

DATABASE

Open Access

Fungal Secretome Database: Integrated platform for annotation of fungal secretomes

Jaeyoung Choi^{1,2,3}, Jongsun Park^{1,2,3,4}, Donghan Kim^{1,2,3}, Kyongyong Jung^{1,2,3}, Seogchan Kang⁶, Yong-Hwan Lee^{1,2,3,4,5*}

Abstract

Background: Fungi secrete various proteins that have diverse functions. Prediction of secretory proteins using only one program is unsatisfactory. To enhance prediction accuracy, we constructed Fungal Secretome Database (FSD).

Description: A three-layer hierarchical identification rule based on nine prediction programs was used to identify putative secretory proteins in 158 fungal/oomycete genomes (208,883 proteins, 15.21% of the total proteome). The presence of putative effectors containing known host targeting signals such as RXLX [EDQ] and RXLR was investigated, presenting the degree of bias along with the species. The FSD's user-friendly interface provides summaries of prediction results and diverse web-based analysis functions through Favorite, a personalized repository.

Conclusions: The FSD can serve as an integrated platform supporting researches on secretory proteins in the fungal kingdom. All data and functions described in this study can be accessed on the FSD web site at <http://fsd.snu.ac.kr/>.

Background

The “secretome” refers to the collection of proteins that contain a signal peptide and are processed via the endoplasmic reticulum and Golgi apparatus before secretion [1]. In organisms from bacteria to humans, secretory proteins are common and perform diverse functions. These functions include immune system [2], roles as neurotransmitters in the nervous system [3], roles as hormones/pheromones [4], acquisition of nutrients [5-7], building and remodeling of cell walls [8], signaling and environmental sensing [9], and competition with other organisms [10-13]. Some secretory proteins in pathogens function as effectors that manipulate and/or destroy host cells with special signatures. In *Plasmodium* and *Phytophthora* species, effectors carry the RXLX [EDQ] or RXLR motifs as host targeting signals [11-13].

With the aid of advanced genome sequencing technologies [14], the rapid increase of sequenced fungal genomes offers many opportunities to study the function and evolution of secretory proteins at the genome level [15,16]. The Comparative Fungal Genomics

Platform (CFGP; <http://cfgp.snu.ac.kr/>) [16] now archives 235 genomes from 120 fungal/oomycete species. The accurate prediction of secretory proteins in sequenced genomes is the key to realizing such opportunities.

The widely used SignalP 3.0 program [17] detected 89.81% of the 2,512 experimentally verified sequences in SPdb [18], a database containing proteins with signal peptides. To improve the accuracy of prediction, we built a hierarchical identification pipeline based on nine prediction programs (Table 1). Through this pipeline, putative secretory proteins, including pathogen effectors, encoded by 158 fungal and oomycete genomes were identified. The Fungal Secretome Database (FSD; <http://fsd.snu.ac.kr/>) was established to support not only the archiving of fungal secretory proteins but also the management and use of the resulting data. The FSD also has a user-friendly web interface and offers several data analysis functions via Favorite, a personalized data repository implemented in the CFGP (<http://cfgp.snu.ac.kr/>) [16].

Construction and content

Evaluation of the pipeline for predicting secretory proteins

To evaluate the capabilities of four programs SignalP 3.0 [17], SigCleave [19], SigPred [20], and RPSP [21] for

* Correspondence: yonglee@snu.ac.kr

¹Fungal Bioinformatics Laboratory, Seoul National University, Seoul 151-921, Korea

Table 1 List of prediction programs used in FSD

Prediction Program	Description	Ref
SignalP 3.0	A program to predict whether a protein has the signal peptidase site I or not	[17]
SigCleave	A program to predict whether a protein has signal peptides or not	[19]
SigPred	A program to predict whether a protein has signal peptides or not	[20]
RPSP	A program to predict whether a protein has signal peptides or not	[21]
TMHMM 2.0c	A program to predict whether a protein has trans-membrane helix(es) or not	[26]
TargetP 1.1b	A program to predict a site where a protein probably resides	[23]
PSort II	A program to predict a site where a protein probably resides	[22]
SecretomeP 1.0f	A program to predict whether a protein is secreted by non-classical pathways or not	[25]
predictNLS	A program to predict whether a protein has nuclear localization signal or not	[28]

predicting signal peptides, we analyzed the secretory proteins collected in SPdb [18]. SignalP 3.0 identified 89.81% of 2,512 proteins; while adding the other three programs, in combination, 87.50% of the proteins, which were not predicted by SignalP 3.0, were identified. The remaining proteins (1.31% of 2,512 proteins) were investigated by using two programs that predicted subcellular localization: PSort II [22] and TargetP 1.1b [23]. We found that 34.38% of the proteins were predicted to be extracellular proteins, increasing the coverage to 99.16%. For the 1,093 characterized fungal/oomycete secretory proteins (Table 2), the combinatory pipeline raised the

prediction coverage from 75.30% to 84.17% in comparison to SignalP 3.0. In addition, 98.14% of 24,921 experimentally unverified sequences in the SPdb were predicted as secretory proteins by the pipeline, while SignalP 3.0 caught 80.22% of them as positive. To assess robustness of the pipeline with non-secretory proteins, we prepared yeast proteins localized in cytosol, endoplasmic reticulum, nucleus, or mitochondrion [24]. When the 1,955 proteins were subjected to the FSD pipeline and SignalP 3.0, the numbers of false positives were almost same (84 and 82, respectively). Together, these results suggest that this ensemble approach could

Table 2 List of references and annotation results of characterized fungal secretory proteins

Title	Total Identified Proteins	Class SP	Class SP ³	Class SL	Putative Secretome	Ref
Crucial Role of Antioxidant Proteins and Hydrolytic Enzymes in Pathogenicity of <i>Penicillium expansum</i> : Analysis Based on Proteomics Approach (Secretory)	21	5	1	0	6	[43]
Crucial Role of Antioxidant Proteins and Hydrolytic Enzymes in Pathogenicity of <i>Penicillium expansum</i> : Analysis Based on Proteomics Approach (Non-secretory)	21	1	2	0	3	[43]
The <i>Phanerochaete chrysosporium</i> secretome: Database predictions and initial mass spectrometry peptide identifications in cellulose-grown medium	49	25	5	0	30	[44]
An analysis of the <i>Candida albicans</i> genome database for soluble secreted proteins using computer-based prediction algorithms (Secretory)	46	28	19	2	49	[45]
An analysis of the <i>Candida albicans</i> genome database for soluble secreted proteins using computer-based prediction algorithms (Non-secretory)	45	0	5	1	6	[45]
The secretome of the maize pathogen <i>Ustilago maydis</i> (Without known functions)	386	352	18	10	380	[46]
The secretome of the maize pathogen <i>Ustilago maydis</i> (With known functions)	168	147	15	5	167	[46]
A Catalogue of the Effector Secretome of Plant Pathogenic Oomycetes	25	22	1	0	23	[11]
Fungal degradation of wood: initial proteomic analysis of extra cellular proteins of <i>Phanerochaete chrysosporium</i> grown on oak substrate	11	8	0	0	8	[47]
Comparative proteomics of extracellular proteins in vitro and in planta from the pathogenic fungus <i>Fusarium graminearum</i>	120	63	8	0	71	[48]
Expression analysis of extracellular proteins from <i>Phanerochaete chrysosporium</i> grown on different liquid and solid substrates	27	16	4	0	20	[49]
Dandruff-associated <i>Malassezia</i> genomes reveal convergent and divergent virulence traits shared with plant and human fungal pathogens	34	28	0	0	28	[50]
Adaptive Evolution Has Targeted the C-Terminal Domain of the RXLR Effectors of Plant Pathogenic Oomycetes	79	79	0	0	79	[41]
Genome, transcriptome, and secretome analysis of wood decay fungus <i>Postia placenta</i> supports unique mechanisms of lignocellulose conversion.	47	29	3	1	33	[51]
Host-Microbe Interactions: Shaping the Evolution of the Plant Immune Response	14	12	0	1	13	[52]
Total	1,093	815	81	20	916	-

compensate for some of the weaknesses of individual programs, resulting in more robust predictions. Additionally, SecretomeP 1.0f [25], which can predict non-classical secretory proteins, was integrated into the FSD.

The FSD contains an identification pipeline that sequentially analyzes proteomes of interest using i) SignalP 3.0; ii) a combination of SigCleave, SigPred, and RPSP to screen those proteins not considered positive by SignalP 3.0; and iii) PSort II and TargetP 1.1b to analyze the negatives from the previous step. Additionally, SecretomeP 1.0f was integrated to provide information related to non-classical secretory proteins. To eliminate potential false positives, we filtered proteins that i) contain more than one transmembrane helix predicted by TMHMM 2.0c [26] and/or ii) the endoplasmic reticulum retention signal ([KRHQSA]-[DENQ]-E-L;

classified as false-positive; Figure 1A) [27]. In addition, iii) nuclear proteins predicted by both predictNLS [28] and PSort II [22] and iv) mitochondrial proteins predicted by PSort II [22] as well as TargetP 1.1b [23] were eliminated because two subcellular localizations are not related to secretory proteins.

Following analysis via the pipeline, the resulting putative secretory proteins after removing potential false positives are divided into four classes: i) SP contains all proteins predicted by SignalP 3.0; ii) SP³ contains the proteins predicted by SigPred, SigCleave, or RPSP but not by SignalP 3.0; iii) SL contains the proteins predicted by PSort II and/or TargetP 1.1b but not by the first two steps; and iv) NS contains the proteins predicted by SecretomeP 1.0f but not by SignalP 3.0 (Figure 1A; Table 3).

