2	E2AK2_HUMAN	P19525	5610
21	Protein enabled homolog	ENAH	ENAH_HUMAN	Q8N8S7	55740
22	Fanconi anemia group J protein	BRIP1	FANCJ_HUMAN	Q9BX63	83990
23	ADP-ribosylation factor-binding protein GGA2	GGA2	GGA2_HUMAN	Q9UJY4	23062
24	Glucosamine 6-phosphate N-acetyltransferase	GNPNAT1	GNA1_HUMAN	Q96EK6	64841
25	Heat shock factor protein 4	HSF4	HSF4_HUMAN	Q9ULV5	3299
26	Interferon-induced protein with tetratricopeptide repeats 3	IFIT3	IFIT3_HUMAN	O14879	3437
27	Interleukin enhancer-binding factor 3	ILF3	ILF3_HUMAN	Q12906	3609
28	Integrin alpha-V	ITGAV	ITAV_HUMAN	P06756	3685
29	Kinesin-like protein KIF3A	KIF3A	KIF3A_HUMAN	Q9Y496	11127
30	Katanin p80 WD40 repeat-containing subunit B1	KATNB1	KTNB1_HUMAN	Q9BVA0	10300
31	Prelamin-A/C	LMNA	LMNA_HUMAN	P02545	4000
32	Mitogen-activated protein kinase kinase kinase 8	MAP3K8	M3K8_HUMAN	P41279	1326
33	Myocyte-specific enhancer factor 2D	MEF2D	MEF2D_HUMAN	Q14814	4209
34	Mitogen-activated protein kinase 3	MAPK3	MK03_HUMAN	P27361	5595
35	Dual specificity mitogen-activated protein kinase kinase 4	MAP2K4	MP2K4_HUMAN	P45985	6416
36	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 13	NDUFA13	NDUAD_HUMAN	Q9P0J0	51079
37	Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform	PPP3CA	PP2BA_HUMAN	Q08209	5530
38	Tyrosine-protein phosphatase non-receptor type 1	PTPN1	PTN1_HUMAN	P18031	5770
39	Rab5 GDP/GTP exchange factor	RABGEF1	RABX5_HUMAN	Q9UJ41	27342
40	Ras-related protein Rap-2a	RAP2A	RAP2A_HUMAN	P10114	5911
41	Ras-related protein Rab-27A	RAB27A	RB27A_HUMAN	P51159	5873
42	Histone-binding protein RBBP4	RBBP4	RBBP4_HUMAN	Q09028	5928
43	Serine incorporator 3	SERINC3	SERC3_HUMAN	Q13530	10955
44	S-phase kinase-associated protein 1	SKP1	SKP1_HUMAN	P63208	6500
45	Sortilin	SORT1	SORT_HUMAN	Q99523	6272
46	Transcription factor PU.1	SPI1	SPI1_HUMAN	P17947	6688
47	Sequestosome-1	SQSTM1	SQSTM_HUMAN	Q13501	8878
48	Serum response factor	SRF	SRF_HUMAN	P11831	6722
49	Signal transducer and activator of transcription 1-alpha/beta	STAT1	STAT1_HUMAN	P42224	6772
50	Signal transducer and activator of transcription 3	STAT3	STAT3_HUMAN	P40763	6774
51	Signal transducer and activator of transcription 6	STAT6	STAT6_HUMAN	P42226	6778
52	Tubulin alpha-4A chain	TUBA4A	TBA4A_HUMAN	P68366	7277
53	Tumor necrosis factor ligand superfamily member 6	FASLG	TNFL6_HUMAN	P48023	356
54	Ubiquitin-conjugating enzyme E2 L3	UBE2L3	UB2L3_HUMAN	P68036	7332

