Journal of Chromatography B, 899 (2012) 86–95

Contents lists available at SciVerse ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb

Application of nanoLC–MS/MS to the shotgun proteomic analysis of the nematocyst proteins from jellyfish *Stomolophus meleagris*

Rongfeng Li[a,b], Huahua Yu[a,*], Ronge Xing[a], Song Liu[a], Yukun Qing[a,b], Kecheng Li[a,b], Bing Li[a,b], Xiangtao Meng[a,b], Jinhui Cui[a,b], Pengcheng Li[a,*]

[a] *Institute of Oceanology, Chinese Academy of Sciences, 7 Nanhai Road, Qingdao 266071, China*
[b] *Graduate University of the Chinese Academy of Sciences, 19 Yuquan Road, Beijing 100039, China*

ARTICLE INFO

Article history:
Received 2 February 2012
Accepted 2 May 2012
Available online 15 May 2012

Keywords:
Jellyfish
Stomolophus meleagris
Nematocyst proteins
Shotgun
NanoLC–MS/MS
Proteomic analysis

ABSTRACT

The nematocyst proteins of jellyfish *Stomolophus meleagris*, a complicated mixture, contain many important bioactive molecules. In present study, to gain comprehensive insight into the protein component and search some novel bioactive molecules in the nematocyst proteins, shotgun proteomic analysis of the nematocyst proteins was carried out by nano liquid chromatography tandem mass spectrometry (nanoLC–MS/MS) for the first time. Digested peptides of the nematocyst proteins were analyzed by nanoLC–MS/MS and all MS/MS spectra were then automatically searched by the SEQUEST program. A total of 181 proteins had been identified, with the molecular weight ranging from 5268.06 to 843,487.57 and the p*I* from 4.49 to 11.39. Bioinformatic analysis was also applied to better understand the identified proteins. In the gene ontology (GO) annotation, all the identified proteins were classified into 13, 9 and 7 groups according to biological process, cellular component and molecular function, respectively. Pathways analysis of the identified proteins was conducted with 33 corresponding pathways found. On the basis of pathways analysis, we also constructed the gene network to analyze the relationship of those genes each other, which contained enzyme–enzyme relation, protein–protein interaction and gene expression interaction.

1. Introduction

The nematocyst, which is predominantly in the surface of jellyfish tentacles, contains many complex and diversiform biomolecules with plenty of activities, such as hemolytic activity, lethal activity, cytotoxic activity, neurotoxic activity, insecticidal activity, antioxidant activity, cardiovascular activity, enzyme activity and other potential activities [1–10]. As a result, increasing attention had been attracted to investigate the jellyfish nematocyst proteins, including bioactivity analysis, isolation, and identification of bioactive molecules. For example, hemolytic proteins CAH1, CaTX-A, CaTX-B, CrTX-A, CrTX-B and CqTX-A had been isolated from the nematocyst of jellyfish *Carybdea alata, Carybdea rastoni* and *Chiropsalmus quadrigatus*. CqTX-A and CrTX-A had also been proved to be lethal to crayfish with the LD_{50} of 80 and 5 μg/kg, respectively [11–14]. Besides, cytotoxic protein ClGp1 had been isolated from the jellyfish Scyphozoa *Cyanea lamarckii* [5]. In addition, neurotoxic protein CmNt had been purified and identified from the nematocyst of jellyfish *Carybdea marsupialis* [6]. Although, such studies had described the biological properties and structures of some nematocyst proteins, the process of the isolation, identification and characterization analysis of those proteins individually was so difficult, time-consuming and incomplete that a globe analysis of the nematocyst proteins is necessary to deeply understand the component, structure of the nematocyst proteins.

With the progress in the proteomic analysis, the proteomic technique began to be employed to analyze the whole component from the cell of organism and seek for novel bioactive proteins instead of individual bioactivity analysis. A classic method of two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) coupled with MS was employed by researchers to separate and identify proteins. However, traditional method of 2D-PAGE has significant limitations for protein analysis including dynamic range, deficiencies in proteome coverage, sensitivity and throughput. But, shotgun is a rapid, sensitive and convenient method to analyze and identify protein mixtures, which is becoming more and more important in the proteomic analysis.

In present study, the nematocyst proteins of jellyfish *Stomolophus meleagris* were identified by shotgun proteomic strategy

* Corresponding authors. Tel.: +86 532 82898707; fax: +86 532 82968951.
E-mail addresses: yuhuahua@qdio.ac.cn (H. Yu), pcli@ms.qdio.ac.cn (P. Li).

1570-0232/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jchromb.2012.05.006

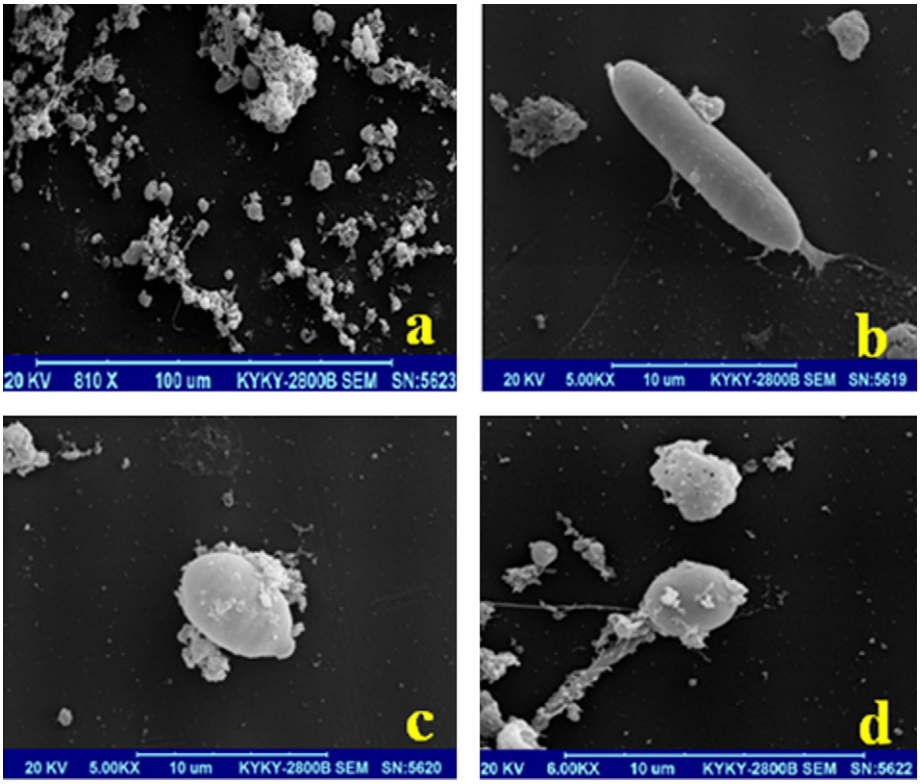

Fig. 1. SEM of nematocysts of the jellyfish *Stomolophus meleagris*. (a) The nematocyst of the jellyfish *Stomolophus meleagris* in SEM. (b) Rod-like and undischarged nematocyst of the jellyfish *Stomolophus meleagris* in SEM. (c) Elliptic and undischarged nematocyst of the jellyfish *Stomolophus meleagris* in SEM. (d) Elliptic and discharged nematocyst of the jellyfish *Stomolophus meleagris* in SEM.

with nanoLC–MS/MS to analyze the component, structure and function of proteins for the first time. Bioinformatic analysis of the identified proteins was also carried out to better understand the proteins. The identified proteins were classified according to cellular component, biological process and molecular function using GO database; pathways of identified proteins were analyzed in the database of KEGG pathway and the gene network was also constructed to analyze the relationship of those proteins each other including enzyme–enzyme relation, protein–protein interaction and gene expression interaction.