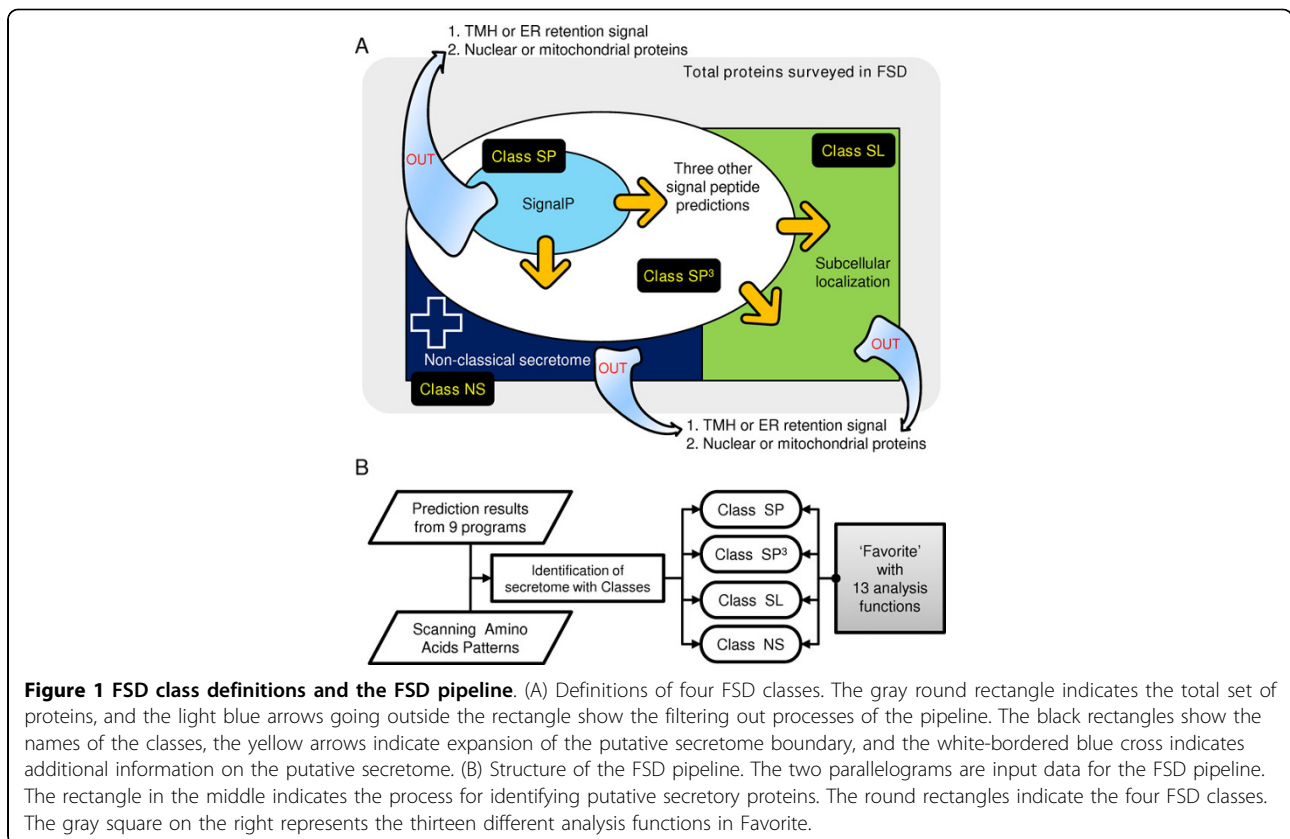


Table 3 Class definitions used in FSD

Class	Description*
Class SP	Proteins which are predicted by SignalP 3.0
Class SP ³	Proteins which are predicted by SigPred, SigCleave, or RPSP
Class SL	Proteins which are predicted by PSort II or TargetP 1.1b, but are not predicted by SignalP 3.0, SigPred, SigCleave, RPSP, or SecretomeP 1.0f
Class NS	Proteins which are predicted by SecretomeP 1.0f, but are not predicted by SignalP 3.0, SigPred, SigCleave, or RPSP

* Proteins as follows were removed from all four classes described in this table: proteins which i) contain more than one trans-membrane helices, ii) have ER retention signals, iii) predicted as mitochondrial proteins by PSort II and TargetP 1.1b, and iv) predicted as nuclear proteins by TargetP 1.1b and predictNLS.

Table 4 List and distribution of secretion-associated proteins of the fungal genomes belonging to the subphylum Pezizomycotina archived in FSD

Species	Size (Mb)	# of ORFs	Class SP	Class SP ³	Class SL	Putative Secretome	Ref
Fungi (Kingdom) ^a							
Ascomycota (Phylum)							
Pezizomycotina (Subphylum)							
<i>Aspergillus clavatus</i>	27.9	9,121	754	732	81	1,567	[53,54]
<i>Aspergillus flavus</i>	36.8	12,604	1,200	990	142	2,332	[55]
<i>Aspergillus fumigatus</i> A1163	29.2	9,929	807	878	67	1,752	[54]
<i>Aspergillus fumigatus</i> AF293	29.4	9,887	781	909	84	1,774	[56]
<i>Aspergillus nidulans</i>	30.1	10,568	922	877	96	1,895	[57]
<i>Aspergillus niger</i> ATCC1015	37.2	11,200	860	883	88	1,831	N
<i>Aspergillus niger</i> CBS513.88	34.0	14,086	1,142	1,320	154	2,616	[58]
<i>Aspergillus oryzae</i>	37.1	12,063	1,060	1,064	145	2,269	[59]
<i>Aspergillus terreus</i>	29.3	10,406	934	916	81	1,931	[53]
<i>Botrytis cinerea</i>	42.7	16,448	1,163	1,287	182	2,632	N
<i>Chaetomium globosum</i> ^b	34.9	11,124	1,121	923	99	2,143	N
<i>Coccidioides immitis</i> H538.4	27.7	10,663	548	957	80	1,585	N
<i>Coccidioides immitis</i> RMSCC 2394	28.8	10,408	575	920	66	1,561	N
<i>Coccidioides immitis</i> RMSCC 3703	27.6	10,465	539	892	65	1,496	N
<i>Coccidioides immitis</i> RS	28.9	10,457	476	855	102	1,433	[60]
<i>Coccidioides posadasii</i> RMSCC 3488	28.1	9,964	546	838	95	1,479	N
<i>Coccidioides posadasii</i> Silveira	27.5	10,125	558	869	91	1,518	N
<i>Cochliobolus heterostrophus</i> C5	34.9	9,633	932	725	83	1,740	N
<i>Cryphonectria parasitica</i>	43.9	11,184	1,040	951	93	2,084	N
<i>Fusarium graminearum</i> GZ3639 ^f	15.1	6,694	373	386	47	806	[61]
<i>Fusarium graminearum</i> MIPS	36.1	13,920	1,370	1,072	118	2,560	N
<i>Fusarium graminearum</i> PH-1	36.6	13,339	1,282	1,004	118	2,404	[61]
<i>Fusarium oxysporum</i>	61.4	17,608	1,613	1,297	147	3,057	N
<i>Fusarium solani</i>	51.3	15,707	1,381	1,242	155	2,778	[62]
<i>Fusarium verticillioides</i>	41.9	14,199	1,347	1,071	116	2,534	N
<i>Histoplasma capsulatum</i> G186AR	29.9	7,454	357	578	96	1,031	N
<i>Histoplasma capsulatum</i> G217B	41.3	8,038	393	583	103	1,079	N
<i>Histoplasma capsulatum</i> H143	39.0	9,547	468	842	87	1,397	N
<i>Histoplasma capsulatum</i> H88	37.9	9,445	492	832	99	1,423	N
<i>Histoplasma capsulatum</i> Nam1	33.0	9,349	398	736	79	1,213	[60]
<i>Magnaporthe oryzae</i>	41.7	11,069	1,573	833	64	2,470	[63]
<i>Microsporum canis</i>	23.3	8,777	564	702	88	1,354	N
<i>Microsporum gypseum</i>	23.3	8,876	629	669	52	1,350	N
<i>Mycosphaerella fijiensis</i>	73.4	10,327	770	778	81	1,629	N
<i>Mycosphaerella graminicola</i>	41.9	11,395	979	913	81	1,973	N
<i>Neosartorya fischeri</i> ^b	32.6	10,403	959	818	84	1,861	[54]
<i>Neurospora crassa</i>	39.2	9,842	817	788	61	1,666	[64]
<i>Neurospora crassa</i> MIPS	34.2	9,572	788	749	78	1,615	N
<i>Neurospora discretadiscrete</i>	37.3	9,948	823	800	88	1,711	N
<i>Neurospora tetrasperma</i>	37.8	10,640	849	895	73	1,817	N
<i>Paracoccidioides brasiliensis</i> Pb01	33.0	9,136	402	808	71	1,281	N
<i>Paracoccidioides brasiliensis</i> Pb03	29.1	9,264	470	823	92	1,385	N
<i>Paracoccidioides brasiliensis</i> Pb18	30.0	8,741	425	743	55	1,223	N
<i>Penicillium chrysogenum</i>	32.2	12,791	947	1,008	127	2,082	[65]
<i>Penicillium marneffeii</i>	28.6	10,638	713	792	109	1,614	N
<i>Podospora anserina</i>	35.7	10,596	1,127	893	124	2,144	[66]
<i>Pyrenophora tritici-repentis</i>	38.0	12,169	1,228	912	123	2,263	N

Table 4: List and distribution of secretion-associated proteins of the fungal genomes belonging to the subphylum Pezizomycotina archived in FSD (Continued)

<i>Sclerotinia sclerotiorum</i>	38.3	14,522	971	1,109	147	2,227	N
<i>Sporotrichum thermophile</i>	38.7	8,806	697	658	66	1,421	N
<i>Stagonospora nodorum</i>	37.2	15,983	1,511	1,309	142	2,962	[67]
<i>Talaromyces stipitatus</i>	35.7	13,252	748	1,116	114	1,978	N
<i>Thielavia terrestris</i>	37.0	9,815	877	855	67	1,799	N
<i>Trichoderma atroviride</i>	36.1	11,100	907	935	86	1,928	N
<i>Trichoderma reesei</i>	33.5	9,129	738	766	70	1,574	[68]
<i>Trichoderma virens</i> GV29-8	38.8	11,643	933	1,009	93	2,035	N
<i>Trichophyton equinum</i>	24.2	8,576	571	699	69	1,339	N
<i>Uncinocarpus reesii</i>	22.3	7,798	485	626	64	1,175	[60]
<i>Verticillium albo-atrum</i> VaMs. 102	32.9	10,239	1,074	815	73	1,962	N
<i>Verticillium dahliae</i> VdLs. 17	33.9	10,575	1,157	861	77	2,095	N
Total	2,059.4	641,257	50,164	52,111	5,578	107,853	-

^a Taxonomy based on [69]

^b Insufficient exon/intron information

^c Incomplete coverage of genome information

System structure of the FSD

To improve the expandability and flexibility of the FSD, we adopted a three-layer structure (i.e., data warehouse, analysis pipeline, and user interface) in its design. The data warehouse was established using the standardized genome warehouse managed by the CFGP (<http://cfgp.snu.ac.kr/>)[16] that has been used in various bioinformatics systems [15,29-35]. The pipeline layer was built with a series of Perl programs.

In addition to the prediction programs described above, ChloroP 1.1 as well as hydropathy plots [36] were included in the FSD to provide additional information on secretory proteins. Whenever new fungal genomes become available, the automated pipeline classifies them based on the predictions of nine programs, thus keeping the FSD current (Figure 1B).

MySQL 5.0.67 and PHP 5.2.9 were used to maintain database and to develop web-based user interfaces that present complex information intuitively. Web pages were serviced through Apache 2.2.11. Favorite, a personal data repository used in the CFGP (<http://cfgp.snu.ac.kr/>)[16], was integrated to provide thirteen functions for further analyses.

Utility and Discussion

Discussion

Secretory proteins in 158 fungal/oomycete genomes

To survey the genome-wide distribution of secretory proteins in fungi and oomycetes, we used the pipeline to analyze all predicted proteins encoded by 158 fungal/oomycete genomes. Of the 1,373,444 open reading frames (ORFs) analyzed, 92,926 (6.77%), 103,224 (7.52%), and 12,733 (0.93%) proteins belonged to classes SP, SP³, and SL, respectively (Table 4, 5, and 6). In total, 208,883 ORFs (15.21%) were denoted putative secretory

proteins. The proteins belonging to class NS were not included in the putative secretome because they represented more than 40% of whole proteome.