Table II  
TABLE SEED GENES DESCRIPTION

	Description
1	ANK repeat and PH domain-containing protein 2"
2	GTPase-activating protein for the ADP ribosylation factor family.
3	Arf-GAP domain and FG repeats-containing protein 1 is a protein that in humans is encoded by the AGFG1 gene.
4	Calcium/phospholipid-binding protein which promotes membrane fusion and is involved in exocytosis.
5	Serine/threonine-protein kinase 13 is an enzyme that in humans is encoded by the AURKC gene.
6	BRCA2 and CDKN1A-interacting protein is a protein that in humans is encoded by the BCCIP gene.
7	Beclin-1 is a protein that in humans is encoded by the BECN1 gene.
8	"Bradykinin receptor B2 is a G-protein coupled receptor for bradykinin
9	"BRCA2 and BRCA2 () are a human gene and its protein product
10	"Calcium-binding chaperone that promotes folding
11	May act as a scaffolding protein within caveolar membranes. Interacts directly with G-protein alpha subunits and can functionally regulate their activity (By similarity).
12	F-actin-capping protein subunit alpha-2 also known as CapZ-alpha2 is a protein that in humans is encoded by the CAPZA2 gene.
13	"Cell division control protein 42 homolog
14	Cyclin-dependent kinase 1 also known as CDK1 or cell division cycle protein 2 homolog is a highly conserved protein that functions as a serine/threonine kinase.
15	Clathrin heavy chain 1 is a protein that in humans is encoded by the CLTC gene.
16	Cullin-2 is a protein that in humans is encoded by the CUL2 gene.
17	Serine/threonine-protein kinase DCLK1 is an enzyme that in humans is encoded by the DCLK1 gene.
18	Functions as a sorting receptor in the Golgi compartment and as a clearance receptor on the cell surface.
19	"The gene polymerase delta 1 (POLD1) encodes the large
20	double-stranded RNA-activated protein kinase"
21	Ena/VASP proteins are actin-associated proteins involved in a range of processes dependent on cytoskeleton.
22	Fanconi anemia group J protein is a protein that in humans is encoded by the BRCA1-interacting protein 1 (BRIP1) gene.
23	ADP-ribosylation factor-binding protein GGA2 is a protein that in humans is encoded by the GGA2 gene.
24	N-acetyltransferase is an enzyme that catalyzes the transfer of acetyl groups from acetyl-CoA to arylamines.
25	Heat shock factor protein 4 is a protein that in humans is encoded by the HSF4 gene.
26	Interferon-induced protein with tetratricopeptide repeats 3 is a protein that in humans is encoded by the IFIT3 gene.
27	Interleukin enhancer-binding factor 3 is a protein that in humans is encoded by the ILF3 gene.
28	Integrin alpha-V is a protein that in humans is encoded by the ITGAV gene.
29	Kinesin-like protein KIF3A is a protein that in humans is encoded by the KIF3A gene.
30	Katanin p80 WD40-containing subunit B1 is a protein that in humans is encoded by the KATNB1 gene.
31	Lamin A/C also known as LMNA is a protein that in humans is encoded by the LMNA gene.
32	Mitogen-activated protein kinase kinase kinase 8 is an enzyme that in humans is encoded by the MAP3K8 gene.
33	Myocyte-specific enhancer factor 2D is a protein that in humans is encoded by the MEF2D gene.
34	Mitogen-activated protein kinase 3 is an enzyme that in humans is encoded by the MAPK3 gene.
35	Dual specificity mitogen-activated protein kinase kinase 4 is an enzyme that in humans is encoded by the MAP2K4 gene.
36	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 13 is an enzyme that in humans is encoded by the NDUFA13 gene.
37	Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform (PP2BA) is a protein that in humans is encoded by the PPP3CA gene.
38	Tyrosine-protein phosphatase non-receptor type 1 is an enzyme that is the founding member of the protein tyrosine phosphatase (PTP) family.
39	Rab5 GDP/GTP exchange factor is a protein that in humans is encoded by the RABGEF1 gene.
40	Ras-related protein Rap-2a is a protein that in humans is encoded by the RAP2A gene.
41	Ras-related protein Rab-27A is a protein that in humans is encoded by the RAB27A gene.
42	"Histone-binding protein RBBP4 (also known as RbAp48
43	Serine incorporator 3 is a protein that in humans is encoded by the SERINC3 gene.
44	Skp1 (S-phase kinase-associated protein 1) is a small protein of approximately 160 amino acids.
45	Functions as a sorting receptor in the Golgi compartment and as a clearance receptor on the cell surface.
46	Transcription factor PU.1 is a protein that in humans is encoded by the SPI1 gene.
47	Sequestosome-1 is a protein that in humans is encoded by the SQSTM1 gene.
48	SRF is a transcription factor that binds to the serum response element.
49	Signal transducer and activator of transcription 1 (STAT1) is a transcription factor which in humans is encoded by the STAT1 gene.
50	Signal transducer and activator of transcription 3 (STAT3) is a transcription factor which in humans is encoded by the STAT3 gene.
51	Signal transducer and activator of transcription 6 (STAT6) is a human gene.
52	Tubulin alpha-4A chain is a protein that in humans is encoded by the TUBA4A gene.
53	"Cytokine that binds to TNFRSF6/FAS
54	"Ubiquitin-conjugating enzyme E2 L3 (UBE2L3)

