IOP Conference Series: Earth and Environmental Science

PAPER • OPEN ACCESS

Molecular and adaptive evolution of *Nep2* gene from carnivorous plant *Nepenthes*

To cite this article: F Alamsyah and M Ito 2020 *IOP Conf. Ser.: Earth Environ. Sci.* **457** 012012

View the article online for updates and enhancements.

This content was downloaded from IP address 180.214.232.71 on 05/08/2020 at 11:07

Molecular and adaptive evolution of *Nep2* gene from carnivorous plant *Nepenthes*

F Alamsyah[1]* and M Ito[2]

[1]Department of Biology, Faculty of Science and Technology, University of Al Azhar Indonesia, Jl. Sisingamangaraja, Kebayoran Baru, Jakarta Selatan, 12110, Indonesia
[2]Department of General System Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Komaba Campus, Meguro-ku, Tokyo, 153-8902, Japan

*Corresponding author: firman.alamsyah@uai.ac.id

Abstract. *Nepenthes* get their nutrient by carnivory using their pitchers. A prey drowned in the pitcher fluid, will be digested by enzymes called nepenthesin, i.e. nepenthesin II. The structure of nepenthesin II-encoding gene might be related to the role of the enzyme. Therefore, the objective of this study was to examine the molecular and adaptive evolutions of *Nep2* gene expressing nepenthesin II. We analyzed 29 *Nepenthes* species that represent most habitat types. Total DNA was extracted from silica-dried leaf samples and amplification of *Nep2* gene was performed using degenerate primers. Homology searching was conducted using BLASTn, followed by computation of isoelectric point of the enzyme, and testing for positive selection using Mega 5. The result showed 29 DNA sequences of *Nep2* gene have no introns. Intron-less *Nep2* gene will produce nepenthesin II rapidly to digest the prey. The gene experienced significant positive selection on *N. sumatrana*, a species inhabits the lowest altitude habitats amongst Sumatran species. An obvious adaptive phenotype is the development of two unusual types of lower pitchers to obtain nutrient in lowland habitats. In conclusion, molecular and adaptive evolutions of *Nep2* gene characterized *Nepenthes* as highly adaptable plants that actively respond to the environmental stress.
Keywords: Nepenthes, Nep2 gene, molecular and adaptive evolution

1. Introduction

Carnivorous pitcher plants *Nepenthes* specifically evolved to inhabit marginal environments with nutrient deficiency [1-3], and to exploit niches where non-carnivorous plant species are less equipped to compete successfully. They augment their nutrient uptake by carnivory with their pitchers [4]. Morphological adaptations, including wetness-dependent peristome and slippery wax crystals, encourage the prey to fall into the pitcher [5-8]. Then the body of the prey is drowned by the pitcher fluid [9, 10] and subsequently digested by enzymes within the pitcher fluid [4].

Nepenthes secrete acid proteinases to digest the protein of their prey that trapped and drowned in their pitcher fluid [11] and absorb the digestion product as a nitrogen source [3, 12]. The acid proteinases inside the pitcher fluid were identified as nepenthesins, which were distinguished into nepenthesin I and II [13]. Both enzymes have optimal activity at acidic pH and are most stable at pH 3 [13, 14]. Moreover, nepenthesin I and II from *N. distillatoria* were quite different from each other in properties as they have different molecular masses and the activity and stability of both enzymes were different at certain temperature and pH. In addition, nepenthesin I and II from *N. gracilis* had only

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

66.6% identity of their amino acid sequences [13]. Despite their differences, these 2 acid proteinases are the only enzymes known to be specialized in prey digestion in the pitcher fluid of *Nepenthes* [14].

The nepenthesin is an aspartic proteinase (AP), making a family of protease enzymes that use an aspartate (Asp) residue for catalysis of their peptide substrates [13]. Aspartic proteinases are found widely in plants [15, 16] and other living organisms including animals, fungi, bacteria and viruses [17, 18]. In plants, APs are distributed in seeds, leaves and flowers [19], as well as in the digestive fluid of carnivorous plants and pitcher fluid of *Nepenthes* [13, 14, 20].