To determine the phylum-level distribution of classes SP, SP³, and SL within fungi, we investigated the proportions of the three classes among subphyla (Figure 2). Class SP³ was the largest, class SP was a little smaller, and the class SL was much smaller; this was consistent over every subphylum. Only in *Plasmodium* species, oomycetes, and the kingdom Metazoa class SP was dominant. Class SL did not exceed 2.10% of the whole genome, except in *Plasmodium* species (4.52%). *Plasmodium* species also showed the lowest variance among the three classes, which may reflect signal peptide-independent types of secretory proteins such as vacuolar transport signals (VTSs) [12]. These results may be partially affected by the composition of the training data for each prediction program and inherent features of each algorithm.

The phylum Basidiomycota had a larger proportion of secretory proteins (17.90%) than other fungal taxonomy such as the subphylum Mucoromycotina (11.99%) and the phyla Ascomycota (12.87%) and Microsporidia (15.10%). Within the phylum Ascomycota, the subphylum Pezizomycotina showed a higher portion of class SP (7.82%) than the subphyla Saccharomycotina and Taphrinomycotina (4.57% and 3.74%, respectively). When considered that subphylum Pezizomycotina contains many pathogenic fungi (47 of 59) compared with subphylum Saccharomycotina (11 of 65), the abundance of secretory proteins in the subphylum Pezizomycotina suggests that pathogens may have larger secretome than saprophytes in general. In fact, *Magnaporthe oryzae* and *Neurospora crassa*, a closely related pair of pathogen and non-pathogen supported by

Table 5 List and distribution of secretion-associated proteins of the fungal genomes belonging to the subphylum Saccharomycotina and Taphrinomycotina archived in FSD

Species	Size (Mb)	# of ORFs	Class SP	Class SP ³	Class SL	Putative Secretome	Ref
Fungi (Kingdom) ^a							
Ascomycota (Phylum)							
Saccharomycotina (Subphylum)							
<i>Candida albicans</i> SC5314	14.3	6,185	321	405	87	813	[70,71]
<i>Candida albicans</i> WO-1	14.5	6,160	310	385	78	773	[72]
<i>Candida dubliniensis</i> ^b	14.5	6,027	308	340	71	719	N
<i>Candida glabrata</i> CBS138	12.3	5,165	231	290	49	570	[73]
<i>Candida guilliermondii</i>	10.6	5,920	279	400	63	742	[72]
<i>Candida lusitanae</i>	12.1	5,941	310	482	50	842	[72]
<i>Candida parapsilosis</i>	13.1	5,733	308	321	83	712	[72]
<i>Candida tropicalis</i>	14.6	6,258	360	373	76	809	[72,74]
<i>Debaryomyces hansenii</i>	12.2	6,354	254	357	74	685	[73]
<i>Eremothecium gossypii</i>	8.8	4,717	204	333	35	572	[75]
<i>Kluyveromyces lactis</i>	10.7	5,327	248	304	60	612	[73]
<i>Kluyveromyces polysporus</i>	14.7	5,367	219	276	58	553	[76]
<i>Kluyveromyces waltii</i>	10.9	4,935	187	280	41	508	[77]
<i>Lodderomyces elongisporus</i>	15.5	5,802	253	351	50	654	[72]
<i>Pichia stipitis</i>	15.4	5,839	263	374	58	695	[78]
<i>Saccharomyces bayanus</i> 623-6C YM4911	11.9	4,966	200	275	44	519	[79]
<i>Saccharomyces bayanus</i> MCYC 623	11.5	9,385	663	767	141	1571	[80]
<i>Saccharomyces castellii</i>	11.4	4,677	177	240	46	463	[79]
<i>Saccharomyces cerevisiae</i> 273614N	12.5	5,354	223	261	51	535	[81]
<i>Saccharomyces cerevisiae</i> 322134S	12.5	5,382	224	290	53	567	[81]
<i>Saccharomyces cerevisiae</i> 378604X	12.5	5,400	232	267	53	552	[81]
<i>Saccharomyces cerevisiae</i> AWRI1631	11.2	5,451	220	364	63	647	N
<i>Saccharomyces cerevisiae</i> BC187	12.5	5,332	226	263	47	536	[81]
<i>Saccharomyces cerevisiae</i> DBVPG1106	12.5	5,318	225	253	52	530	[81]
<i>Saccharomyces cerevisiae</i> DBVPG1373	12.4	5,349	229	260	48	537	[81]
<i>Saccharomyces cerevisiae</i> DBVPG1788	12.4	5,347	227	263	46	536	[81]
<i>Saccharomyces cerevisiae</i> DBVPG1853	12.5	5,359	224	265	51	540	[81]
<i>Saccharomyces cerevisiae</i> DBVPG6040	12.6	5,364	221	271	50	542	[81]
<i>Saccharomyces cerevisiae</i> DBVPG6044	12.5	5,890	224	268	48	540	[81]
<i>Saccharomyces cerevisiae</i> DBVPG6765	12.2	5,377	230	263	48	541	[81]
<i>Saccharomyces cerevisiae</i> K11	12.5	5,375	228	270	52	550	[81]
<i>Saccharomyces cerevisiae</i> L_1374	12.4	5,346	225	264	55	544	[81]
<i>Saccharomyces cerevisiae</i> L_1528	12.4	5,346	227	258	48	533	[81]
<i>Saccharomyces cerevisiae</i> M22	10.8	6,755	249	399	62	710	[82]
<i>Saccharomyces cerevisiae</i> NCYC110	12.5	5,408	226	264	57	547	[81]
<i>Saccharomyces cerevisiae</i> NCYC361	12.6	5,360	228	261	49	538	[81]
<i>Saccharomyces cerevisiae</i> RM11-1a	11.7	5,696	264	283	63	610	N
<i>Saccharomyces cerevisiae</i> S288C	12.2	6,692	394	425	99	918	[83]
<i>Saccharomyces cerevisiae</i> SK1	12.4	5,433	233	269	55	557	[81]
<i>Saccharomyces cerevisiae</i> UWOPS03_461_4	12.6	5,329	218	268	51	537	[81]
<i>Saccharomyces cerevisiae</i> UWOPS05_217_3	12.6	5,350	217	264	47	528	[81]
<i>Saccharomyces cerevisiae</i> UWOPS05_227_2	12.6	5,334	220	266	51	537	[81]
<i>Saccharomyces cerevisiae</i> UWOPS83_787_3	12.6	5,392	225	269	51	545	[81]
<i>Saccharomyces cerevisiae</i> UWOPS87_2421	12.6	5,368	226	266	56	548	[81]
<i>Saccharomyces cerevisiae</i> W303	12.4	5,467	237	271	52	560	[81]
<i>Saccharomyces cerevisiae</i> Y12	12.6	5,370	223	268	57	548	[81]
<i>Saccharomyces cerevisiae</i> Y55	12.3	5,415	239	262	60	561	[81]

Table 5: List and distribution of secretion-associated proteins of the fungal genomes belonging to the subphylum Saccharomycotina and Taphrinomycotina archived in FSD (Continued)

<i>Saccharomyces cerevisiae</i> Y9	12.6	5,377	223	271	49	543	[81]
<i>Saccharomyces cerevisiae</i> Yllc17_E5	12.5	5,376	227	265	47	539	[81]
<i>Saccharomyces cerevisiae</i> YJM789	12.0	5,903	293	303	59	655	[84]
<i>Saccharomyces cerevisiae</i> YJM975	12.4	5,341	223	255	45	523	[81]
<i>Saccharomyces cerevisiae</i> YJM978	12.4	5,353	224	258	47	529	[81]
<i>Saccharomyces cerevisiae</i> YJM981	12.5	5,351	224	256	54	534	[81]
<i>Saccharomyces cerevisiae</i> YPS128	12.4	5,364	230	269	54	553	[81]
<i>Saccharomyces cerevisiae</i> YPS163	10.7	6,648	229	368	67	664	[82]
<i>Saccharomyces cerevisiae</i> YPS606	12.5	5,354	224	270	51	545	[81]
<i>Saccharomyces cerevisiae</i> YS2	12.6	5,383	221	254	50	525	[81]
<i>Saccharomyces cerevisiae</i> YS4	12.5	5,398	215	267	54	536	[81]
<i>Saccharomyces cerevisiae</i> YS9	12.6	5,373	226	265	51	542	[81]
<i>Saccharomyces kluyveri</i>	11.0	2,968	120	180	29	329	[79]
<i>Saccharomyces kudriavzevii</i>	11.2	3,768	187	195	28	410	[79]
<i>Saccharomyces mikatae</i>	11.5	9,016	575	630	154	1359	[80]
<i>Saccharomyces mikatae</i> WashU	10.8	3,100	161	154	24	339	[79]
<i>Saccharomyces paradoxus</i>	11.9	8,939	581	615	138	1334	[80]
<i>Yarrowia lipolytica</i>	20.5	6,524	409	464	75	948	[73]
Taphrinomycotina (Subphylum)							
<i>Pneumocystis carinii</i> ^{b, c}	6.3	4,020	129	333	35	497	N
<i>Schizosaccharomyces japonicus</i>	11.3	5,172	207	312	25	544	N
<i>Schizosaccharomyces octosporus</i>	11.2	4,925	190	263	26	479	N
<i>Schizosaccharomyces pombe</i>	12.6	5,058	192	288	36	516	[85]
Total	853.1	383,828	17,389	21,403	3,937	42,729	-

^a Taxonomy based on [69]

^b Insufficient exon/intron information

^c Incomplete coverage of genome information

recent phylogenomic studies [37-39], contain 22.31% and 16.93% of secretory proteins, respectively. Moreover, the same tendency was found in comparison with 158 fungal/oomycete genomes archived in the FSD (pathogens and saprophytes showed 14.06% and 11.70%, respectively).

Effectors encoded by fungal/oomycete and *Plasmodium* genomes

Phytophthora species, a group that includes many important plant pathogens, uses a RXLR signal to secrete effectors to host cells [40]. RXLR effectors were tightly co-located with signal peptides predicted by the SignalP 3.0 with high confidence values (HMM and NN for 0.93 and 0.65, respectively) [41]. With the same conditions, we identified 734 putative RXLR effectors from three *Phytophthora* species, similar to a previous study [42]. However, 153 fungal genomes showed that only 0.04% of the total proteome contained this motif, suggesting that the use of RXLR for secretion is oomycete-specific.

The motivation of finding the RXLR pattern in oomycetes was the RXLX [EDQ] motif of the VTS in the malaria pathogen, *Plasmodium falciparum*. Once

P. falciparum invades the human erythrocyte, it secretes the proteins that carry the pentameric VTS of the RXLX [EDQ] motif from the parasitophorous vacuole to the host cytoplasm [12,13]. To determine how many VTSs could be detected by our pipeline, we investigated 217 proteins of *P. falciparum* [13]. Of these, 115 proteins (53.00%) were classified as secretory proteins, defined in the FSD by the RXLX [EDQ] motif. Comparing our result to that predicted by SignalP 3.0 alone (41 out of 217), we found that our pipeline demonstrated high fidelity in detecting proteins containing VTSs.