Table III  
SUMMARY DATABASE DATA

	BioGrid	Apid	String
n. of seed genes	54	54	54
n. of interacting proteins	3898	4505	6019
n. of interactions	36301	7219	275510
n. interaction with seed genes	5834	4293	11994

Table IV  
INTERSECTION INTERACTOME

	Symbol A	Uniprot A	Symbol B	Uniprot B
0	CALR	P27797	K1C17	Q04695
1	CALR	P27797	MARE3	Q9UPY8
2	CALR	P27797	PERF	P14222
3	CALR	P27797	VWF	P04275
4	CALR	P27797	CALX	P27824
5	STAT1	P42224	PGFRA	P16234
6	STAT1	P42224	EP300	Q09472
7	STAT1	P42224	IRF1	P10914
8	STAT1	P42224	IKBA	P25963
9	STAT1	P42224	KS6B1	P23443
10	STAT1	P42224	FANCE	Q9HB96
11	STAT1	P42224	STAT2	P52630
12	STAT1	P42224	FOS	P01100
13	STAT1	P42224	JAK1	P23458
14	RBBP4	Q09028	FOG1	Q8IX07
15	RBBP4	Q09028	MTA1	Q13330
16	RBBP4	Q09028	HAT1	O14929
17	RBBP4	Q09028	CBP	Q92793
18	RBBP4	Q09028	TERA	P55072
19	RBBP4	Q09028	LIN9	Q5TKA1
20	RBBP4	Q09028	MBD3	O95983
21	MEF2D	Q14814	EP300	Q09472
22	MEF2D	Q14814	HDAC5	Q9UQL6
23	ITAV	P06756	TSP1	P07996
24	ITAV	P06756	MK01	P28482
25	M3K8	P41279	MP2K4	P45985
26	BRCA2	P51587	PCNA	P12004
27	BRCA2	P51587	SMC3	Q9UQE7
28	BRCA2	P51587	FLNA	P21333
29	BRCA2	P51587	FANCE	Q9HB96
30	BRCA2	P51587	BRCC3	P46736
31	BRCA2	P51587	MCPH1	Q8NEM0
32	BRCA2	P51587	XRCC3	O43542
33	BRCA2	P51587	BUB1B	O60566
34	BRCA2	P51587	KAT2B	Q92831
35	BRCA2	P51587	BCCIP	Q9P287
36	BRCA2	P51587	FANCG	O15287
37	BRCA2	P51587	ERCC5	P28715
38	LMNA	P02545	NU153	P49790
39	LMNA	P02545	PCBP1	Q15365
40	LMNA	P02545	WDR18	Q9BV38
41	LMNA	P02545	LBR	Q14739
42	LMNA	P02545	TM201	Q5SNT2
43	LMNA	P02545	ELYS	Q8WYP5
44	LMNA	P02545	P121A	Q96HA1
45	LMNA	P02545	NUP98	P52948
46	LMNA	P02545	NU133	Q8WUM0
47	LMNA	P02545	GANP	O60318
48	LMNA	P02545	MLIP	Q5VWP3
49	SQSTM	Q13501	UBC	P0CG48
50	SQSTM	Q13501	TRAF6	Q9Y4K3
51	SQSTM	Q13501	MLP3C	Q9BXW4
52	SQSTM	Q13501	GRIA1	P42261
53	SQSTM	Q13501	FUS	P35637
54	SQSTM	Q13501	MTOR	P42345
55	SQSTM	Q13501	KS6B1	P23443
56	SQSTM	Q13501	HSP74	P34932
57	SQSTM	Q13501	ULK2	Q8IYT8
58	SQSTM	Q13501	MK01	P28482
59	SQSTM	Q13501	ISG15	P05161