Some plant APs have been purified and well characterized, such as from barley called phytepsin [15, 21], and from rice called oryzasin [22]. Those 2 plant APs were identified as intracellular vacuolar enzymes and shared a plant-specific insertion sequence in the middle of their DNA sequences. In addition, the cloning of AP homologs from the pitcher tissue of *Nepenthes alata* that also belong to the vacuolar aspartic proteinase, contain a so-called plant-specific insertion [20]. In contrast, nepenthesins from *N. gracilis* do not have any plant-specific insertion sequences. Instead, they have a specific insertion namely the nepenthesin aspartic proteinase (NAP). Therefore, APs from the pitcher fluid of *Nepenthes* plants are clearly belong to a novel subfamily of APs [13]

The plant AP gene has undergone both the gain and loss of introns during molecular evolution. Rice AP oryzasin 1 gene comprises 14 exons and 13 introns [22]. Other plant AP gene from *Fagopyrum esculentum* (Genbank: AM422870), which is in the same order Caryophyllales with *Nepenthes*, also comprises the same number of exons and introns as oryzasin 1 gene. Non-plant AP genes have different composition of intron and exons, for instance, human cathepsin D [23], rat renin [24] and bovine chymosin [25] are all composed of 9 exons and 8 introns. Otherwise, yeast proteinase gene has no introns [26]. Therefore, the molecular evolution of APs as well as their structure-function relationships and physiological roles have become an interested field of study.

In the present study, we have tried to isolate the nepenthesin-encoding genes to study their molecular evolution and structure-function relationships, and to detect positive selection operates on the gene. For these purposes, we designed degenerate primers for the amplification of both nepenthesin I and II based on the alignment of sequences of the genes and their homologs available in the Genbank. However, we only succeeded to amplify the nepenthesin II-encoding gene, namely *Nep2*. The study of *Nep1* of *Nepenthes* will be conducted in the future.

2. Materials and methods

2.1. Plant materials

We analyzed 29 *Nepenthes* species in total (table 1) that represent most of habitat types and geographical areas, including Peninsular Malaysia and Thailand (3 species), Sumatera (12 species), Borneo (7 species), Philippines (3 species), Sulawesi (2 species with species *N. tentaculata* also distributed in Borneo), New Guinea (1 species), and outlying areas (1 species endemic to India).

2.2. Amplification and sequencing of Nep2 gene

Total DNA was extracted from silica-dried leaf samples with a QIAGEN DNeasy Mini Plant Kit (Qiagen) following the manufacturer's protocol. Amplification was performed using 2 pairs of degenerate primers (table 2). Ex-Taq buffer and Ex-Taq DNA polymerase (Takara Bio) were used for the amplification of the *Nep2* gene. The polymerase chain reaction (PCR) protocol consisted of an initial 90-s pre-denaturation at 96°C; 40 cycles of 45-s at 96°C (denaturation), 80-s at 58.5°C (annealing), and 70-s at 72°C (extension); and a final 7-min extension at 72°C.

The PCR products were cleaned using Wizard SV Gel and PCR Clean Up System (Promega) and were used for autocycle sequencing reaction following the manufacturer's (Beckman Coulter) instructions. Autocycle sequencing products were cleaned by ethanol precipitation. Both forward and reverse sequences were analyzed with a CEQ8000 automated sequencer (Beckman Coulter), using the same primers as for PCR. Six internal primers (table 2) were designed to get better sequences of the *Nep2* gene. The DNA genomic sequences of the *Nep2* gene were deposited in the Genbank (table 1).

IOP Conf. Series: Earth and Environmental Science **457** (2020) 012012 doi:10.1088/1755-1315/457/1/012012

Table 1. Plant materials examined in the study of *Nep2* gene. The materials were collected from the Indonesian Carnivorous Plant Society (*Komunitas Tanaman Karnivora Indonesia / KTKI*).