2. Materials and methods

2.1. Sample collection and preparation

Jellyfish *Stomolophus meleagris* were collected from the coast of Qingdao, China in August 2010. The tentacles were excised manually from living specimens as soon as possible after capture and were then immediately frozen at −80 °C for subsequent use. The extraction method of nematocyst proteins was the same as we described before [3]. Briefly, the nematocyst was sonicated 30 min for 10 s intervals in cold extraction buffer (4 °C, 20 mM Tris–HCl, pH 7.8, 0.15 M NaCl, 1 mM EDTA, 5 μg/mL Pepstatin A and 0.5 mM PMSF). After centrifugation at 10,000 × *g* for 15 min, 4 °C, the supernatant was considered to be the nematocyst proteins. Protein concentration of the sample was determined using a Bradford assay with bovine serum albumin (BSA) protein as a standard [15].

2.2. The scanning of nematocysts by SEM

The preparation of nematocyst suspension sample, used for scanning electron microscopy (SEM), was performed as follows: the nematocysts were separated from the tentacles and then fixed in 2.5% glutaraldehyde 100 mM, pH 7.4 phosphate buffer for 1 h, followed by dehydrating with 30–70% ethanol, drying in acetate-substituted ethanol for 30 min and gold coating in vacuum. The nematocysts were then observed with scanning electron microscopy.

2.3. SDS–PAGE analysis

The nematocyst proteins were analyzed by SDS–PAGE according to the Laemmli method [16]. Brifely, the sample with same volume of 2× loading buffer was mixed and boiled for 5 min and centrifuged at 10,000 × *g* for 15 min and then about 20 μg proteins were separated using 5% stacking gel (5% acrylamide/bis-acrylamide, 10%SDS, 10% AP, 5 μL TEMED, pH 6.8 Tris–HCl) and 12% separation gel (12% acrylamide/bis-acrylamide, 10%SDS, 10% AP, 10 μL TEMED, pH 8.8 Tris–HCl). The electrophoresized gel was visualized using the Silver stain method [17], and the molecular weights determined by comparison with standard proteins of 116 kDa β-galactosidase, 66.2 kDa bovine serum albumin, 45.0 kDa ovalbumin, 35.0 kDa lactate dehydrogenase, 25.0 kDa REase bsp981, 18.4 kDa β-lactoglobulin and 14.4 kDa lysozyme (Fermentas).

2.4. Sample digestion

Sample digestion was performed as follows: the nematocyst proteins were dissolved in 6 M guanidine hydrochloride, 100 mM, pH 8.3, Tris–HCl and then reacted at 37 °C for 2.5 h with 10 mM DTT. Subsequently, 1 M indoleacetic acid was added with the terminal concentration of 50 mM and reacted at room temperature in the dark for 40 min. Then, 200 μL, 100 mM NH_4HCO_3 was added and centrifuged at 4 °C, 10,000 × *g* for 2 h with ultrafiltration device

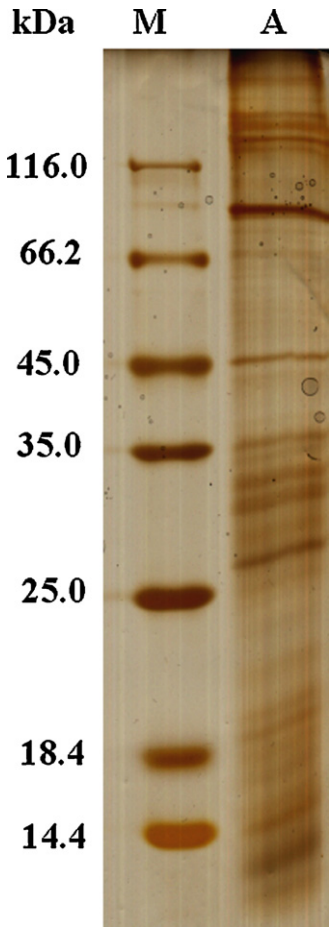

Fig. 2. SDS–PAGE analysis of the nematocyst proteins of jellyfish *Stomolophus meleagris*. M: marker, C: crude nematocyst proteins.

(MWCO 3K). Whereafter, the sample was adjusted to pH 8.0–8.5 and then digested with trypsin (trypsin: proteins = 1:20) for 20 h at 37 °C. Finally, the digested proteins were freeze-dried and stored at −80 °C for use.

2.5. Identification by nanoLC–MS/MS

Ettan MDLC system (GE Healthcare) was applied for desalting and separation of tryptic peptides mixtures. In this system, the digested peptides were desalted on RP trap columns (Zorbax 300 SB C18, Agilent Technologies), and then separated on a analytical RP column (150 μm i.d., 100 mm length, Column technology Inc., Fremont, CA), which was equilibrated with 95% buffer A for 15 min. Buffer A (0.1% formic acid in HPLC grade water) and the buffer B (0.1% formic acid in acetonitrile) were selected. 20 μg of tryptic peptide mixtures was loaded onto the columns, and separation was done at a flow rate of 2 μL/min by using a linear gradient of 4–50% buffer B for 120 min. A Finnigan LTQ linear ion trap MS (Thermo Electron) equipped with an electrospray interface was connected to the LC setup for eluted peptides detection. The positive ion mode was employed and the mass spectrometer with the application of a spray voltage was set at 3.0 kV. The spray temperature was set at 200 °C for peptides. Normalized collision energy was set to 35% and the full scan ranged from *m*/*z* 300 to 1800. Data-dependent MS/MS spectra were obtained simultaneously. Each scan cycle consisted of one full MS scan in profile mode followed by five MS/MS scans in centroid mode with the following dynamic exclusion settings: repeat count 2, repeat duration 30 s, exclusion duration 90 s. Each sample was analyzed in triplicate.

2.6. Database search and bioinformatics analysis

All MS/MS spectra were automatically searched from the non-redundant International Protein Index (IPI) *Cnidaria* database (version 3.26, 67,687 entries) using the BioworksBrowser rev. 3.1(Thermo Electron, San Jose, CA). Protein identification results were extracted from SEQUEST out files with BuildSummary

The peptides were constrained to be tryptic and up to two missed cleavages were allowed. Carbamidomethylation of cysteines were treated as a fixed modification, whereas oxidation of methionine residues was considered as variable modifications. The mass tolerance allowed for the precursor ions was 2.0 Da and fragment ions was 0.2 Da, respectively. The protein identification criteria were based on Delta CN ($\geq$0.1) and cross-correlation scores (Xcorr, one charge $\geq$ 1.9, two charges $\geq$ 2.2, three charges $\geq$ 3.75) and the false discovery rate was under 1%.