In class SP, the proportions of proteins possessing the RXLX [EDQ] but not the RXLR motif were 96.75%, 56.18%, and 93.21% in fungi, oomycetes, and *Plasmodium* species, respectively (Figure 3A). There were similar proportions of the RXLX [EDQ] motif in classes SP³ and SL across the three groups (Figure 3B and 3C). Taken together, these data show that the RXLR motif, with signal peptides predicted by SignalP 3.0, is oomycete-specific [41]. It is interesting that fungal genomes have significantly higher numbers of the RXLX [EDQ] motif than *Plasmodium* species (t-test based on amino acid frequency in each genome; $P = 2.2e^{-16}$), suggesting

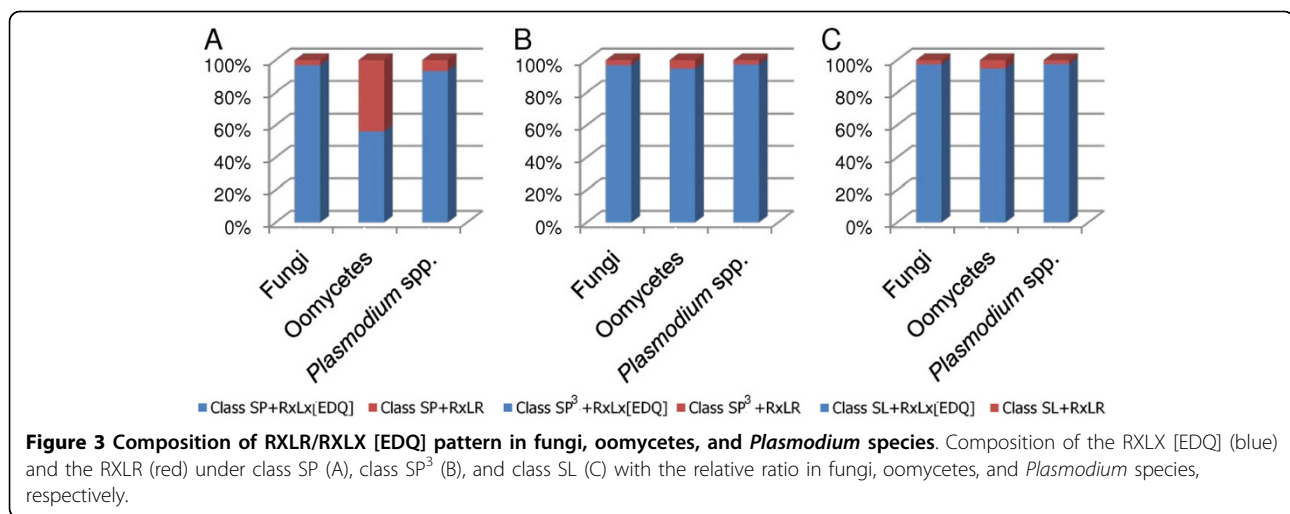
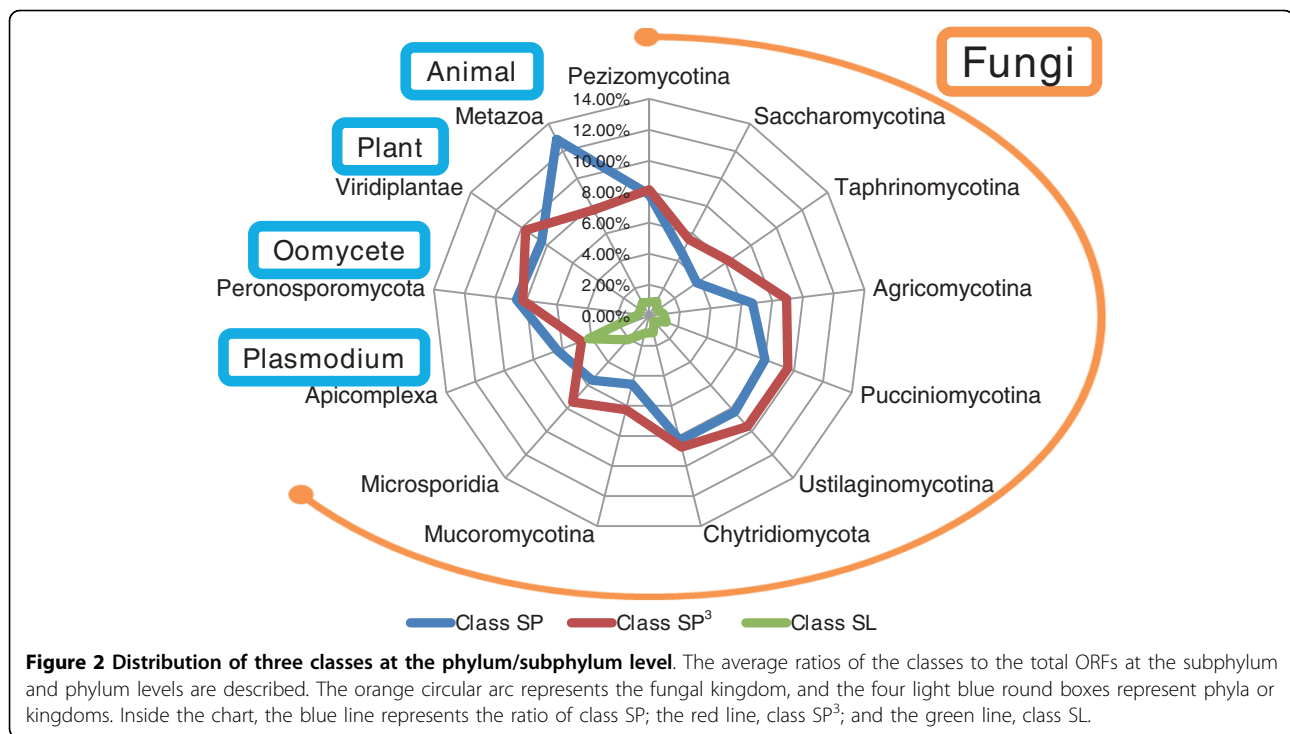
Table 6 List and distribution of secretion-associated proteins of the fungal genomes belonging to the phyla Basidiomycota, Chytridiomycota, and Microsporidia, the subphylum Mucoromycotina, and the phylum Peronosporomycota (oomycetes) archived in FSD

Species	Size (Mb)	# of ORFs	Class SP	Class SP ³	Class SL	Putative Secretome	Ref
Fungi (Kingdom) ^a							
Basidiomycota (Phylum)							
Agricomycotina (Subphylum)							
<i>Coprinus cinereus</i>	36.3	13,410	1,189	1,032	119	2,340	N
<i>Cryptococcus neoformans</i> Serotype A	18.9	6,980	377	549	56	982	N
<i>Cryptococcus neoformans</i> Serotype B	19.0	6,870	331	529	44	904	N
<i>Cryptococcus neoformans</i> Serotype D B-3501A	18.5	6,431	342	523	39	904	[86]
<i>Cryptococcus neoformans</i> Serotype D JEC21	19.1	6,475	344	541	38	923	[86]
<i>Laccaria bicolor</i>	64.9	20,614	1,190	2,024	256	3,470	[87]
<i>Moniliophthora perniciosa</i>	26.7	13,560	843	1,127	126	2,096	N
<i>Phanerochaete chrysosporium</i>	35.1	10,048	793	933	83	1,809	[88]
<i>Pleurotus ostreatus</i>	34.3	11,603	1,039	1,058	106	2,203	N
<i>Postia placenta</i>	90.9	17,173	1,057	1,808	202	3,067	[51]
<i>Schizophyllum commune</i>	38.5	13,181	975	1,175	119	2,269	N
Pucciniomycotina (Subphylum)							
<i>Melampsora laricis-populina</i>	21.9	16,694	1305	1483	233	3,021	N
<i>Puccinia graminis</i>	88.7	20,567	1,931	2,020	230	4,181	N
<i>Sporobolomyces roseus</i>	21.2	5,536	187	592	43	822	N
Ustilaginomycotina (Subphylum)							
<i>Malassezia globosa</i>	9.0	4,286	211	378	37	626	[50]
<i>Ustilago maydis</i> 521	19.7	6,689	789	583	10	1382	[89]
<i>Ustilago maydis</i> FB1	19.3	6,950	481	717	34	1232	[89]
<i>Ustilago maydis</i> MIPS	19.7	6,787	574	687	34	1295	N
Chytridiomycota (Phylum)							
<i>Batrachochytrium dendrobatidis</i> JAM81	24.3	8,732	806	750	108	1,664	N
<i>Batrachochytrium dendrobatidis</i> JEL423	23.9	8,818	650	785	91	1,526	N
Mucoromycotina (Subphylum <i>incertae sedis</i>)							
<i>Mucor circinelloides</i>	36.6	10,930	580	623	83	1286	N
<i>Phycomyces blakesleeanus</i>	55.9	14,792	642	1,085	221	1,948	N
<i>Rhizopus oryzae</i>	46.1	17,482	750	994	202	1,946	[90]
Microsporidia (Phylum)							
<i>Antonosporea locustae</i> ^b	6.1	2,606	166	208	62	436	N
<i>Encephalitozoon cuniculi</i>	2.5	1,996	90	135	34	259	[91]
Alveolata (Kingdom)							
Apicomplexa (Phylum)							
<i>Plasmodium berghei</i>	18.0	12,175	844	554	569	1,967	N
<i>Plasmodium chabaudi</i>	16.9	15,007	1,027	643	661	2,331	N
<i>Plasmodium falciparum</i> 3D7	21.0	5,387	212	283	267	762	[92]
<i>Plasmodium knowlesi</i>	23.5	5,103	305	280	81	666	N
Stramenopila (Kingdom)							
Peronosporomycota (Phylum)							
<i>Hyaloperonospora parasitica</i>	83.6	14,789	868	1,235	132	2,235	N
<i>Phytophthora capsici</i>	107.8	17,414	1,485	1,179	136	2,800	N
<i>Phytophthora infestans</i> ^b	228.5	22,658	1,668	1,923	153	3,744	[93]
<i>Phytophthora ramorum</i>	66.7	15,743	1,670	1,372	91	3,133	[94]
<i>Phytophthora sojae</i>	86.0	19,027	2,040	1,662	96	3,798	[94]
Total	1,449.1	386,513	27,761	31,470	4,796	64,027	-

^a Taxonomy based on [69]

^b Insufficient exon/intron information

^c Incomplete coverage of genome information



that the RXLX [EDQ] motif may be one of fungal-specific signatures of effectors.

Utility

FSD web interfaces

To support the browsing of the global patterns of archived data, the FSD prepares diverse charts and tables. For example, intersections of prediction results are summarized in a chart for each genome (Figure 4). Despite of the many programs, all prediction results for each protein are displayed on one page, allowing users to browse them easily (Figure 5).