Species	Source	Genbank accession number of *Nep2*	Altitude (m) [4]
Nepenthes adnata Tamin & M. Hotta ex Schlauer	West Sumatra, cult. KTKI	AB769066	600 – 1200
Nepenthes alba Ridl.	cult. KTKI	AB769067	1600 – 2187
Nepenthes bellii K. Kondo	cult. KTKI	AB769068	0 – 800
Nepenthes campanulata Sh. Kurata	cult. KTKI	AB769069	300 – 500
Nepenthes chaniana C. Clarke, Chi. C. Lee & S. McPherson	cult. KTKI	AB769070	1100 – 1800
Nepenthes copelandii Merr. ex Macfarlane	cult. KTKI	AB769071	1400 – 1600
Nepenthes densiflora Danser	cult. KTKI	AB769072	1700 – 3200
Nepenthes diatas Jebb & Cheek	cult. KTKI	AB769073	2400 – 2900
Nepenthes ephippiata Danser	cult. KTKI	AB769074	1300 – 2000
Nepenthes faizaliana J.H. Adam & Wilcock	cult. KTKI	AB769075	400 – 1600
Nepenthes glabrata J.R. Turnbull & A.T. Middleton	cult. KTKI	AB769076	1600 – 2100
Nepenthes khasiana Hook. f.	cult. KTKI	AB769077	500 – 1500
Nepenthes lingulata Chi. C. Lee, Hernawati & Akhriadi	North Sumatra,cult. KTKI	AB769078	1700 – 2100
Nepenthes longifolia J. Nerz & Wistuba	cult. KTKI	AB769094	300 – 1100
Nepenthes naga Akhriadi, Hernawati, Primaldhi & M. Hambali	North Sumatra, cult. KTKI	AB769079	1500 – 2000
Nepenthes ovata J. Nerz & Wistuba	North Sumatra, cult. KTKI	AB769080	1700 – 2100
Nepenthes papuana Danser	Papua New Guinea, cult. KTKI	AB769081	0 – 1300
Nepenthes platychila Chi. C. Lee	cult. KTKI	AB769082	900 – 1400
Nepenthes rajah Hook. f.	Sabah, Borneo, cult. KTKI	AB769093	1500 – 2650
Nepenthes sanguinea Lindl.	cult. KTKI	AB769083	300 – 1800
Nepenthes spathulata Danser	Lampung, Sumatra, cult. KTKI	AB769084	1100 – 2900
Nepenthes spectabilis Danser	North Sumatra, cult. KTKI	AB769085	1400 – 2200
Nepenthes stenophylla Mast.	cult. KTKI	AB769086	800 – 2600
Nepenthes sumatrana (Miq.) Beck	North Sumatra, cult. KTKI	AB769087	0 – 800
Nepenthes talangensis J. Nerz & Wistuba	cult. KTKI	AB769088	1800 – 2500
Nepenthes tentaculata Hook. f.	cult. KTKI	AB769089	400 – 2550
Nepenthes thai Cheek	cult. KTKI	AB769090	500 – 600
Nepenthes tobaica Danser	North Sumatra, cult. KTKI	AB769091	380 – 1800
Nepenthes ventricosa Blanco	cult. KTKI	AB769092	1000 – 2000

Table 2. Primers used for amplification and sequencing of *Nep2* gene.

Primer name	Sequence (5'→ 3')
NepF1	GGYAGYGATCTYATYTGGAC
NepR2	ACYWGC RWKTYTTGCTGCTG
NepF5	ATGGCCTCRYCRCTRTAYTC
NepR6	YTACGACGCACCACAYTGAG
Internal primers	
NepF3	TRTCGCTTCCTTCTCAACTCG
NepR4	TGGATKAGGGTYGTAYTMGG
NepF7	CAAGCGTTCACTGACCAGA
NepR8	CTTGGCAATACTGGCTCTCGCA
NepF9	TTGATCTCTCCAGCTGAAGG
NepF10	GAGGAGCTCTGCAACATAGCA

2.3. Identification of the Nep2 gene and characterization of nepenthesin II

To identify the *Nep2* gene, homology searching was conducted using BLASTn program on NCBI server. To characterize the nepenthesins II, DNA sequences of the *Nep2* genes were translated into amino acid sequences using Expasy translate tool and followed by computation of isoelectric point (pI) of each enzyme on the Expasy server. The pI is the pH at which a certain molecule or surface holds no net electrical charge [27]. Other characterization of nepenthesin II including determination of prepro-nepenthesin II form, along with their active sites, cystein residues, N-glycosylation sites, acid and basic residues, and the NAP-specific insertion, that conducted by comparison with nepenthesin II from *N. gracilis* [13].

2.4. Test of positive selection on Nep2 gene

Testing for positive selection on *Nep2* gene by computing the average number of synonymous and nonsynonymous substitutions were performed with MEGA 5 [28] by using codon-based Z-test of selection based on Nei-Gojobori method [29] for sequence pairs, which involved 29 nucleotide sequences of the *Nep2* genes. All ambiguous positions were removed from each sequence pair.

The probability computed must be <0.05 for null hypothesis (strict neutrality) rejection at 5% level, where the number of synonymous substitutions per synonymous site (d_S) and the number of nonsynonymous substitutions per nonsynonymous site (d_N) are not the same, and the alternative hypothesis is $d_N>d_S$, indicated positive selection. The difference in synonymous and non-synonymous substitutions should be significant at the 5% level. For estimating the variance of the difference between d_N and d_S, the bootstrap method was conducted with 1000 replications.