Reannotations of the identified proteins were searched against the database swissprot. Further bioinformatics of gene ontology was analyzed according to biological process, cellular components and biological functions using GSEABase. Pathways of the identified proteins were searched against the database KEGG pathway (GenMAPP v2.1) and collected all the enrichment *p*-value of the gene in each pathway. The network of the corresponding genes was analyzed by KEGGSOAP according to enzyme–enzyme relation, protein–protein interaction and gene expression interaction.

3. Results

3.1. The scanning of nematocysts by SEM

As seen from Fig. 1a, SEM examination of the nematocysts of jellyfish *Stomolophus meleagris* revealed that the purity of the nematocyst preparation from the jellyfish was about 95% and it was composed of two main types of elliptic and rod-like nematocyst. The elliptic nematocyst, which account for more than 90%, is about 6–8 μm long. However, the rod-like nematocyst only occupied less than 10%, with about 12–15 μm long. The strange nematocyst with a long tail, like a tadpole, is also observed from Fig. 1d, which is the nematocyst after discharge.

3.2. Identification by nanoLC–MS/MS

In Fig. 2, it is seen that plenty of protein bands had been clearly observed in the SDS–PAGE, which indicated that the sample was suitable for the shotgun analysis. Therefore, the sample, after being digested by trypsin, was analyzed by nanoLC–MS/MS and 181 proteins had been identified (see Table 1). 86.28% (156/181) of the identified proteins had single unique peptide, while the others had more than one and even to12 unique peptides, such as the TBA2 and TBA3.

3.3. Characterization of identified proteins

The physical and chemical properties of the identified proteins such as molecular mass (MW) and p*I* distribution were carried out. Most of the identified proteins' MW ranged from 20 to100 kDa and p*I* from 5 to 10 observed in Fig. 3. However, there are also some proteins, whose MW are extra large to more than 300 kDa and even to 843.4876 kDa. Only few proteins' p*I* is more than 11 or less than 4 (see Fig. 4).

Table 1
Shotgun analysis result of the nematocyst proteins of jellyfish *Stomolophus meleagris*.

Protein name	Gene symbol	Gi no.	No. of unique peptides	MW	pI
Tubulin alpha-2/alpha-4 chain	TBA2	gi|32967406	12	42,659	6.07
Tubulin alpha-3 chain	TBA3	gi|156394507	12	50,203	4.91
Tubulin alpha-3 chain	TBA3	gi|156394505	12	50,235	4.91
Actin	ACT	gi|312861909	8	41,767	5.29
Tubulin beta chain	TBB	gi|221129327	7	50,049	4.74
Myosin heavy chain	MYS	gi|221113859	4	238,150	5.59
Clathrin heavy chain 1	CLH1	gi|156394133	4	191,233	5.6
Histone H2A	H2AV	gi|156398753	3	13,362	10.58
78 kDa glucose-regulated protein	GRP78	gi|221132017	3	73,995	5.4
ATP synthase subunit beta	ATPB	gi|156364605	3	44,840	4.92
40S ribosomal protein S9	RS9	gi|156361995	3	22,194	10.53
Clathrin heavy chain 1	CLH1	gi|221130531	3	192,025	5.61
AP-2 complex subunit alpha-2	AP2A2	gi|156406763	2	104,984	6.39
Transitional endoplasmic reticulum ATPase	TERA	gi|156370042	2	89,172	4.97
Histone H4	H4	gi|156322163	2	20,084	11.09
Tubulin alpha chain	TBA	gi|221105656	2	8513.3	4.11
Tubulin beta chain	TBB	gi|163638946	2	13,028	6.3
Heat shock protein HSP 90-beta	HS90B	gi|14041148	2	84,319	4.8
Myosin-10	MYH10	gi|221130996	2	220,047	5.53
Elongation factor 1-alpha	EF1A	gi|312861905	2	50,742	9.29
TopBP1-interacting checkpoint and replication regulator	TICRR	gi|156369983	2	206,947	9.41
Coatomer subunit gamma-2	COPG2	gi|156382581	2	98,428	5.49
Copine-3	CPNE3	gi|156398026	2	58,303	5.7
Transitional endoplasmic reticulum ATPase TER94	TERA	gi|221109190	2	57,474	4.94
Histone H3.3	H33	gi|156356095	2	15,312	11.27
Histone H2B.1/H2B.2	H2B1	gi|221136709	1	14,563	10.34
Succinate dehydrogenase ubiquinone flavoprotein subunit	DHSA	gi|221129803	1	72,044	6.63
UPF0727 protein v1g158749	U727	gi|224493391	1	9350.8	5.76
26S protease regulatory subunit 6A	PRS6A	gi|156373935	1	47,820	5.05
Dickkopf-related protein 3	DKK3	gi|221115495	1	21,888	8.33
60S ribosomal protein L17	RL17	gi|585868	1	21,161	10.19
ATP synthase subunit alpha	ATPA	gi|221131162	1	58,779	9.24
NAD(P) transhydrogenase, mitochondrial	NNTM	gi|156375124	1	105,419	5.87
AP-1 complex subunit mu-1	AP1M1	gi|156395641	1	48,973	7.15
WD repeat-containing protein 65	WDR65	gi|156369738	1	98,157	5.57
Ras-like GTP-binding protein RHO	RHO	gi|156379567	1	21,706	6
Hypoxanthine-guanine phosphoribosyltransferase	HPRT	gi|221113545	1	19,649	5.74
6-phosphogluconate dehydrogenase, decarboxylating	6PGD	gi|156367416	1	53,058	7.62
T-complex protein 1 subunit alpha	TCPA	gi|156380526	1	54,432	6.38
NAD(P) transhydrogenase	NNTM	gi|221120408	1	113,384	8.24
ADP, ATP carrier protein	ADT	gi|156401103	1	33,259	9.67
40S ribosomal protein S3	RS3	gi|37724563	1	26,506	9.6
T-complex protein 1 subunit epsilon	TCPE	gi|156407978	1	59,655	5.81
Myosin-7 (fragment)	MYH7	gi|156395115	1	18,720	8.9
Anaphase-promoting complex subunit 10	APC10	gi|156376950	1	20,472	5.71
Histone H2B.1/H2B.2	H2B1	gi|221136304	1	7354.5	9.87
Protein regulator of cytokinesis 1	PRC1	gi|156405735	1	57,869	5.62
F-actin-capping protein subunit beta	CAPZB	gi|156351294	1	15,460	4.49
40S ribosomal protein S2	RS2	gi|216296537	1	12,203	10.39
Clathrin heavy chain 1	CLH1	gi|225547767	1	26,542	5.27
Vitrin	VITRN	gi|156358477	1	136,520	6.06
26S protease regulatory subunit S10B	PRS10	gi|156363865	1	44,236	6.32
rRNA 2′-O-methyltransferase fibrillarin	FBRL	gi|156380891	1	31,428	10.15
Putative insulin-like peptide receptor	HTK7	gi|2497558	1	168,278	7.53
Transportin-1	TNPO1	gi|156382619	1	99,708	4.89
Protein l(2)37Cc	L2CC	gi|156407434	1	30,242	6.03
14-3-3 protein gamma	1433G	gi|27368037	1	28,099	4.99
C-1-tetrahydrofolate synthase, cytoplasmic	C1TC	gi|221116976	1	32,655	7.12
HSPB1-associated protein 1	HBAP1	gi|221121200	1	59,851	5.23
SID1 transmembrane family member 1	SIDT1	gi|221113792	1	243,121	8.81
Calmodulin	CALM	gi|221128663	1	18,945	4.28
Rab GDP dissociation inhibitor alpha	GDIA	gi|221131840	1	49,885	5.42
Sodium/potassium-transporting ATPase subunit alpha	AT1A	gi|221104433	1	65,868	5.11
Replication factor C subunit 4	RFC4	gi|156383785	1	39,514	6.62
Multifunctional protein ADE2	PUR6	gi|221124646	1	43,815	6.35
Collagen alpha-1(XXVII) chain	CORA1	gi|120616160	1	250,391	5.18
Eukaryotic translation initiation factor 3 subunit G	EIF3G	gi|224488026	1	32,847	5.77
Dickkopf-related protein 3	DKK3	gi|37498690	1	21,221	8.42
PHD finger protein 14	PHF14	gi|156392036	1	24,397	7.89
Ubiquitin	UBIQ	gi|55228560	1	14,748	9.72
26S proteasome non-ATPase regulatory subunit 1	PSMD1	gi|221117036	1	101,292	5.5
Regulation of nuclear pre-mRNA domain-containing protein 1B	RPR1B	gi|156382492	1	41,564	9.08
ATP-dependent DNA helicase PIF1	PIF1	gi|221104481	1	136,939	6.74
Brain-specific angiogenesis inhibitor 1	BAI1	gi|221110288	1	31,493	7.28
Transcription factor RFX3	RFX3	gi|156392417	1	55,496	5.86