The SNUGB interface (<http://genomebrowser.snu.ac.kr/>)[15] provides several fields: i) signal peptides predicted by four different programs; ii) effector patterns, such as RXLR and RXLX [EDQ]; iii) nucleotide localization signals predicted by predictNLS; iv) transmembrane helices predicted by TMHMM 2.0c; and v) hydropathy plots (Figure 6). The users can readily compare secretome-related information with diverse genomic contexts. **The personalized virtual space, Favorite, supports in-depth analyses in the FSD**

The FSD allows users to collect proteins of interest and save them into the Favorite, which provides thirteen

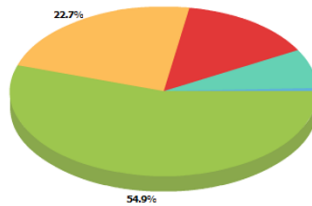
Detail Information on **Magnaporthe oryzae 6A**

Introduction Class SigPep Clvg Site AA Dist SubLoc Pattern

Class Distribution of 4 Classes

11,069 proteins are in **Magnaporthe oryzae 6A** genome.
 8,554 proteins have their classes.
 2,515 proteins were not included any class.

Distribution among classes



Class	# Proteins (Ratio)*
Class SP	1,573 (14.21 %)
Class SP ³	833 (7.53 %)
Class NS	6,084 (54.96 %)
Class SL	64 (0.58 %)

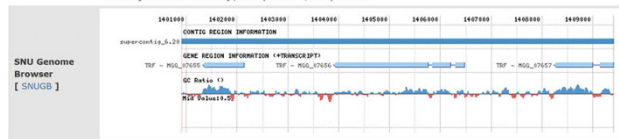
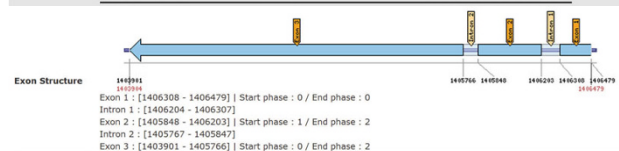
* The ratios in the table: (# Proteins in a Class) / (Total # Proteins in the genome)

Figure 4 Screenshot of genome-level analysis functions for an example fungal genome. This screenshot shows the ORF numbers and ratios of each class through the pie chart in the left and the table in the right. The numbers in the table provide links to the list of putative secretory proteins belonging to each group. This figure shows the result from *M. oryzae*.

Contig Browser

Basic Information

Class	Class SP
Locus Name	MGG_07656.6
Description	chitinase
Species Name	<i>Magnaporthe oryzae 70-15</i> [Magnaporthe oryzae 6A PROTEINS]
Amino Sequence	[797 aa]



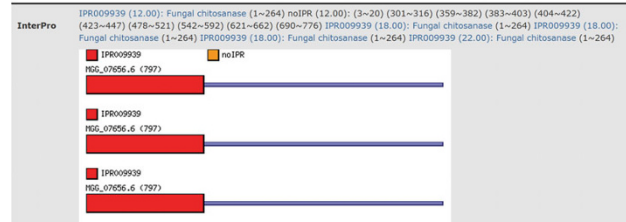
Signal Peptide

Method	NN Method : 21 ~ 22, Seq : TTA-RD	Position	: 9 ~ 21
Prediction	: 'Signal peptide'	Probability	: 8.105
Anchor	: 0	Signal Peptide	: ALSSSLFAATTA
Probability	: 0.885 (21~22)	Mature Peptide	: RDIPENKILRESIARGDCEKLLASFRSVSTDEKTSGYCDLLSAGV
Position	: 21 ~ 22		
Probability	: 0.792484700833211	Position	: 21 ~ 22
Secretion	: Yes	Signal Peptide	: NRSSIATLALSSSLFAATTA

Sub-cellular Localization

	PLANT network	Non-PLANT network		
TargetP	mTP Score : 0.1 SP Score : 0.556 Other Score : 0.044 Location : Secretory RC : 5 TP_LENGTH : 0	Length : 797 cTP Score : 0.425 mTP Score : 0.234 SP Score : 0.802 Other Score : 0.024 Location : Secretory RC : 3 TP_LENGTH : 21	PredictNLS	There is no predictNLS annotation data.
PSortIII	extracellular, including cell wall	44.4	Score : 0.449	
	nuclear	33.3	cTP : N	
	mitochondrial	22.2	CS Score : 1.026	
			cTP Length : 17	

Functional Annotation



GO Terms

There are no GO terms.

Other Predictions

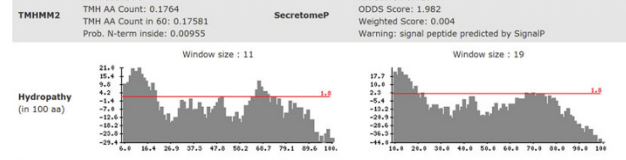


Figure 5 One page summary for a protein. The web page shows a one page summary of amino acid sequence, exon structure, and genome context via the SNUGB [15], along with 12 predictions, including signal peptides and subcellular localization.

functions: i) classes distribution of proteins; ii) comparisons of predicted signal peptides generated by the four programs; iii) distributions and lists of proteins with predicted signal peptide cleavage sites; iv) compositions of amino acids near the cleavage sites; v) analyses of subcellular localization predictions; vi) lists and ratios of proteins that have chloroplast transit peptides, as determined by ChloroP 1.1; vii) analyses of proteins detected

by SecretomeP 1.0f; viii) lists and distribution charts of proteins with trans-membrane helices, as predicted by TMHMM 2.0c; ix) hydropathy plots for proteins; x) analyses of proteins believed to be targeted to the nucleus of a host cell supported by predictNLS; xi) distributions and lists of proteins with a specific amino acid patterns; xii) lists of functional domains predicted by InterPro Scan; xiii) domain architecture of InterPro Scan (Figure



7). From these result pages, users can collect and store proteins in Favorite again, for further analyses. Additionally, Favorites created in the FSD can be shared with the CFGP (<http://cfgp.snu.ac.kr/>)[16], permitting users to use the 22 bioinformatics tools provided in the CFGP web site.

Conclusions

Given the availability of large number of fungal genomes and diverse prediction programs for secretory proteins, a three-layer classification rule was established and implemented in a web-based database, the FSD. With the aid of an automated pipeline, the FSD

classifies putative secretory proteins from 158 fungal/oomycetes genomes into four different classes, three of which are defined as the putative secretome. The proportion of fungal secretory proteins and host targeting signals varies considerably by species. It is interesting that fungal genomes have high proportions of the RXLX [EDQ] motif, characterized as host targeting signal in *Plasmodium* species. Summaries of the complex prediction results from twelve programs help users to readily access to the information provided by the FSD. Favorite, a personalized virtual space in the CFGP, serves thirteen different analysis tools for further in-depth analyses. Moreover, 22 bioinformatics tools

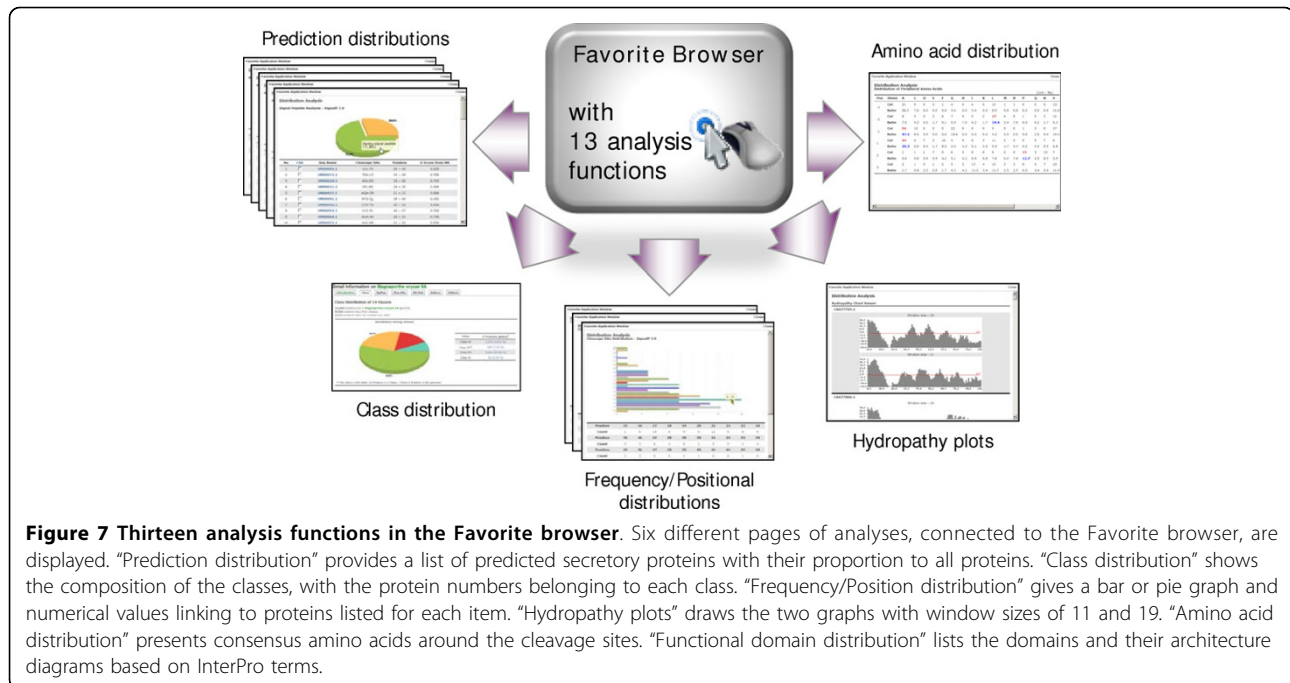


Figure 7 Thirteen analysis functions in the Favorite browser. Six different pages of analyses, connected to the Favorite browser, are displayed. "Prediction distribution" provides a list of predicted secretory proteins with their proportion to all proteins. "Class distribution" shows the composition of the classes, with the protein numbers belonging to each class. "Frequency/Position distribution" gives a bar or pie graph and numerical values linking to proteins listed for each item. "Hydropathy plots" draws the two graphs with window sizes of 11 and 19. "Amino acid distribution" presents consensus amino acids around the cleavage sites. "Functional domain distribution" lists the domains and their architecture diagrams based on InterPro terms.

provided by the CFGP can be utilized via the Favorite. Given these features, the FSD can serve as an integrated environment for studying secretory proteins in the fungal kingdom.

Availability and requirements

All data and functions described in this paper can be freely accessed through the FSD web site at <http://fsd.snu.ac.kr/>.

Acknowledgements

This work was supported by the National Research Foundation of Korea grants (2009-0063340 and 2009-0080161) and grants from the Biogreen21 (20080401-034-044-009-01-00), the TDPAF (309015-04-SB020), and the Crop Functional Genomics Center (2009K001198). JC is grateful for the graduate fellowship through the Brain Korea 21 Program.