	Symbol A	Uniprot A	Symbol B	Uniprot B
60	SQSTM	Q13501	KLHL3	Q9UH77
61	SQSTM	Q13501	KPCI	P41743
62	SQSTM	Q13501	GBRAP	O95166
63	SQSTM	Q13501	DVL2	O14641
64	SQSTM	Q13501	INSR	P06213
65	SQSTM	Q13501	GRP78	P11021
66	AURKC	Q9UQB9	H31	P68431
67	IFIT3	O14879	IFIT5	Q13325
68	MP2K4	P45985	MK01	P28482
69	KTNB1	Q9BVA0	MSH3	P20585
70	SKP1	P63208	IKBA	P25963
71	SKP1	P63208	FBXL4	Q9UKA2
72	SKP1	P63208	FBX5	Q9UKT4
73	TNFL6	P48023	FADD	Q13158
74	TNFL6	P48023	SRC	P12931
75	TNFL6	P48023	SNX18	Q96RF0
76	TNFL6	P48023	TR10B	O14763
77	TNFL6	P48023	PK3CA	P42336
78	TNFL6	P48023	SNX33	Q8WV41
79	TNFL6	P48023	P85A	P27986
80	TNFL6	P48023	NCF1	P14598
81	ANXA4	P09525	DJB11	Q9UBS4
82	ANXA4	P09525	PADI4	Q9UM07
83	NDUAD	Q9P0J0	NDUA9	Q16795
84	NDUAD	Q9P0J0	NDUF4	Q9P032
85	NDUAD	Q9P0J0	ATP5J	P18859
86	NDUAD	Q9P0J0	UCRI	P47985
87	NDUAD	Q9P0J0	TIM23	O14925
88	NDUAD	Q9P0J0	ADT3	P12236
89	NDUAD	Q9P0J0	YMEL1	Q96TA2
90	NDUAD	Q9P0J0	ACAD9	Q9H845
91	NDUAD	Q9P0J0	ECSIT	Q9BQ95
92	NDUAD	Q9P0J0	NDUV3	P56181
93	NDUAD	Q9P0J0	NDUA3	O95167
94	NDUAD	Q9P0J0	NDUS6	O75380
95	NDUAD	Q9P0J0	NDUS8	O00217
96	NDUAD	Q9P0J0	NDUAC	Q9UI09
97	NDUAD	Q9P0J0	NDUS3	O75489
98	STAT3	P40763	P85B	O00459
99	STAT3	P40763	TWST1	Q15672
100	STAT3	P40763	CBL	P22681
101	STAT3	P40763	MYOD1	P15172
102	STAT3	P40763	MK01	P28482
103	STAT3	P40763	EP300	Q09472
104	BCCIP	Q9P287	GRWD1	Q9BQ67