Table 1 (*Continued*)

Protein name	Gene symbol	Gi no.	No. of unique peptides	MW	p*I*
Alpha-actinin-1	ACTN1	gi|156384739	1	103,474	4.84
V-type proton ATPase subunit B	VATB	gi|221103824	1	55,979	5.55
Protein turtle homolog B	TUTLB	gi|156394459	1	27,677	9.07
Sodium/potassium-transporting ATPase subunit alpha	ATNA	gi|256010164	1	37,127	5.44
Tetratricopeptide repeat protein 21B	TT21B	gi|156407868	1	148,872	7.96
Electron transfer flavoprotein subunit beta	ETFB	gi|14029143	1	22,474	8.92
Peptidyl-prolyl cis-trans isomerase	PPIA	gi|140427776	1	17,542	9.1
Serine/threonine-protein kinase 17A	ST17A	gi|156372524	1	29,930	5.9
40S ribosomal protein S18	RS18	gi|156372342	1	17,856	10.42
Developmentally-regulated GTP-binding protein 1	DRG1	gi|156380796	1	40,462	8.86
24-Dehydrocholesterol reductase	DHC24	gi|156370088	1	61,198	6.71
Neurocalcin homolog	NCAH	gi|221113879	1	13,114	4.89
Ankyrin repeat and zinc finger domain-containing protein 1	ANKZ1	gi|156389420	1	32,576	9.04
Abnormal spindle-like microcephaly-associated protein homolog	ASPM	gi|221090853	1	43,207	7.07
Substance-K receptor	NK2R	gi|156356506	1	42,127	9.58
Kinesin-like protein KIF11	KIF11	gi|221124886	1	120,500	5.38
Alpha-aminoadipic semialdehyde synthase, mitochondrial	AASS	gi|156376340	1	95,323	6.55
Calcium-binding mitochondrial carrier protein Aralar1	CMC1	gi|156388071	1	78,131	8.7
Receptor-type tyrosine-protein phosphatase S	PTPRS	gi|221104221	1	115,819	8.51
Alpha-actinin-1	ACTN1	gi|32698308	1	28,971	6.02
Regulator of G-protein signaling 22	RGS22	gi|221107552	1	69,492	7.94
Ras suppressor protein 1	RSU1	gi|156394252	1	17,501	5.49
Coatomer subunit alpha	COPA	gi|156380509	1	137,568	7.87
Disintegrin and metalloproteinase domain-containing protein 22	ADA22	gi|221110446	1	43,227	9.33
Collagen alpha-4(VI) chain	CO6A4	gi|156363859	1	43,085	8.78
A disintegrin and metalloproteinase with thrombospondin motifs 17	ATS17	gi|221124121	1	843,488	5.43
Anoctamin-3	ANO3	gi|221124569	1	108,830	8.09
60 kDa heat shock protein, mitochondrial	CH60	gi|221110951	1	61,282	5.28
cGMP-dependent protein kinase, isozyme 1	KGP1	gi|156376630	1	75,055	8.53
Uncharacterized protein C7orf36	CG036	gi|221119409	1	25,419	4.67
Protein transport protein Sec61 subunit alpha isoform 2	S61A2	gi|221129307	1	52,151	8.65
Ras-related C3 botulinum toxin substrate 1	RAC1	gi|156376906	1	21,459	8.59
DnaJ homolog subfamily C member 13	DJC13	gi|221122839	1	56,113	7.78
Cell growth-regulating nucleolar protein	LYAR	gi|221119327	1	40,635	9.41
Pescadillo homolog	PESC	gi|221128259	1	61,266	8.98
Protein GRINL1A	GRL1A	gi|156371785	1	129,277	8.01
Heat shock cognate 71 kDa protein	HSP7C	gi|156373042	1	71,977	5.38
Heat shock protein 70 A1	HSP71	gi|156393957	1	69,865	5.45
Carbonyl reductase [NADPH] 1	CBR1	gi|221116884	1	31,375	5.54
Thrombospondin-1	TSP1	gi|221103079	1	39,584	8.45
Alanine aminotransferase 2	ALAT2	gi|221122695	1	49,362	8.08
U5 small nuclear ribonucleoprotein 200 kDa helicase	U520	gi|156406753	1	243,862	5.56
Histone-binding protein RBBP4	RBBP4	gi|221110561	1	39,228	5.27
Gelsolin-like protein 1	GELS1	gi|156382137	1	41,903	5.65
60S ribosomal protein L3	RL3	gi|156359547	1	46,323	10.09
40S ribosomal protein S23	RS23	gi|156408299	1	15,875	10.57
Tetratricopeptide repeat protein 28	TTC28	gi|221128805	1	159,163	7.51
Nucleolar protein 58	NOP58	gi|156406016	1	57,194	8.98
60S ribosomal protein L5	RL5	gi|156401420	1	33,824	9.65
Transposon TX1 uncharacterized 149 kDa protein	YTX2	gi|221106744	1	84,886	8.41
Dynein light chain Tctex-type 1	DYLT1	gi|221130745	1	44,750	6.46
60S ribosomal protein L23a	RL23A	gi|156402495	1	17,156	10.53
DNA replication licensing factor mcm2	MCM2	gi|221121676	1	108,959	5.17
60S ribosomal protein L7a	RL7A	gi|221117472	1	30,180	10.41
ATP-binding cassette sub-family A member 3	ABCA3	gi|221112746	1	73,892	5.15
Protein phosphatase 1D	PPM1D	gi|221116210	1	73,892	5.15
Alpha-enolase	ENOA	gi|156383570	1	46,812	6.06
Intron-binding protein aquarius	AQR	gi|156366211	1	163,399	6.42
Putative ankyrin repeat protein FPV031	V031	gi|156379996	1	14,723	6.73
Interferon-induced very large GTPase 1	GVIN1	gi|221110504	1	175,575	8.44
Probable citrate synthase 2, mitochondrial	CISY2	gi|221108566	1	52,073	7.61
Protein tyrosine phosphatase type IVA 3	TP4A3	gi|156372813	1	17,374	9.14
UPF0727 protein C6orf115 homolog	CF115	gi|221122293	1	9516	6.83
Epoxide hydrolase 4	EPHX4	gi|221,115,194	1	39,142	8.66
RING finger protein 213	RN213	gi|156,374,392	1	532,684	5.74
Kielin/chordin-like protein	KCP	gi|156365813	1	132,603	8.97
ADP-ribosylation factor GTPase-activating protein 2	ARFG2	gi|221121256	1	55,447	8.78
Probable ubiquitin carboxyl-terminal hydrolase FAF-X	USP9X	gi|221132909	1	268,337	5.69
Uncharacterized protein C1orf83	CA083	gi|156359603	1	25,000	9.77
Myosin-11	MYH11	gi|193089960	1	30,978	8.23
Surfeit locus protein 4	SURF4	gi|221102167	1	31,231	8.92
Casein kinase II subunit alpha	CSK21	gi|221112300	1	41,281	7.23
Centrosomal protein of 152 kDa	CE152	gi|156399598	1	230,187	5.34
ATPase family AAA domain-containing protein 3	ATAD3	gi|221116423	1	36,412	9.28
Neuroligin-4, X-linked	NLGNX	gi|156367572	1	62,391	6.36
Dynactin subunit 1	DCTN1	gi|156406663	1	124,544	5.19