Author details

¹Fungal Bioinformatics Laboratory, Seoul National University, Seoul 151-921, Korea. ²Department of Agricultural Biotechnology, Seoul National University, Seoul 151-921, Korea. ³Center for Fungal Pathogenesis, Seoul National University, Seoul 151-921, Korea. ⁴Center for Fungal Genetic Resources, Seoul National University, Seoul 151-921, Korea. ⁵Center for Agricultural Biomaterials, Seoul National University, Seoul 151-921, Korea. ⁶Department of Plant Pathology, The Pennsylvania State University, University Park, PA 16802, USA.

Authors' contributions

JC, JP, and YHL designed this project, JC and JP constructed the database and developed the pipeline with nine prediction programs. DK generated basic data from the twelve programs and JP, JC, and DK managed genome sequences for FSD. JC developed thirteen analysis functions of FSD. JC and JP constructed web-based interfaces. JC, JP, SK, and YHL wrote the manuscript. All the authors read and confirmed the manuscript.

Received: 16 May 2009

Accepted: 11 February 2010 Published: 11 February 2010

References

1. Lippincott-Schwartz J, Roberts TH, Hirschberg K: **Secretory protein trafficking and organelle dynamics in living cells.** *Annu Rev Cell Dev Biol* 2000, **16**:557-589.
2. Abbas KA, Lichtman HA, Pillai S: *Cellular and Molecular Immunology* Saunders, 6 2006.
3. Cho WJ, Jeremic A, Rognlien KT, Zhvania MG, Lazrshvili I, Tamar B, Jena BP: **Structure, isolation, composition and reconstitution of the neuronal fusion pore.** *Cell Biol Int* 2004, **28**(10):699-708.
4. Cho SJ, Jeftinija K, Glavaski A, Jeftinija S, Jena BP, Anderson LL: **Structure and dynamics of the fusion pores in live GH-secreting cells revealed using atomic force microscopy.** *Endocrinology* 2002, **143**(3):1144-1148.
5. Suarez MB, Sanz L, Chamorro MI, Rey M, Gonzalez FJ, Llobell A, Monte E: **Proteomic analysis of secreted proteins from *Trichoderma harzianum*: Identification of a fungal cell wall-induced aspartic protease.** *Fungal Genet Biol* 2005, **42**(11):924-934.
6. Van den Wymelenberg A, Mingos P, Sabat G, Martinez D, Aerts A, Salamov A, Grigoriev I, Shapiro H, Putnam N, Belinky P, Dosoretz C, Gaskell J, Kersten P, Cullen D: **Computational analysis of the *Phanerochaete chrysosporium* v2.0 genome database and mass spectrometry identification of peptides in ligninolytic cultures reveal complex mixtures of secreted proteins.** *Fungal Genet Biol* 2006, **43**(5):343-356.
7. Vinzant TB, Adney WS, Decker SR, Baker JO, Kinter MT, Sherman NE, Fox JW, Himmel ME: **Fingerprinting *Trichoderma reesei* hydrolases in a commercial cellulase preparation.** *Appl Biochem Biotechnol* 2001, **91**:93-99-107.
8. Lesage G, Bussey H: **Cell wall assembly in *Saccharomyces cerevisiae*.** *Microbiol Mol Biol Rev* 2006, **70**(2):317-343.
9. Waters CM, Bassler BL: **Quorum sensing: cell-to-cell communication in bacteria.** *Annu Rev Cell Dev Biol* 2005, **21**:319-346.
10. Cornelis GR, Van Gijsegem F: **Assembly and function of type III secretory systems.** *Annu Rev Microbiol* 2000, **54**:735-774.
11. Kamoun S: **A catalogue of the effector secretome of plant pathogenic oomycetes.** *Annu Rev Phytopathol* 2006, **44**:41-60.
12. Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, Lopez-Estrano C, Haldar K: **A host-targeting signal in virulence proteins reveals a secretome in malarial infection.** *Science* 2004, **306**(5703):1934-1937.
13. Marti M, Good RT, Rug M, Knuepfer E, Cowman AF: **Targeting malaria virulence and remodeling proteins to the host erythrocyte.** *Science* 2004, **306**(5703):1930-1933.

14. Bouws H, Wattenberg A, Zorn H: **Fungal secretomes—nature's toolbox for white biotechnology.** *Appl Microbiol Biotechnol* 2008, **80**(3):381-388.
15. Jung K, Park J, Choi J, Park B, Kim S, Ahn K, Choi J, Choi D, Kang S, Lee Y-H: **SNUGB: a versatile genome browser supporting comparative and functional fungal genomics.** *BMC Genomics* 2008, **9**:585.
16. Park J, Park B, Jung K, Jang S, Yu K, Choi J, Kong S, Kim S, Kim H, Kim JF, Blair JE, Lee K, Kang S, Lee YH: **CFGP: a web-based, comparative fungal genomics platform.** *Nucleic Acids Res* 2008, **36**:D562-571.
17. Bendtsen JD, Nielsen H, von Heijne G, Brunak S: **Improved prediction of signal peptides: SignalP 3.0.** *J Mol Biol* 2004, **340**(4):783-795.
18. Choo KH, Tan TW, Ranganathan S: **SPdb—a signal peptide database.** *BMC Bioinformatics* 2005, **6**:249.
19. Rice P, Longden I, Bleasby A: **EMBOSS: the European Molecular Biology Open Software Suite.** *Trends Genet* 2000, **16**(6):276-277.
20. Bradford JR: **Protein Design for Biopharmaceutical Development at GlaxoSmithKline: In silico Methods for Prediction of Signal Peptides and their Cleavage Sites, and Linear Epitopes.** *MRes thesis, The University of Leeds, Department of Biological Sciences* 2001.
21. Plewczynski D, Slabinski L, Ginalski K, Rychlewski L: **Prediction of signal peptides in protein sequences by neural networks.** *Acta Biochim Pol* 2008, **55**(2):261-267.
22. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K: **WoLF PSORT: protein localization predictor.** *Nucleic Acids Res* 2007, **35**(Web Server issue):W585-587.
23. Emanuelsson O, Nielsen H, Brunak S, von Heijne G: **Predicting subcellular localization of proteins based on their N-terminal amino acid sequence.** *J Mol Biol* 2000, **300**(4):1005-1016.
24. Kumar A, Agarwal S, Heyman JA, Matson S, Heidtman M, Piccirillo S, Umansky L, Drawid A, Jansen R, Liu Y, Cheung KH, Miller P, Gerstein M, Roeder GS, Snyder M: **Subcellular localization of the yeast proteome.** *Genes Dev* 2002, **16**(6):707-719.
25. Bendtsen JD, Jensen LJ, Blom N, Von Heijne G, Brunak S: **Feature-based prediction of non-classical and leaderless protein secretion.** *Protein Eng Des Sel* 2004, **17**(4):349-356.
26. Sonnhammer EL, von Heijne G, Krogh A: **A hidden Markov model for predicting transmembrane helices in protein sequences.** *Proc Int Conf Intell Syst Mol Biol* 1998, **6**:175-182.
27. Raykhel I, Alanen H, Salo K, Juvansuu J, Nguyen VD, Latva-Ranta M, Ruddock L: **A molecular specificity code for the three mammalian KDEL receptors.** *J Cell Biol* 2007, **179**(6):1193-1204.
28. Cokol M, Nair R, Rost B: **Finding nuclear localization signals.** *EMBO Rep* 2000, **1**(5):411-415.
29. Park J, Park J, Jang S, Kim S, Kong S, Choi J, Ahn K, Kim J, Lee S, Kim S, Park B, Jung K, Kim S, Kang S, Lee YH: **FTFD: an informatics pipeline supporting phylogenomic analysis of fungal transcription factors.** *Bioinformatics* 2008, **24**(7):1024-1025.
30. Jeon J, Park SY, Chi MH, Choi J, Park J, Rho HS, Kim S, Goh J, Yoo S, Choi J, Park JY, Yi M, Yang S, Kwon MJ, Han SS, Kim BR, Khang CH, Park B, Lim SE, Jung K, Kong S, Karunakaran M, Oh HS, Kim H, Kim S, Park J, Kang S, Choi WB, Kang S, Lee YH: **Genome-wide functional analysis of pathogenicity genes in the rice blast fungus.** *Nat Genet* 2007, **39**(4):561-565.
31. Park J, Lee S, Choi J, Ahn K, Park B, Park J, Kang S, Lee YH: **Fungal Cytochrome P450 Database.** *BMC Genomics* 2008, **9**(1):402.
32. Choi J, Park J, Jeon J, Chi MH, Goh J, Yoo SY, Park J, Jung K, Kim H, Park SY, Rho HS, Kim S, Kim BR, Han SS, Kang S, Lee YH: **Genome-wide analysis of T-DNA integration into the chromosomes of *Magnaporthe oryzae*.** *Mol Microbiol* 2007, **66**(2):371-382.
33. Lee W, Park J, Choi J, Jung K, Park B, Kim D, Lee J, Ahn K, Song W, Kang S, Lee YH, Lee S: **IMGD: an Integrated Platform Supporting Comparative Genomics and Phylogenetics of Insect Mitochondrial Genomes.** *BMC Genomics* 2009, **10**:148.
34. Park J, Park B, Veeraraghavan N, Jung K, Lee YH, Blair J, Geiser DM, Isard S, Mansfield MA, Nikolaeva E, Park SY, Russo J, Kim SH, Greene M, Ivors KL, Balci Y, Peiman M, Erwin DC, Coffey MD, Rossman A, Farr D, Cline E, Crünwald NJ, Luster DG, Schrandt J, Martin F, Ribeiro OK, Makalowska I, Kang S: **Phytophthora Database: A Forensic Database Supporting the Identification and Monitoring of *Phytophthora*.** *Plant disease* 2008, **92**(6):966-972.
35. Xi H, Park J, Ding G, Lee YH, Li Y: **SysPIMP: the web-based systematical platform for identifying human disease-related mutated sequences from mass spectrometry.** *Nucleic Acids Res* 2009, **37**:D913-920.
36. Kyte J, Doolittle RF: **A simple method for displaying the hydrophobic character of a protein.** *J Mol Biol* 1982, **157**(1):105-132.
37. Fitzpatrick DA, Logue ME, Stajich JE, Butler G: **A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis.** *BMC Evol Biol* 2006, **6**:99.
38. Wang H, Xu Z, Gao L, Hao B: **A fungal phylogeny based on 82 complete genomes using the composition vector method.** *BMC Evol Biol* 2009, **9**:195.
39. Soanes DM, Alam I, Cornell M, Wong HM, Hedeler C, Paton NW, Rattray M, Hubbard SJ, Oliver SG, Talbot NJ: **Comparative genome analysis of filamentous fungi reveals gene family expansions associated with fungal pathogenesis.** *PLoS One* 2008, **3**(6):e2300.
40. Whisson SC, Boevink PC, Moleleki L, Avrova AO, Morales JG, Gilroy EM, Armstrong MR, Grouffaud S, van West P, Chapman S, Hein I, Toth IK, Pritchard L, Birch PR: **A translocation signal for delivery of oomycete effector proteins into host plant cells.** *Nature* 2007, **450**(7166):115-118.
41. Win J, Morgan W, Bos J, Krasileva KV, Cano LM, Chaparro-Garcia A, Ammar R, Staskawicz BJ, Kamoun S: **Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes.** *Plant Cell* 2007, **19**(8):2349-2369.
42. Jiang RH, Tripathy S, Govers F, Tyler BM: **RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members.** *Proc Natl Acad Sci USA* 2008, **105**(12):4874-4879.
43. Qin G, Tian S, Chan Z, Li B: **Crucial role of antioxidant proteins and hydrolytic enzymes in pathogenicity of *Penicillium expansum*: analysis based on proteomics approach.** *Mol Cell Proteomics* 2007, **6**(3):425-438.
44. Wymelenberg AV, Sabat G, Martinez D, Rajangam AS, Teeri TT, Gaskell J, Kersten PJ, Cullen D: **The *Phanerochaete chrysosporium* secretome: database predictions and initial mass spectrometry peptide identifications in cellulose-grown medium.** *J Biotechnol* 2005, **118**(1):17-34.
45. Lee SA, Wormsley S, Kamoun S, Lee AF, Joiner K, Wong B: **An analysis of the *Candida albicans* genome database for soluble secreted proteins using computer-based prediction algorithms.** *Yeast* 2003, **20**(7):595-610.
46. Mueller O, Kahmann R, Aguilar G, Trejo-Aguilar B, Wu A, de Vries RP: **The secretome of the maize pathogen *Ustilago maydis*.** *Fungal Genet Biol* 2008, **45**(Suppl 1):S63-70.
47. Abbas A, Koc H, Liu F, Tien M: **Fungal degradation of wood: initial proteomic analysis of extracellular proteins of *Phanerochaete chrysosporium* grown on oak substrate.** *Curr Genet* 2005, **47**(1):49-56.
48. Paper JM, Scott-Craig JS, Adhikari ND, Cuomo CA, Walton JD: **Comparative proteomics of extracellular proteins in vitro and in planta from the pathogenic fungus *Fusarium graminearum*.** *Proteomics* 2007, **7**(17):3171-3183.
49. Sato S, Liu F, Koc H, Tien M: **Expression analysis of extracellular proteins from *Phanerochaete chrysosporium* grown on different liquid and solid substrates.** *Microbiology* 2007, **153**(Pt 9):3023-3033.
50. Xu J, Saunders CW, Hu P, Grant RA, Boekhout T, Kuramae EE, Kronstad JW, Deangelis YM, Reeder NL, Johnstone KR, Leland M, Fieno AM, Begley WM, Sun Y, Lacey MP, Chaudhary T, Keough T, Chu L, Sears R, Yuan B, Dawson TL Jr: **Dandruff-associated *Malassezia* genomes reveal convergent and divergent virulence traits shared with plant and human fungal pathogens.** *Proc Natl Acad Sci USA* 2007, **104**(47):18730-18735.
51. Martinez D, Challacombe J, Morgenstern I, Hibbett D, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Duenas FJ, Martinez AT, Kersten P, Hammel KE, Vanden Wymelenberg A, Gaskell J, Lindquist E, Sabat G, Bondurant SS, Larrondo LF, Canessa P, Vicuna R, Yadav J, Doddapaneni H, Subramanian V, Pisabarro AG, Lavin JL, Oguiza JA, Master E, Henrissat B, Coutinho PM, Harris P, Magnuson JK, et al: **Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion.** *Proc Natl Acad Sci USA* 2009, **106**(6):1954-1959.
52. Chisholm ST, Coaker G, Day B, Staskawicz BJ: **Host-microbe interactions: shaping the evolution of the plant immune response.** *Cell* 2006, **124**(4):803-814.
53. Wortman JR, Fedorova N, Crabtree J, Joardar V, Maiti R, Haas BJ, Amedeo P, Lee E, Angiuoli SV, Jiang B, Anderson MJ, Denning DW, White OR, Nierman WC: **Whole genome comparison of the *A. fumigatus* family.** *Med Mycol* 2006, **44**(6):3-7.