Table V  
TABLE GO SEED GENES

	name	p-value	Z-score	Combined scores
1	phosphorylation of RNA polymerase II C-terminal domain (GO:0070816)	4.63401377456252E-09	-5.8422783646	112.1124012979
2	positive regulation of transcription from RNA polymerase II promoter (GO:0045944)	1.07612517030482E-05	-7.6742730333	87.7902966967
3	protein phosphorylation (GO:0006468)	1.02313139285069E-07	-4.8796598132	78.5392359548
4	positive regulation of transcription from RNA polymerase II promoter in response to nitrogen starvation (GO:0036278)	2.07862541206201E-05	-6.5201399981	70.2950549436
5	negative regulation of transcription from RNA polymerase II promoter (GO:0000122)	5.63097445662691E-05	-7.114962448	69.6173671915
6	JUN phosphorylation (GO:0007258)	2.13796852558072E-07	-4.4346214806	68.107979056
7	protein phosphorylation involved in mitotic spindle assembly (GO:1990801)	1.87778205182594E-07	-4.3834643142	67.8911142801
8	protein autophosphorylation (GO:0046777)	4.94613140111728E-06	-4.7497270292	58.026963046
9	peptidyl-tyrosine phosphorylation (GO:0018108)	2.53536426470294E-06	-4.4828395717	57.7621644755
10	positive regulation of pseudohyphal growth by positive regulation of transcription from RNA polymerase II promoter (GO:1900461)	0.0001202605	-6.3653693336	57.4528730224

Table VI  
TABLE PATHWAY SEED GENES

	name	p-value	Z-score	Combined scores
1	Pathways in cancer_Homo sapiens_hsa05200	9.31990332970234E-08	-2.1097791489	34.1542198539
2	Pancreatic cancer_Homo sapiens_hsa05212	9.36134966954956E-07	-1.9738643325	27.4002099207
3	Hepatitis B_Homo sapiens_hsa05161	2.6418518011837E-06	-1.8853872229	24.2159708956
4	MAPK signaling pathway_Homo sapiens_hsa04010	5.35937831060745E-06	-1.921408818	23.3194904951
5	Influenza A_Homo sapiens_hsa05164	7.51016066732915E-06	-1.8657376661	22.014312058
6	Proteoglycans in cancer_Homo sapiens_hsa05205	0.000017482	-1.8580379718	20.3535754896
7	Chagas disease (American trypanosomiasis)_Homo sapiens_hsa05142	8.91394059299544E-06	-1.6973464052	19.7365643322
8	Viral carcinogenesis_Homo sapiens_hsa05203	1.84785833752542E-05	-1.8042758285	19.6646184922
9	Osteoclast differentiation_Homo sapiens_hsa04380	2.83175828478206E-05	-1.6570939665	17.3531338261
10	HTLV-I infection_Homo sapiens_hsa05166	6.68505344817913E-05	-1.7181794317	16.5169469496

Table VII  
TABLE GO UNION INTERACTOME GENES

	name	p-value	Z-score	Combined scores
1	protein polyubiquitination (GO:0000209)	3.89716131252174E-61	-4.4654733514	621.1359216247
2	DNA damage induced protein phosphorylation (GO:0006975)	2.09211195660224E-49	-5.1060077824	572.3247303187
3	protein autophosphorylation (GO:0046777)	8.20040724969569E-51	-4.7985776644	553.4087138117
4	negative regulation of protein kinase activity by protein phosphorylation (GO:0100002)	2.93477942878535E-51	-4.5875568575	533.7861298223
5	protein polyubiquitination involved in nucleus-associated proteasomal ubiquitin-dependent protein catabolic process (GO:1901044)	1.37139823372652E-58	-3.9433084089	525.3831648925
6	protein ubiquitination involved in ubiquitin-dependent protein catabolic process (GO:0042787)	8.45735940621878E-52	-4.372501393	514.203487136
7	I-kappaB phosphorylation (GO:0007252)	6.79016439270026E-48	-4.5971421846	499.2892194785
8	protein phosphorylation (GO:0006468)	2.15948560920604E-45	-4.8331834519	497.0758043837
9	peptidyl-tyrosine phosphorylation (GO:0018108)	4.75752658067187E-48	-4.5114783348	491.5903338678
10	protein phosphorylation involved in double-strand break repair via nonhomologous end joining (GO:1990804)	1.10062993278406E-46	-4.3621002145	461.6106688222