Table 1 (*Continued*)

Protein name	Gene symbol	Gi no.	No. of unique peptides	MW	p*I*
Splicing factor 3B subunit 1	SF3B1	gi|156389470	1	148,778	6.56
40S ribosomal protein S14	RS14	gi|156386752	1	16,350	10.31
Sensor protein qseC	QSEC	gi|156345568	1	64,742	11.39
Elongation factor 2	EF2	gi|32967446	1	28,196	6.12
WD repeat-containing protein C10orf79	CJ079	gi|221123146	1	85,428	5.81
2′,3′-Cyclic-nucleotide 2′-phosphodiesterase	CNPD	gi|156344550	1	321,862	5.6
Protein translocase subunit secA	SECA	gi|221124288	1	98,684	5.49
SCL-interrupting locus protein	STIL	gi|156363514	1	144,244	6.55
Ras-related protein Rab-11B	RB11B	gi|221119156	1	58,220	9.1
Isocitrate dehydrogenase [NADP] cytoplasmic	IDHC	gi|156387427	1	44,330	6.84
Heterogeneous nuclear ribonucleoprotein U-like protein 1	HNRL1	gi|221103220	1	137,593	8.58
Cation-independent mannose-6-phosphate receptor	MPRI	gi|156399857	1	251,772	7.43
Kinesin-like protein KIF6	KIF6	gi|156359854	1	27,272	9.86
Peroxiredoxin-1	PRDX1	gi|221111162	1	22,056	5.64
Transmembrane emp24 domain-containing protein 9	TMED9	gi|221130607	1	25,327	6.76
Serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B alpha isoform	2ABA	gi|221102149	1	55,432	5.5
Leucyl-tRNA synthetase, cytoplasmic	SYLC	gi|221090747	1	61,100	6.99
Afadin	AFAD	gi|156358471	1	124,338	6.93
L-Threonine 3-dehydrogenase, mitochondrial	TDH	gi|156385006	1	39,774	6.41
Uncharacterized protein FLJ43738	YC006	gi|156359645	1	64,014	6.98
Ellis–van Creveld syndrome protein homolog	EVC	gi|156372704	1	47,632	5.73
Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15	DHX15	gi|221107484	1	82,754	7.17
Casein kinase II subunit alpha	CSK21	gi|156385394	1	43,617	6.57
cAMP-dependent protein kinase catalytic subunit alpha	KAPCA	gi|221106878	1	40,410	8.86
Glutamine synthetase 2 cytoplasmic	GLNA2	gi|156382657	1	41,665	5.93
GTP-binding nuclear protein Ran	RAN	gi|156373234	1	24,817	6.44
Alpha-centractin	ACTZ	gi|156405461	1	42,544	6.44
60S ribosomal protein L6	RL6	gi|221108921	1	9372.9	10.01
A disintegrin and metalloproteinase with thrombospondin motifs 10	ATS10	gi|156,363,512	1	185,398	8.68
Casein kinase II subunit beta	CSK2B	gi|156,373,129	1	25,708	5.98

3.4. Bioinformatics analysis

3.4.1. Gene ontology analysis

Classification of the identified proteins had been analyzed according to the biological process, cellular component and molecular function. As seen from Fig. 5a, in the biological process, the largest proportions of proteins (14%), except the unidentified 16% proteins, are related to the transport like electron transfer flavoprotein, myosin, dickkopf-related protein, and peptidyl-prolyl cis-trans isomerase; 11% proteins belong to cell organization and biogenesis including kinesin-like protein KIF11, DNA replication licensing factor MCM2; developmental processes proteins,