54. Fedorova ND, Khaldi N, Joardar VS, Maiti R, Amedeo P, Anderson MJ, Crabtree J, Silva JC, Badger JH, Albarraq A, Angiuoli S, Bussey H, Bowyer P, Cotty PJ, Dyer PS, Egan A, Galens K, Fraser-Liggett CM, Haas BJ, Inman JM, Kent R, Lemieux S, Malavazi I, Orvis J, Roemer T, Ronning CM, Sundaram JP, Sutton G, Turner G, Venter JC, et al: **Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus***. *PLoS Genet* 2008, **4(4)**:e1000046.
55. Payne GA, Nierman WC, Wortman JR, Pritchard BL, Brown D, Dean RA, Bhatnagar D, Cleveland TE, Machida M, Yu J: **Whole genome comparison of *A. flavus* and *A. oryzae***. *Med Mycol* 2006, **44(6)**:9-11.
56. Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, Arroyo J, Berriman M, Abe K, Archer DB, Bermejo C, Bennett J, Bowyer P, Chen D, Collins M, Coulson R, Davies R, Dyer PS, Farman M, Fedorova N, Fedorova N, Feldblyum TV, Fischer R, Fosker N, Fraser A, Garcia JL, Garcia MJ, Goble A, Goldman GH, Gomi K, Griffith-Jones S, et al: **Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus***. *Nature* 2005, **438(7071)**:1151-1156.
57. Galagan JE, Calvo SE, Cuomo C, Ma LJ, Wortman JR, Batzoglou S, Lee SI, Basturkmen M, Spevak CC, Clutterbuck J, Kapitonov V, Jurka J, Scacciocchio C, Farman M, Butler J, Purcell S, Harris S, Braus GH, Draht O, Busch S, D'Enfert C, Bouchier C, Goldman GH, Bell-Pedersen D, Griffiths-Jones S, Doonan JH, Yu J, Vienken K, Pain A, Freitag M, et al: **Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae***. *Nature* 2005, **438(7071)**:1105-1115.
58. Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G, Schaap PJ, Turner G, de Vries RP, Albang R, Albermann K, Andersen MR, Bendtsen JD, Benen JA, van den Berg M, Breestraat S, Caddick MX, Contreras R, Cornell M, Coutinho PM, Danchin EG, Debets AJ, Dekker P, van Dijk PW, van Dijk A, Dijkhuizen L, Driessen AJ, d'Enfert C, Geysens S, Goosen C, Groot GS, et al: **Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88**. *Nat Biotechnol* 2007, **25(2)**:221-231.
59. Machida M, Asai K, Sano M, Tanaka T, Kumagai T, Terai G, Kusumoto K, Arima T, Akita O, Kashiwagi Y, Abe K, Gomi K, Horiuchi H, Kitamoto K, Kobayashi T, Takeuchi M, Denning DW, Galagan JE, Nierman WC, Yu J, Archer DB, Bennett JW, Bhatnagar D, Cleveland TE, Fedorova ND, Gotoh O, Horikawa H, Hosoyama A, Ichinomiya M, Igarashi R, et al: **Genome sequencing and analysis of *Aspergillus oryzae***. *Nature* 2005, **438(7071)**:1157-1161.
60. Sharpton TJ, Stajich JE, Rounsley SD, Gardner MJ, Wortman JR, Joardar VS, Maiti R, Kodira CD, Neafsey DE, Zeng Q, Hung CY, McMahan C, Muszewska A, Grynberg M, Mandel MA, Kellner EM, Barker BM, Galiani JN, Orbach MJ, Kirkland TN, Cole GT, Henn MR, Birren BW, Taylor JW: **Comparative genomic analyses of the human fungal pathogens *Coccidioides* and their relatives**. *Genome Res* 2009, **19(10)**:1722-1731.
61. Cuomo CA, Guldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma LJ, Baker SE, Rep M, Adam G, Antoniw J, Baldwin T, Calvo S, Chang YL, Decaprio D, Gale LR, Gnerre S, Goswami RS, Hammond-Kosack K, Harris LJ, Hillburn K, Kennell JC, Kroken S, Magnuson JK, Mannhaupt G, Mauceli E, Mewes HW, Mitterbauer R, Muehlbauer G, et al: **The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization**. *Science* 2007, **317(5843)**:1400-1402.
62. Coleman JJ, Rounsley SD, Rodriguez-Carres M, Kuo A, Wasmann CC, Grimwood J, Schmutz J, Taga M, White GJ, Zhou S, Schwartz DC, Freitag M, Ma LJ, Danchin EG, Henrissat B, Coutinho PM, Nelson DR, Straney D, Napoli CA, Barker BM, Gribskov M, Rep M, Kroken S, Molnar I, Rensing C, Kennell JC, Zamora J, Farman ML, Selker EU, Salamov A, et al: **The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion**. *PLoS Genet* 2009, **5(8)**:e1000618.
63. Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu JR, Pan H, Read ND, Lee YH, Carbone I, Brown D, Oh YY, Donofrio N, Jeong JS, Soanes DM, Djonovic S, Kolomiets E, Rehmeier C, Li W, Harding M, Kim S, Lebrun MH, Bohnert H, Coughlan S, Butler J, Calvo S, Ma LJ, et al: **The genome sequence of the rice blast fungus *Magnaporthe grisea***. *Nature* 2005, **434(7036)**:980-986.
64. Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen CB, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-Pedersen D, Nelson MA, Werner-Washburne M, Selitrennikoff CP, Kinsey JA, Braun EL, Zelter A, Schulte U, Kothe GO, Jedd G, Mewes W, et al: **The genome sequence of the filamentous fungus *Neurospora crassa***. *Nature* 2003, **422(6934)**:859-868.
65. van den Berg MA, Albang R, Albermann K, Badger JH, Daran JM, Driessen AJ, Garcia-Estrada C, Fedorova ND, Harris DM, Heijne WH, Joardar V, Kiel JA, Kovalchuk A, Martin JF, Nierman WC, Nijland JG, Pronk JT, Roubos JA, van der Klei IJ, van Peij NN, Veenhuis M, von Dohren H, Wagner C, Wortman J, Bovenberg RA: **Genome sequencing and analysis of the filamentous fungus *Penicillium chrysogenum***. *Nat Biotechnol* 2008, **26(10)**:1161-1168.
66. Espagne E, Lespinet O, Malagnac F, Da Silva C, Jaillon O, Porcel BM, Couloux A, Aury JM, Segurens B, Poulain J, Anthouard V, Grossetete S, Khalili H, Coppin E, Dequard-Chablat M, Picard M, Contamine V, Arnaise S, Bourdais A, Berteaux-Lecellier V, Gautheret D, de Vries RP, Battaglia E, Coutinho PM, Danchin EG, Henrissat B, Khoury RE, Sainsard-Chanet A, Boivin A, Pinan-Lucarre B, et al: **The genome sequence of the model ascomycete fungus *Podospira anserina***. *Genome Biol* 2008, **9(5)**:R77.
67. Hane JK, Lowe RG, Solomon PS, Tan KC, Schoch CL, Spatafora JW, Crous PW, Kodira C, Birren BW, Galagan JE, Torriani SF, McDonald BA, Oliver RP: **Dothideomycete Plant Interactions Illuminated by Genome Sequencing and EST Analysis of the Wheat Pathogen *Stagonospora nodorum***. *Plant Cell* 2007, **19(11)**:3347-3368.
68. Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, Danchin EG, Grigoriev IV, Harris P, Jackson M, Kubicek CP, Han CS, Ho I, Larrondo LF, de Leon AL, Magnuson JK, Merino S, Misra M, Nelson B, Putnam N, Robbertse B, Salamov AA, Schmoll M, Terry A, Thayer N, Westerholm-Parvinen A, et al: **Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*)**. *Nat Biotechnol* 2008, **26(5)**:553-560.
69. Hobbitt DS, Binder M, Bischoff JF, Blackwell M, Cannon PF, Eriksson OE, Huhndorf S, James T, Kirk PM, Lucking R, Thorsten Lumbsch H, Lutzoni F, Matheny PB, McLaughlin DJ, Powell MJ, Redhead S, Schoch CL, Spatafora JW, Stalpers JA, Vilgalys R, Aime MC, Aptroot A, Bauer R, Begerow D, Benny GL, Castlebury LA, Crous PW, Dai YC, Gams W, Geiser DM, et al: **A higher-level phylogenetic classification of the Fungi**. *Mycol Res* 2007, **111(Pt 5)**:509-547.
70. Jones T, Federspiel NA, Chibana H, Dungan J, Kalman S, Magee BB, Newport G, Thorstenson YR, Agabian N, Magee PT, Davis RW, Scherer S: **The diploid genome sequence of *Candida albicans***. *Proc Natl Acad Sci USA* 2004, **101(19)**:7329-7334.
71. van het Hoog M, Rast TJ, Martchenko M, Grindle S, Dignard D, Hogues H, Cuomo C, Berriman M, Scherer S, Magee BB, Whiteway M, Chibana H, Nantel A, Magee PT: **Assembly of the *Candida albicans* genome into sixteen supercontigs aligned on the eight chromosomes**. *Genome Biol* 2007, **8(4)**:R52.
72. Butler G, Rasmussen MD, Lin MF, Santos MA, Sakthikumar S, Munro CA, Rheinbay E, Grabherr M, Forche A, Reedy JL, Agrafioti I, Arnaud MB, Bates S, Brown AJ, Brunke S, Costanzo MC, Fitzpatrick DA, de Groot PW, Harris D, Hoyer LL, Hube B, Klis FM, Kodira C, Lennard N, Logue ME, Martin R, Neiman AM, Nikolaou E, Quail MA, Quinn J, et al: **Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes**. *Nature* 2009, **459(7247)**:657-662.
73. Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, De Montigny J, Marck C, Neuvéglise C, Talla E, Goffard N, Frangeul L, Aigle M, Anthouard V, Babour A, Barbe V, Barnay S, Blanchin S, Beckerich JM, Beyne E, Bleykasten C, Boisrame A, Boyer J, Cattolico L, Confanioli F, De Daruvar A, Despons L, Fabre E, Fairhead C, Ferry-Dumazet H, et al: **Genome evolution in yeasts**. *Nature* 2004, **430(6995)**:35-44.
74. Blandin G, Ozier-Kalogeropoulos O, Wincker P, Artiguenave F, Dujon B: **Genomic exploration of the hemiascomycetous yeasts: 16. *Candida tropicalis***. *FEBS Lett* 2000, **487(1)**:91-94.
75. Dietrich FS, Voegelé S, Brachat S, Lerch A, Gates K, Steiner S, Mohr C, Pohlmann R, Luedi P, Choi S, Wing RA, Flavier A, Gaffney TD, Philippsen P: **The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome**. *Science* 2004, **304(5668)**:304-307.
76. Scannell DR, Frank AC, Conant GC, Byrne KP, Woolfit M, Wolfe KH: **Independent sorting-out of thousands of duplicated gene pairs in two yeast species descended from a whole-genome duplication**. *Proc Natl Acad Sci USA* 2007, **104(20)**:8397-8402.
77. Kellis M, Birren BW, Lander ES: **Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae***. *Nature* 2004, **428(6983)**:617-624.