3. Results

3.1. Nep2 genomic DNA sequences

In the present study, we have succeeded to amplify for the first time the genomic DNA sequences of APs *Nep2* gene for all species examined. All of the amplified DNA sequences are most similar to the aspartic proteinase *Nep2* cDNA of *N. gracilis* and *N. mirabilis* from the Genbank,with identities between 94%−98% and 93%−98%, respectively. The *Nep2* genes from 29 *Nepenthes* species varied in length between 1314−1317 bps, but mostly 1317 bp. Some *Nep2* genes with 1314 bp in length have a deletion in their propeptide (*N. khasiana*) or in enzyme (*N. stenophylla, N. tentaculata,* and *N. rajah*). Based on the alignment of 29 *Nep2* gene sequences together with the *Nep2* cDNA sequences from *N. gracilis* and *N. mirabilis*, all the DNA sequences of the *Nep2* genes show no introns (data not shown).

3.2. Nepenthesin II characterization

The prepro-form of nepenthesin II from 29 *Nepenthes* species, most composed of 438 amino acids, including 24 residues putative signal sequence, 55 residues putative propeptide, and 359 residues enzyme (figure 1). Nepenthesin II contain 12 cysteine residues per molecule protein, 2 active site sequence motifs: aspartic acid-threonine-glycine (D-T-G) and aspartic acid-serine-glycine (D-S-G); as well as the so-called flap tyrosine residue, assigned to residue 96 (figure 2). In addition, nepenthesin II appear to have 22 residues of NAP-specific insertion, assigned to residue 70-91 (figure 2), except for *N. stenophylla* that has a deletion within its NAP-specific insertion sequence. Moreover, nepenthesin II contains 1 or 2 potential N-glycosylation sites (table 3) and nepenthesin II also contains different number of acidic (aspartic acid and glutamic acid) and basic (histidine and arginine) residues (figure 2). The number of acidic and basic residues are between 28–35 residues and 1–4 residues, respectively (table 3).

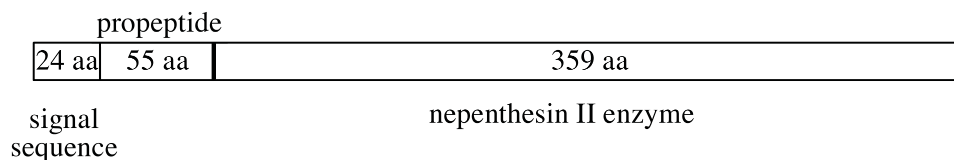

Figure 1. Illustration of the prepro-form of nepenthesin II, composed of 438 amino acids, including 24 residues putative signal sequence, 55 residues putative propeptide, and 359 residues enzyme.

A. QSSSGIETPVYAGDGAYLMNVSIGTPVSSFSAIMDTGSDLIWTQCEPCTQ : 50
CFSQPTPIFNPQDSSSFSTLPCDSQYCQDLPSETCNNNECQYTYGYGGGS : 100
TTQGYMATETFTFETGSVPNIAFGCGEDNQGFGQGNGAGLIGMGRGPLSL : 150
PSQLGVGQFSYCMTSYGSSSPSTLALGSAASGVPEGSPSTTLIHSSSIPT : 200
YYYITLQGITVGGDNLGIPSSTFRLQDDGTGGMIIDSGTTLTYLPQDAYY : 250
AVAQAFTDQINLSTVDESSSGLSTCFQQPSDGSTVQVPEISMQFDGGVLN : 300
LGEENVLISPAEGVICLAMGSSSQLGISIFGNIQQQETQVLYDLQNSAVS : 350
FVPAQCGAS : 359

B. QSSSGIQTPVYAGNGEYLMNASIGTPASSFSAIMDTGSDLIWTQCEPCTQ : 50
CFSQPTPIFNPQDSSSFSTLPCESQYCQDLPSETCNNNECQYTYGYGDGS : 100
STQGYMATETFTFETGSVPNIAFGCGQDNQGFGQGNGAGLIGMGWGPLSL : 150
PSQLGVGQFSYCMTSYGSSSPSTLALGSAASGVPEGSPSTTLIHSSPNPT : 200
YYYITLQGITVGGDNLGIPSSTFQLQGDGTGGMIIDSGTTLTYLPQDAYN : 250
AVAQAFTDQIHLSTVE-SSSGLNTCFQQPSDGSSVQVPEISMQFDGGVLN : 299
LGEENILISPAEGVICLAMGSSSQQGISIFGNIQQQETLVLYDLQNVAVS : 349
FVPTQCGAS : 358

Figure 2. Representative of amino acid sequences of nepenthesin II enzymes from *N. sanguinea* (A) and *N. tentaculata* (B). Twelve cysteine (C) residues, 2 active sites (DTG and DSG), and a flap tyrosine (Y) residue are highlighted in yellow, green, and pink, respectively; Twenty-two residues of NAP-specific insertion are marked with blue line, the acidic (D and E) and basic (H and R) residues are in red and green font color, respectively; and the N-glycosylation sites (NVS, NLS, and NAS) are underlined.