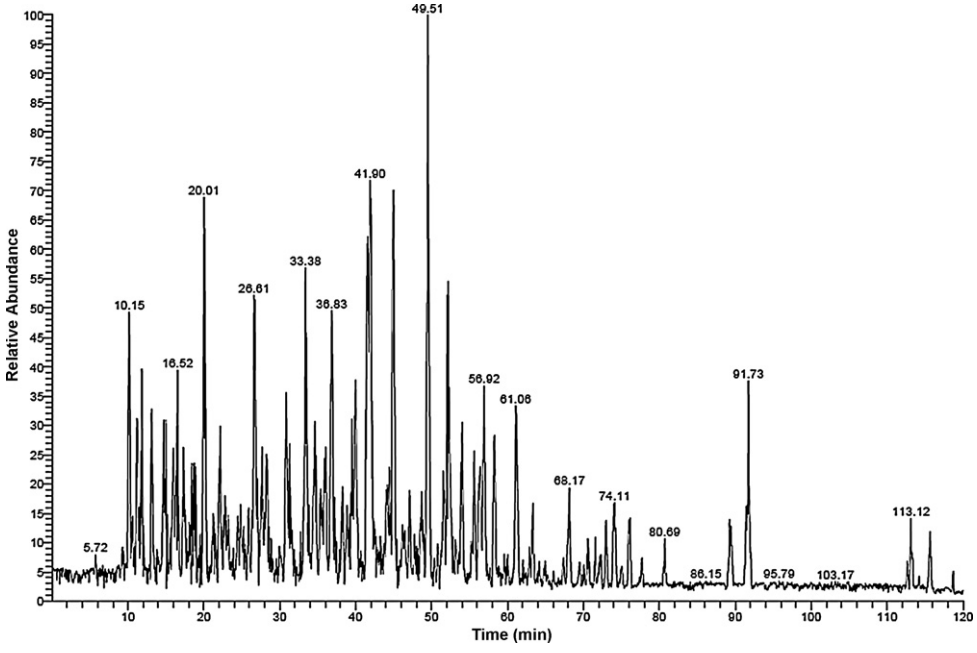

Fig. 3. MS spectrum of the nematocyst proteins of jellyfish *Stomolophus meleagris*.

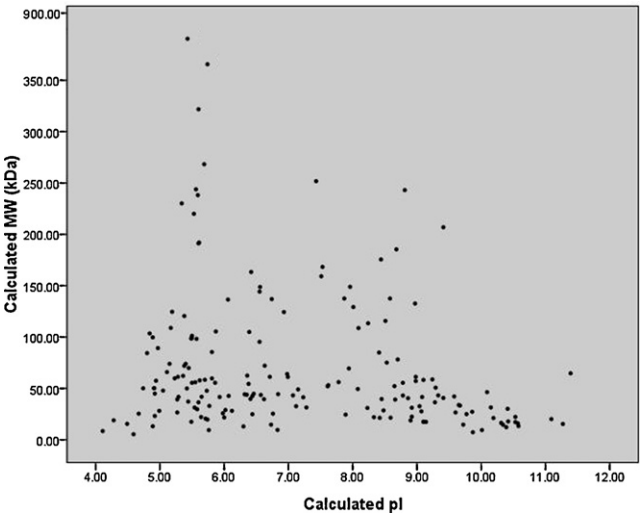

Fig. 4. Distribution of the molecular mass and p*I* of the identified proteins.

like splicing factor 3B, and abnormal spindle-like microcephaly associated protein, account for 10%; 9% proteins play important roles in signal transduction such as ras-like GTP-binding protein RHO, hypoxanthine-guanine phosphoribosyltransferase and chordin-like protein; 9% proteins are related to protein metabolism as ubiquitin carboxyl-terminal hydrolase, receptor-type tyrosine-protein phosphatase S and F-actin-capping protein; 8% proteins take part in cell cycle and proliferation as kinesin-like protein KIF11, alpha-aminoadipic semialdehyde synthase and calcium-binding mitochondrial carrier protein aralar1; in addition, some other important proteins are related to stress response, DNA metabolism, RNA metabolism, cell adhesion, death, other metabolic processes and biological processes, which account for 3%, 3%, 6%, 2%, 2%, 7% and 16%, respectively. Fig. 5b shows the classification of the identified proteins according to cellular component. Membrane proteins are the major component in all the identified proteins except some unclassified proteins, which account for 25% in total including 6% of plasma membrane and 19% of other membrane proteins; 11% proteins make up of cytoskeleton and 17%, 5%, 4%, 2%, 3% proteins located in nucleus, ER/Golgi, mitochondrion, cytosol and non-structural extracellular, respectively. Classification by molecular function revealed a notable identification of proteins participating in transcription regulatory activity, enzyme regulator activity, transporter activity, signal transduction activity, nucleic acid binding activity and cytoskeletal activity, which account for 1%, 2%, 4%, 6%, 10% and 11%, respectively. However, approximately 66% protein functions had not been identified (Fig. 5c).

3.4.2. Pathways analysis

As is seen from Table 2, 33 pathways have been found out in the KEGG pathway database including tight junction, metabolic pathways, DNA replication, p53 signaling pathway, vascular smooth muscle contraction, chemokine signaling pathway, natural killer cell mediated cytotoxicity, B cell receptor signaling pathway, neurotrophin signaling pathway, leukocyte transendothelial migration pathway and so on. Approximately 18.18% (6/33) pathways of tight junction, regulation of actin cytoskeleton, DNA replication, p53 signaling pathway, adheres junction and lysine biosynthesis have got the enrichment $p < 0.05$, which means these results are very reliably in statistics.

3.4.3. Gene network analysis

It is clearly seen from Fig. 6, the gene network analysis was conducted including enzyme–enzyme relation, protein–protein interaction and gene expression interaction. 19.89% (36/181) genes had interaction and some relations were indicated from the experiment result such as TNPO1 and RAN; some relations were found in literatures such as PSMD1 and USP9X and some other relations were predicated from database like KIF11 and PRC1. Through relations exist among those genes; the connection degree varies significantly each other (Fig. 7). The higher connection degrees, which we call it hub gene, the more important of the gene is. Therefore, RAC1, KIF11 and MCM2 might play an important role in the network.

4. Discussion

The shotgun proteomic strategy based on digesting proteins into peptides and sequencing them using tandem mass spectrometry and automated database searching has become the

Table 2
Pathways analysis of the identified proteins.

KEGG pathways	No. of involved genes	KEGG pathways	No. of involved genes
Tight junction	4	Cell cycle	1
Regulation of actin cytoskeleton	3	Lysosome	1
Metabolic pathways	2	Endocytosis	1
DNA replication	2	Cardiac muscle contraction	1
p53 signaling pathway	2	Vascular smooth muscle contraction	1
Focal adhesion	2	Wnt signaling pathway	1
Adherens junction	2	Axon guidance	1
Leukocyte transendothelial migration	2	VEGF signaling pathway	1
Lysine biosynthesis	1	Toll-like receptor signaling pathway	1
Lysine degradation	1	Natural killer cell mediated cytotoxicity	1
Arachidonic acid metabolism	1	B cell receptor signaling pathway	1
ABC transporters	1	Fc epsilon RI signaling pathway	1
Proteasome	1	Fc gamma R-mediated phagocytosis	1
Nucleotide excision repair	1	Neurotrophin signaling pathway	1
Mismatch repair	1	Epithelial cell signaling in Helicobacter pylori infection	1
MAPK signaling pathway	1	Hypertrophic cardiomyopathy (HCM)	1
Chemokine signaling pathway	1		