78. Jeffries TW, Grigoriev IV, Grimwood J, Laplaza JM, Aerts A, Salamov A, Schmutz J, Lindquist E, Dehal P, Shapiro H, Jin YS, Passoth V, Richardson PM: **Genome sequence of the lignocellulose-bioconverting and xylose-fermenting yeast *Pichia stipitis***. *Nat Biotechnol* 2007, **25**(3):319-326.
79. Cliften P, Sudarsanam P, Desikan A, Fulton L, Fulton B, Majors J, Waterston R, Cohen BA, Johnston M: **Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting**. *Science* 2003, **301**(5629):71-76.
80. Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES: **Sequencing and comparison of yeast species to identify genes and regulatory elements**. *Nature* 2003, **423**(6937):241-254.
81. Liti G, Carter DM, Moses AM, Warringer J, Parts L, James SA, Davey RP, Roberts IN, Burt A, Koufopanov V, Tsai IJ, Bergman CM, Bensasson D, O'Kelly MJ, van Oudenaarden A, Barton DB, Bailes E, Nguyen AN, Jones M, Quail MA, Goodhead I, Sims S, Smith F, Blomberg A, Durbin R, Louis EJ: **Population genomics of domestic and wild yeasts**. *Nature* 2009, **458**(7236):337-341.
82. Doniger SW, Kim HS, Swain D, Corcuera D, Williams M, Yang SP, Fay JC: **A catalog of neutral and deleterious polymorphism in yeast**. *PLoS Genet* 2008, **4**(8):e1000183.
83. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG: **Life with 6000 genes**. *Science* 1996, **274**(5287):563-547.
84. Gu Z, David L, Petrov D, Jones T, Davis RW, Steinmetz LM: **Elevated evolutionary rates in the laboratory strain of *Saccharomyces cerevisiae***. *Proc Natl Acad Sci USA* 2005, **102**(4):1092-1097.
85. Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R, Stewart A, Sgouros J, Peat N, Hayles J, Baker S, Basham D, Bowman S, Brooks K, Brown D, Brown S, Chillingworth T, Churcher C, Collins M, Connor R, Cronin A, Davis P, Feltwell T, Fraser A, Gentles S, Goble A, Hamlin N, Harris D, Hidalgo J, Hodgson G, Holroyd S, et al: **The genome sequence of *Schizosaccharomyces pombe***. *Nature* 2002, **415**(6874):871-880.
86. Loftus BJ, Fung E, Roncaglia P, Rowley D, Amedeo P, Bruno D, Vamathevan J, Miranda M, Anderson IJ, Fraser JA, Allen JE, Bosdet IE, Brent MR, Chiu R, Doering TL, Donlin MJ, D'Souza CA, Fox DS, Grinberg V, Fu J, Fukushima M, Haas BJ, Huang JC, Janbon G, Jones SJ, Koo HL, Krzywinski MI, Kwon-Chung JK, Lengeler KB, Maiti R, et al: **The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans***. *Science* 2005, **307**(5713):1321-1324.
87. Martin F, Aerts A, Ahren D, Brun A, Danchin EG, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V, Salamov A, Shapiro HJ, Wuyts J, Blaudez D, Buee M, Brokstein P, Canback B, Cohen D, Courty PE, Coutinho PM, Delaruelle C, Detter JC, Deveau A, DiFazio S, Duplessis S, Fraissinet-Tachet L, Lucic E, Frey-Klett P, Fourrey C, Feussner I, et al: **The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis**. *Nature* 2008, **452**(7183):88-92.
88. Martinez D, Larrondo LF, Putnam N, Gelpke MD, Huang K, Chapman J, Helfenbein KG, Ramaiya P, Detter JC, Larimer F, Coutinho PM, Henrissat B, Berka R, Cullen D, Rokhsar D: **Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78**. *Nat Biotechnol* 2004, **22**(6):695-700.
89. Kamper J, Kahmann R, Bolker M, Ma LJ, Brefort T, Saville BJ, Banuett F, Kronstad JW, Gold SE, Muller O, Perlin MH, Wosten HA, de Vries R, Ruiz-Herrera J, Reynaga-Pena CG, Snetselaar K, McCann M, Perez-Martin J, Feldbrugge M, Basse CW, Steinberg G, Ibeas JI, Holloman W, Guzman P, Farman M, Stajich JE, Sentandreu R, Gonzalez-Prieto JM, Kennell JC, Molina L, et al: **Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis***. *Nature* 2006, **444**(7115):97-101.
90. Ma LJ, Ibrahim AS, Skory C, Grabherr MG, Burger G, Butler M, Elias M, Idnurm A, Lang BF, Sone T, Abe A, Calvo SE, Corrochano LM, Engels R, Fu J, Hansberg W, Kim JM, Kodira CD, Koehrsen MJ, Liu B, Miranda-Saavedra D, O'Leary S, Ortiz-Castellanos L, Poulter R, Rodriguez-Romero J, Ruiz-Herrera J, Shen YQ, Zeng Q, Galagan J, Birren BW, et al: **Genomic analysis of the basal lineage fungus *Rhizopus oryzae* reveals a whole-genome duplication**. *PLoS Genet* 2009, **5**(7):e1000549.
91. Katinka MD, Duprat S, Cornillot E, Metenier G, Thomarat F, Prensier G, Barbe V, Peyretailade E, Brottier P, Wincker P, Delbac F, El Alaoui H, Peyret P, Saurin W, Gouy M, Weissenbach J, Vivares CP: **Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi***. *Nature* 2001, **414**(6862):450-453.
92. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Perlea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, et al: **Genome sequence of the human malaria parasite *Plasmodium falciparum***. *Nature* 2002, **419**(6906):498-511.
93. Haas BJ, Kamoun S, Zody MC, Jiang RH, Handsaker RE, Cano LM, Grabherr M, Kodira CD, Raffaele S, Torto-Alalibo T, Bozkurt TO, Ah-Fong AM, Alvarado L, Anderson VL, Armstrong MR, Avrova A, Baxter L, Beynon J, Boevink PC, Bollmann SR, Bos JJ, Bulone V, Cai G, Cakir C, Carrington JC, Chawner M, Conti L, Costanzo S, Ewan R, Fahlgren N, et al: **Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans***. *Nature* 2009, **461**(7262):393-398.
94. Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RH, Aerts A, Arredondo FD, Baxter L, Bensasson D, Beynon JL, Chapman J, Damasceno CM, Dorrance AE, Dou D, Dickerman AW, Dubchak IL, Garbelotto M, Gijzen M, Gordon SG, Govers F, Grunwald NJ, Huang W, Ivors KL, Jones RW, Kamoun S, Kramps K, Lamour KH, Lee MK, McDonald WH, Medina M, et al: ***Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis**. *Science* 2006, **313**(5791):1261-1266.

doi:10.1186/1471-2164-11-105

Cite this article as: Choi et al.: Fungal Secretome Database: Integrated platform for annotation of fungal secretomes. *BMC Genomics* 2010 11:105.

Submit your next manuscript to BioMed Central
and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