3.3. Synonymous and nonsynonymous substitutions

There was a total of 438 positions in the final dataset of *Nep2*. Table 4 shows the relative abundance of synonymous and nonsynonymous substitutions that have occurred in DNA sequences of the *Nep2*

genes of all sequence pairs by using codon-based Z-test of selection. It shows that the probability (P) of rejecting the null hypothesis of strict neutrality ($d_N = d_S$) in favor of the alternative hypothesis of positive selection ($d_N > d_S$) with values of P less than 0.05, correspond to several of all sequence pairs involving *N. sumatrana*. This result suggests that the evolution of the *Nep2* gene from *N. sumatrana* has been under positive selection.

Table 3. The numbers of N-glycosylation sites, acidic and basic residues, and the isoelectric point of 29 *Nepenthes* species.

No.	Species	N-glycosylation site	Acidic residues	Basic residues	Isoelectric point (pI)
1.	*N. khasiana*	1	30	4	3.45
2.	*N. papuana*	1	32	4	3.44
3.	*N. adnata*	2	32	4	3.42
4.	*N. bellii*	2	30	3	3.34
5.	*N. sumatrana*	1	30	3	3.33
6.	*N. sanguinea*	2	31	3	3.32
7.	*N. stenophylla*	1	33	3	3.29
8.	*N. tentaculate*	1	28	2	3.24
9.	*N. ventricosa*	1	29	2	3.21
10.	*N. glabrata*	2	30	2	3.19
11.	*N. tobaica*	1	32	2	3.18
12.	*N. longifolia*	1	33	2	3.15
13.	*N. rajah*	1	29	1	3.03
14.	*N. densiflora*	2	30	1	3.00
15.	*N. naga*	2	31	1	3.00
16.	*N. spathulate*	1	31	1	2.99
17.	*N. faizaliana*	2	31	1	2.99
18.	*N. spectabilis*	1	31	1	2.98
19.	*N. ovata*	2	32	1	2.98
20.	*N. lingulata*	2	32	1	2.98
21.	*N. platychila*	2	32	1	2.98
22.	*N. campanulate*	2	32	1	2.98
23.	*N. copelandii*	1	31	1	2.98
24.	*N. chaniana*	2	32	1	2.97
25.	*N. thai*	1	32	1	2.97
26.	*N. alba*	2	32	1	2.97
27.	*N. talangensis*	1	32	1	2.97
28.	*N. diatas*	2	32	1	2.96
29.	*N. ephippiata*	1	35	1	2.93

In contrast, the Z values calculated from the dN – dS of some sequence pairs involving *N. chaniana*, *N. naga*, *N. spathulata*, *N. stenophylla*, *N. glabrata*, *N. densiflora*, *N. lingulata*, *N. platychila*, *N. diatas*, and *N. ventricosa*, were positive, but the corresponding P values were above 0.05. This result suggests that the evolution of the *Nep2* genes of that 10 *Nepenthes* species have been under strong purifying selection.

IOP Conf. Series: Earth and Environmental Science **457** (2020) 012012 doi:10.1088/1755-1315/457/1/012012