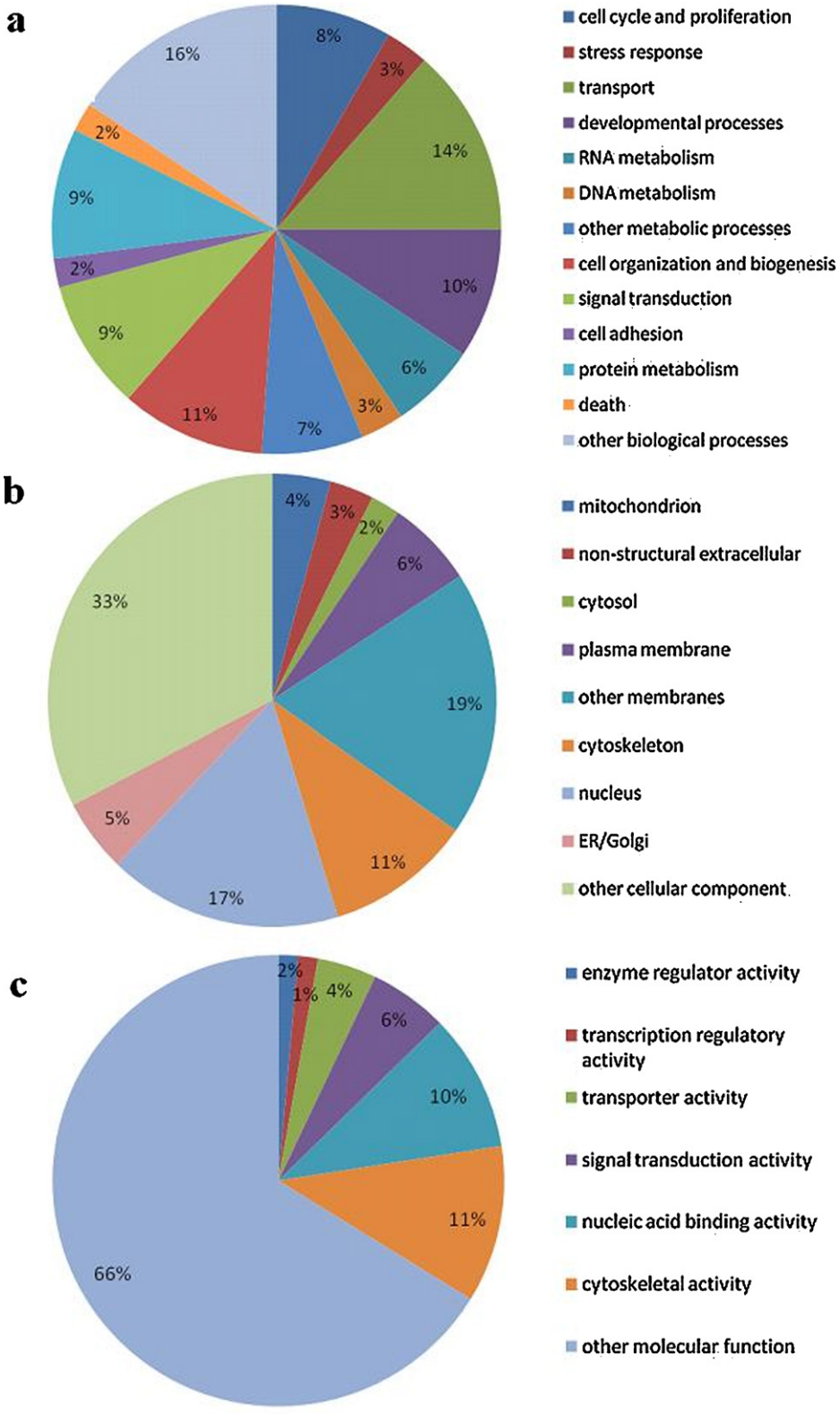

Fig. 5. Bioinformatics analysis by DAVID: (a) Classification of identified proteins in the nematocyst of jellyfish *Stomolophus meleagris* by biological process. (b) Classification of identified proteins in the nematocyst of jellyfish *Stomolophus meleagris* by cellular component. (c) Classification of identified proteins in the nematocyst of jellyfish *Stomolophus meleagris* by molecular function.

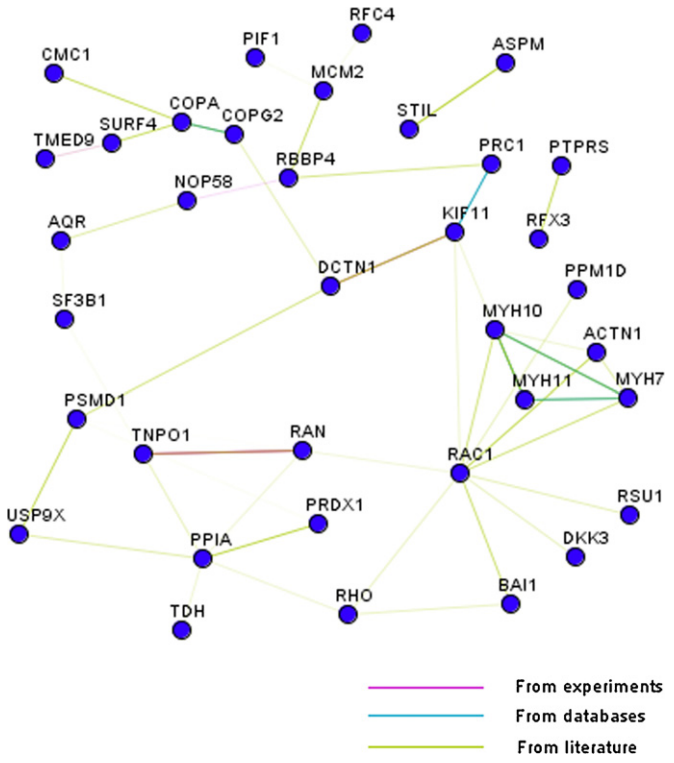

Fig. 6. Gene network analysis of the identified proteins.

method of choice for identifying proteins in most large scale studies [18]. In this study, we employed the method of shotgun to identify the nematocyst proteins of jellyfish *Stomolophus meleagris* by nanoLC–MS/MS. Shotgun proteomic analysis is proved to be a high sensitive and dynamic monitoring method compared with traditional two-dimensional electrophoresis and MALDI-TOF-mass spectrometry. It is also a simple method that we can enzymolysis the sample with the trypsin directly instead of two-dimensional electrophoresis and then the digested proteins are separated by two columns and monitored by auto MS/MS detection. Finally, all the MS spectra are analyzed by searching the database and the identified results can be visioned obviously.

In total, 181 proteins had been identified from the nematocyst proteins, including proteases like enolase, hypoxanthine phosphoribosyltransferase, phosphoribosylaminoimidazole carboxylase, protein phosphatase and so on. However, no known proteins like CAH1, CfTX-1, CfTX-2, CqTX-A and CrTX-A had been identified in the nematocyst proteins of jellyfish *Stomolophus meleagris*, which might be attributed to the low homologous between those bioactive proteins from different species and the undetectable amount of those bioactive proteins in the nematocyst proteins. Although shotgun proteomic analysis is a high sensitive and dynamic monitoring method, there must be some proteins unidentified in this procedure. The reasons might be account for as follows: firstly, there were only few studies about the nematocyst proteins of the jellyfish and very few proteins had been identified until now. Therefore, the database of jellyfish is very incomplete and many proteins are absent in the database. As a result, it is difficult to find out all digested proteins from the database. Secondly, the number of the identified proteins is related to the content and abundance of the nematocyst proteins. As is seen from the SDS–PAGE analysis (Fig. 2), there are not so many proteins in the nematocyst proteins that some trace proteins might not be detected in this procedure.