Table VIII  
TABLE PATHWAY UNION INTERACTOME GENES

	name	p-value	Z-score	Combined scores
1	Pathways in cancer_Homo sapiens_hsa05200	3.8711317937372E-67	-2.1097791489	322.6266981199
2	PI3K-Akt signaling pathway_Homo sapiens_hsa04151	2.88557786564708E-57	-2.0411124312	265.7275800213
3	Focal adhesion_Homo sapiens_hsa04510	3.21769790586913E-57	-1.8717166077	243.4704334055
4	Epstein-Barr virus infection_Homo sapiens_hsa05169	3.81279867318203E-50	-1.8739385463	213.2371371782
5	HTLV-I infection_Homo sapiens_hsa05166	4.06216501311762E-47	-1.8180992957	194.2089727586
6	Regulation of actin cytoskeleton_Homo sapiens_hsa04810	3.45496212779622E-43	-1.7725193541	173.3016057246
7	Ubiquitin mediated proteolysis_Homo sapiens_hsa04120	2.51131807025633E-47	-1.5498977888	166.3051046931
8	Chemokine signaling pathway_Homo sapiens_hsa04062	1.30341051874194E-41	-1.7206600668	161.9846671005
9	Endocytosis_Homo sapiens_hsa04144	1.86276507391408E-38	-1.7406520184	151.2211833638
10	Cell cycle_Homo sapiens_hsa04110	3.4378367915527E-42	-1.5562396069	148.5800023373

Table IX  
TABLE GO INTERSECTION INTERACTOME GENES

	name	p-value	Z-score	Combined scores
1	positive regulation of DNA repair by positive regulation of transcription from RNA polymerase II promoter (GO:1900404)	0.000080164	-6.5547763617	61.8209546233
2	positive regulation of transcription from RNA polymerase II promoter (GO:0045944)	0.0003513477	-7.5627048815	60.1517450293
3	chromatin-mediated maintenance of transcription (GO:0048096)	7.68362889417625E-05	-5.5192835016	52.2887730279
4	negative regulation of apoptotic process (GO:0043066)	7.41030696629734E-05	-5.3982103827	51.3372700866
5	canonical Wnt signaling pathway involved in negative regulation of apoptotic process (GO:0044336)	5.25561259088494E-05	-5.0725774554	49.9832957838
6	positive regulation of telomeric RNA transcription from RNA pol II promoter (GO:1901582)	0.0008334223	-6.2522565886	44.3283121603
7	positive regulation of mating type switching by positive regulation of transcription from RNA polymerase II promoter (GO:1900423)	0.0008334223	-6.2508905336	44.3186268709
8	positive regulation of transcription from RNA polymerase II promoter involved in stress response to cadmium ion (GO:1902901)	0.0008334223	-6.2399389475	44.240980453
9	positive regulation of transcription from RNA polymerase II promoter by histone modification (GO:1903757)	0.0008334223	-6.2372913549	44.222209101
10	positive regulation of pseudohyphal growth by positive regulation of transcription from RNA polymerase II promoter (GO:1900461)	0.0008334223	-6.2362620639	44.2149114587

Table X  
TABLE PATHWAY INTERSECTION INTERACTOME GENES

	name	p-value	Z-score	Combined scores
1	Hepatitis B_Homo sapiens_hsa05161	1.55352121304851E-07	-1.9363651286	30.3575028435
2	Pathways in cancer_Homo sapiens_hsa05200	2.83137462216903E-06	-2.0830857456	26.6108959452
3	Herpes simplex infection_Homo sapiens_hsa05168	1.09939173071666E-05	-1.7296779688	19.7497543435
4	Prolactin signaling pathway_Homo sapiens_hsa04917	0.000056854	-1.8745841591	18.3241051173
5	Leishmaniasis_Homo sapiens_hsa05140	6.07588904885948E-05	-1.7442630435	16.9343471974
6	HTLV-I infection_Homo sapiens_hsa05166	0.0001165908	-1.7681393637	16.0137559724
7	Osteoclast differentiation_Homo sapiens_hsa04380	0.000109779	-1.7347402159	15.8156984131
8	Hepatitis C_Homo sapiens_hsa05160	0.0001144371	-1.6837978025	15.2812817341
9	Jak-STAT signaling pathway_Homo sapiens_hsa04630	0.000292006	-1.7273717903	14.0586235087
10	Pancreatic cancer_Homo sapiens_hsa05212	0.0005496337	-1.6782809247	12.5976104724