Table 4. The probabilities (P) computed from codon-based Z-test of positive selection (below diagonal) for analysis between sequences of Nep2 gene. P<0.05 for null hypothesis (strict neutrality) rejectionat 5% level are highlighted. The test statistic (dN - dS) is shown above the diagonal. dS and dN are the numbers of synonymous and nonsynonymous substitutions per site, respectively.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
N. faizaliana (1)		-2,93	-2,24	-2,47	-2,95	-1,72	-2,61	-2,51	-2,13	-2,36	-1,92	-3,33	-2,16	-2,14	-1,90	-2,31	-2,91	-2,55	-2,26	-3,02	-3,10	-2,10	-2,77	-1,25	-3,04	-3,00	-1,29	-2,83	-2,38
N. ephippiata (2)	1,00		-2,27	-2,78	-3,24	-2,09	-2,25	-2,20	-1,14	-3,40	-2,51	-3,60	-2,77	-3,27	-2,02	-1,93	-2,79	-2,52	-2,53	-3,31	-2,93	-2,55	-2,45	-1,12	-3,11	-3,11	-1,46	-2,86	-2,65
N. bellii (3)	1,00	1,00		-1,71	-2,00	-1,02	-1,11	-1,38	-1,02	-1,92	-1,29	-3,60	-0,96	-2,27	-1,40	-1,24	-1,33	-2,53	-0,53	-3,42	-2,09	-2,27	-1,92	-0,91	-1,75	-3,00	-0,35	-1,90	-1,76
N. adnata (4)	1,00	1,00	1,00		-1,71	-1,36	-1,62	-1,13	-1,90	-2,69	-0,61	-3,53	-1,28	-1,68	-1,87	-1,01	-1,25	-1,99	-2,17	-3,21	-1,51	-2,36	-0,27	-0,72	-1,33	-1,97	-0,41	-1,41	-1,29
N. spectabilis (5)	1,00	1,00	1,00	1,00		-1,68	-0,95	-1,34	-2,15	-3,02	-1,85	-3,91	-1,89	-2,56	-2,11	-1,81	-1,99	-2,69	-2,19	-4,06	-0,82	-3,26	-2,61	-1,92	-0,88	-2,99	-0,94	-2,31	-0,97
N. chaniana (6)	1,00	1,00	1,00	1,00	1,00		-0,74	-0,34	-0,61	-1,62	0,94	-3,42	0,08	-1,35	-1,17	0,91	-0,87	-0,92	-1,08	-3,16	-1,50	-3,02	-0,55	0,55	-2,03	-2,00	2,91	-0,81	-1,56
N. naga (7)	1,00	1,00	1,00	1,00	1,00	1,00		-0,61	-1,29	-2,42	-0,91	-3,51	-0,85	-1,95	-1,35	-0,94	-0,81	-2,32	-2,32	-3,54	-1,29	-3,28	-0,63	0,12	-1,50	-1,91	-0,56	-1,10	-1,30
N. spathulata (8)	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-1,02	-2,34	-0,03	-3,19	-0,45	-1,32	-1,31	0,17	-1,13	-1,60	-2,16	-3,38	-0,81	-2,89	-0,34	-0,49	-1,48	-2,27	1,29	-1,05	-1,48
N. stenophylla (9)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-2,08	-0,04	-4,15	-0,62	-1,34	-1,46	0,08	-1,63	-2,29	-1,67	-3,73	-1,79	-1,79	-1,13	0,15	-2,25	-2,42	1,06	-1,42	-1,96
N. rajah (10)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-1,64	-4,72	-2,07	-2,71	-2,48	-2,10	-2,48	-3,35	-2,66	-4,06	-3,24	-3,26	-2,26	-1,72	-3,36	-2,69	-1,17	-2,66	-2,84
N. glabrata (11)	1,00	1,00	1,00	1,00	1,00	0,17	1,00	1,00	1,00	1,00		-3,55	-0,17	-1,66	-1,65	0,45	-0,26	-1,77	-1,29	-3,71	-1,84	-2,84	0,21	0,44	-2,39	-2,00	1,73	-1,02	-2,07
N. khasiana (12)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-3,72	-4,06	-3,19	-3,36	-4,00	-3,26	-4,26	-3,19	-3,72	-4,22	-3,61	-2,85	-3,58	-3,88	-3,08	-3,81	-3,53
N. densiflora (13)	1,00	1,00	1,00	1,00	1,00	0,47	1,00	1,00	1,00	1,00	1,00	1,00		-2,13	-1,40	-0,88	-0,18	-2,35	-1,68	-3,71	-1,75	-2,92	-1,69	-0,59	-1,77	-2,82	0,69	-1,82	-1,50
N. tentaculata (14)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-2,12	-1,59	-1,64	-2,24	-2,21	-3,69	-2,25	-2,49	-1,91	-1,04	-2,41	-3,08	-0,86	-2,46	-2,10
N. talangensis (15)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-1,56	-1,52	-1,33	-1,51	-3,73	-2,78	-3,03	-1,40	-0,60	-2,00	-1,65	-0,70	-1,08	-1,48
N. lingulata (16)	1,00	1,00	1,00	1,00	1,00	0,18	1,00	0,43	0,47	1,00	0,33	1,00	1,00	1,00	1,00		-0,70	-1,48	-1,84	-3,79	-1,64	-2,88	-0,50	0,33	-2,17	-2,57	2,31	-1,25	-1,80
N. platychila (17)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-1,99	-1,87	-3,09	-1,84	-2,66	-0,59	-0,03	-1,92	-2,21	0,76	-1,49	-1,46
N. tobaica (18)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-2,34	-3,48	-2,40	-2,89	-1,63	-1,18	-2,19	-2,43	-1,33	-2,05	-1,96
N. sanguinea (19)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-3,31	-2,25	-2,88	-2,20	-1,34	-1,62	-1,93	-0,72	-2,16	-1,79
N. papuana (20)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-4,09	-3,71	-3,70	-2,83	-4,25	-3,50	-3,70	-4,12	-3,15
N. alba (21)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-3,40	-1,74	-1,31	-1,69	-3,10	-0,86	-2,48	-1,10
N. copelandii (22)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-2,95	-0,91	-3,65	-3,40	-2,49	-3,36	-2,62
N. diatas (23)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,42	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-0,40	-1,78	-2,32	0,22	-1,33	-1,32
N. ventricosa (24)	1,00	1,00	1,00	1,00	1,00	0,29	0,45	1,00	0,44	1,00	0,33	1,00	1,00	1,00	1,00	0,37	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-1,59	-2,06	0,95	-0,37	-1,15
N. thai (25)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-2,16	-1,66	-2,04	-0,35
N. campanulata (26)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-1,45	-2,87	-1,78
N. sumatrana (27)	1,00	1,00	1,00	1,00	1,00	**0,00**	1,00	0,10	0,15	1,00	**0,04**	1,00	0,25	1,00	1,00	**0,01**	0,22	1,00	1,00	1,00	1,00	1,00	0,41	0,17	1,00	1,00		-0,04	-1,12
N. ovata (28)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		-1,72
N. longifolia (29)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	