To gain deeply comprehensive insight and the possible interactions of the identified proteins, bioinformation of identified 181 proteins including gene ontology analysis, pathway analysis and gene network analysis had been analyzed. In gene ontology analysis, most of the proteins (84%) participate in the biological processes that have been known. However, it is difficult to identify other 16% proteins' biological process, which may be little studied before. In the analysis of cellular component, only 3% proteins were extracellular, which indicated that the process of preparing the nematocyst of jellyfish *Stomolophus meleagris* might be so successful that about 97% identified proteins are intracellular. However, 66% proteins could not be identified in the analysis of molecular function, which can be attributed to the function of these proteins undiscovered. The identified proteins were subjected to query against the KEGG reference pathway database and the largest proportions of pathways were related to signaling pathway, including p53, MAPK, chemokine, wnt, VEGF, toll-like receptor, epithelial cell, fc epsilon RI, neurotrophin, B cell receptor signaling pathway, which indicated that signaling proteins played a significant role in the nematocyst. In addition, two pathways were involved in muscle contraction of cardiac and vascular smooth. All these pathways also indicated that the nematocyst was a complex cell in the jellyfish of *Stomolophus meleagris*. The network analysis was based on the pathway of the KEGG database and 36 genes were predicated to be related to other genes. The degrees in Fig. 7 represented the conjunctions of gene, which mean that the higher of the degrees, the more important of the gene was. RAC1, the highest degree gene, had relations with other 10 genes of ACTN1, BAI1, DKK3, KIF11, MYH10, MYH7, PPM1D, RAN, RHO and RSU1, which indicated that RAC1 might play a very important role in the whole network.

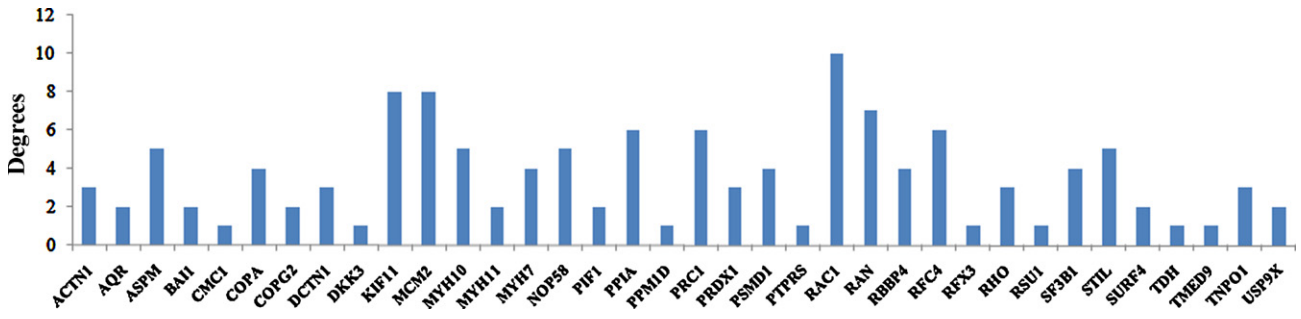

Fig. 7. The conjunction degree analysis of genes in the network.

5. Conclusions

In present study, we used the highly sensitive nanoLC–MS/MS method to analyze the nematocyst proteins of jellyfish *Stomolophus meleagris* for the first time, with a total of 181 proteins identified by this shotgun proteomics analysis. Bioinformatic analysis was employed to better understand the identified proteins including classification by biological process, cellular component and molecular function and construction of pathways and gene network analysis, which would be significant to the study of the nematocyst proteins of jellyfish *Stomolophus meleagris* in the future.

Conflict of interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors gratefully acknowledge financial grants: the National Natural Science Foundation of China (Grant No. 41006095), the National Science and Technology Support Program (2011BAE06B04-05), the Knowledge Innovation Program of the Chinese Academy of Sciences (KZC-YW-205-03) and Guangdong Province & Chinese Academy of Sciences overall strategy cooperative project (2010B080703027).

References

[1] H. Yu, R. Xing, S. Liu, C. Li, Z. Guo, P. Li, Int. J. Biol. Macromol. 40 (2007) 276.
[2] J. Feng, H. Yu, C. Li, R. Xing, S. Liu, L. Wang, S. Cai, P. Li, Toxicon 55 (2010) 118.
[3] R. Li, H. Yu, J. Feng, R. Xing, S. Liu, L. Wang, Y. Qin, K. Li, P. Li, Int. J. Biol. Macromol. 49 (2011) 14.
[4] D. Brinkman, J. Burnell, Toxicon 51 (2008) 853.
[5] H. Helmholz, S. Naatz, A. Prange, J. Chromatogr. B 871 (2008) 60.
[6] J. Sanchez-Rodriguez, E. Torrens, L. Segura-Puertas, Arch. Toxicol. 80 (2006) 163.
[7] H. Yu, X. Liu, X. Dong, C. Li, R. Xing, S. Liu, P. Li, Bioorg. Med. Chem. Lett. 15 (2005) 4949.
[8] C. Li, H. Yu, S. Liu, R. Xing, Z. Guo, P. Li, Bioorg. Med. Chem. Lett. 15 (2005) 5370.
[9] H. Yu, X. Liu, R. Xing, S. Liu, Z. Guo, P. Wang, C. Li, P. Li, Food Chem. 95 (2006) 123.
[10] K.L. Winter, G.K. Isbister, J.J. Schneider, N. Konstantakopoulos, J.E. Seymour, W.C. Hodgson, Toxicol. Lett. 179 (2008) 118.
[11] J.J. Chung, L.A. Ratnapala, I.M. Cooke, A.A. Yanagihara, Toxicon 39 (2001) 981.
[12] H. Nagai, K. Takuwa, M. Nakao, E. Ito, M. Miyake, M. Noda, T. Nakajima, Biochem. Biophys. Res. Commun. 275 (2000) 582.
[13] H. Nagai, K. Takuwa, M. Nakao, B. Sakamoto, G.L. Crow, T. Nakajima, Biochem. Biophys. Res. Commun. 275 (2000) 589.
[14] H. Nagai, K. Takuwa-Kuroda, M. Nakao, N. Oshiro, S. Iwanaga, T. Nakajima, Biosci. Biotechnol. Biochem. 66 (2002) 97.
[15] M.M. Bradford, Anal. Biochem. 72 (1976) 248.
[16] U.K. Laemmli, Nature 227 (1970) 680.
[17] R.C. Switzer, C.R. Merril, S. Shifrin, Anal. Biochem. 98 (1979) 231.
[18] A.I. Nesvizhskii, R. Aebersold, Mol. Cell. Proteomics 4 (2005) 1419.