IOP Conf. Series: Earth and Environmental Science **457** (2020) 012012 doi:10.1088/1755-1315/457/1/012012

4. Discussion

4.1. Structural and functional relationships of the Nep2 gene

In the present study, it is revealed that the *Nep2* genes from 29 *Nepenthes* species have no introns. This is the first report presenting structural feature of nepenthesin-encoding gene from *Nepenthes* species. This structural feature may have relationship with the function of the gene. Genes whose expression levels are changed promptly in response to environmental stress have significantly lower intron densities in some eukaryotes. In addition, introns could slow down the regulatory responses and were selected in genes whose transcripts need fast adjustment for survival from environmental stress [30].

In some cases, transcription occurs at 1200-1500 nucleotides per minute [31] with half-lives for splicing reactions are less than 1 minute for the first intron, but 2–8 minutes for the subsequent introns [30, 32]. Consequently, splicing of 2 or more introns takes longer time than the transcription itself [30]. Therefore, intron-less *Nep2* gene will produce its protein rapidly for digesting the trapped prey. This rapid production of nepenthesin II enzyme may help to avoid putrefaction of trapped prey, which resulted in an accumulation of ammonium that may harm the pitcher to die [4]. This result is concordant to the result of immuno-histochemical staining of *Nepenthes* pitcher tissue, which indicated that nepenthesins were directly secreted into the pitcher fluid and functioned without accumulation in the pitcher tissue [13]. The rapid production of nepenthesin II enzyme is also corroborated by the small quantity of fluid contained in newly opened pitchers, which usually less than 1/6 of the total volume of the mature one [4]. Thus, *Nep2* gene is supposed to have adapted specifically to produce extracellular nepenthesin II digestive enzymes rapidly by removing its introns during the process of molecular evolution.

As an extracellular proteinase, nepenthesin II of the genus *Nepenthes* is synthesized in the endoplasmic reticulum (ER), travel to Golgi apparatus and then to plasma membrane for secretion. This route is known as the secretory pathway. The signal sequence of nepenthesin II is recognized by specific cellular components that facilitate the proper routing of that protein. As synthesized in the ER and secreted via Golgi apparatus and plasma membrane, the signal sequence of nepenthesin II is included in the ER signal type, which is usually located near the amino terminus [33]. As an ER signal type, the signal sequence of nepenthesin II is composed of mostly (67%) nonpolar amino acids: methionine (M), alanine (A), valine (V), glycine (G), leucine (L), isoleucine (I), and proline (P) (figure 3).

Based on the alignment of amino acid residues of nepenthesin II signal sequences from 29 *Nepenthes* species, there are some substitutions within the signal sequences, including valine (V) to leucine (L) and alanine (A) to glycine (G) (*N. khasiana*), valine (V) to alanine (A) (*N. longifolia*), leucine (L) to glycine (G) (*N. tobaica*), leucine (L) to valine (V) (*N. papuana*), and glycine (G) to alanine (A) (*N. sanguinea*) (figure 3). However, those substitutions of amino acids were expected to have no change of the protein properties, since the substituted amino acids are also nonpolar.

All the enzyme of nepenthesin II examined contain 12 cysteine residues, which would form 6 disulphide bonds expected to contribute greatly to the stability of the enzyme [13]. Moreover, the stabilized structure with 6 disulphide bonds allows the protein to be resistant to protease degradation [34]. These structures suggest that nepenthesin II enzyme can remains in the pitcher fluid without digestion [14], which indicated by the 85% of the original activity after 30 days at pH 3 [13].

```
N.adnata      : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.bellii      : MASSLYSVVLGLAIVSAIVAPTCS : 24
N.platychila  : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.khasiana    : MASSLYSVVLGLAILSAIVGPTSS : 24
N.ovata       : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.densiflora  : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.diatas      : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.lingulata   : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.talangensis : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.sumatrana   : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.longifolia  : MASSLYYAVLGLAIVSAIVAPTSS : 24
N.campanulata : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.tentaculata : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.naga        : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.spectabilis : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.tobaica     : MASSLYSVVLGGAIVSAIVAPTSS : 24
N.ephippiata  : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.glabrata    : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.copelandii  : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.thai        : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.faizaliana  : MASSLYTVVLGLAIVSAIVAPTSS : 24
N.papuana     : MASSLYSVVLGVAIVSAIVAPTSS : 24
N.chaniana    : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.alba        : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.spathulata  : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.stenophylla : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.sanguinea   : MASSLYSVVLALAIVSAIVAPTSS : 24
N.rajah       : MASSLYSVVLGLAIVSAIVAPTSS : 24
N.ventricosa  : MASSLYSVVLGLAIVSAIVAPTCS : 24
                MAsSLY VVLG AI6SAIVAPTsS
```

Figure 3. Alignment of nepenthesin II signal sequences from 29 *Nepenthes* species. Nonpolar amino acids are highlighted in red, and amino acid substitutions are in white font color.

All nepenthesin II of 29 *Nepenthes* species contain an approximately 22 residues of NAP-specific insertion, preceding the flap tyrosine residue (figure 2). This insertion contains 4 cysteine residues as well as 4 acidic residues, except for *N. sanguinea*, *N. khasiana*, *N. campanulata*, *N. faizaliana*, and *N. spathulate* that have only 3 acidic residues. In addition, *N. copelandii* and *N. bellii* have only 2 acidic residues within their NAP-specific insertion. Overall, the differences of acidic and basic residues number would determine the pI of the enzymes which vary among nepenthesins II, with the highest point reach pH 3.45 (*N. khasiana*) and the lowest one at pH 2.95 (*N. thai*) (table 3). The sequences of NAP-specific insertion of nepenthesins II are not conserved, since some substitutions appeared within the sequence insertions of some *Nepenthes* species (figure 2). On the contrary, the flap tyrosine residues following the NAP-specific insertions, as well as the 2 active site motifs: aspartic acid-threonine-glycine (D-T-G) and aspartic acid-serine-glycine (D-S-G), are conserved among the 29 nepenthesin II enzymes (figure 2).

During the synthesis of prepro-nepenthesin II in the ER and following the travel to Golgi apparatus, the enzyme is attached with carbohydrate in the process of glycosylation. Since nepenthesin II has the N-glycosylation motif site(s), the carbohydrate, which is oligosaccharide chains, is attached to a nitrogen of asparagine (N) side chains, in the sequence motif of asparagine-leucine-serine (N-L-S) and asparagine-valine-serine (N-V-S), within the nepenthesin II sequence of most *Nepenthes* species. In addition, the sequence motif of asparagine-alanine-serine (N-A-S) [33], which resulted from the

substitution of valine (V) to alanine (A), could act as an N-glycosylation site within the nepenthesin II sequence from *N. tentaculata* (figure 2).

4.2. Adaptive evolution of nepenthesin II enzyme

The present study is the first study that concern to the adaptive evolution on *Nep2* gene by using statistical analysis based on the relative abundance of synonymous and nonsynonymous substitutions. Adaptive evolution after gene duplication has been reported in several gene families [34, 35]. In this study, *Nep2* gene, a member of a gene family of aspartic proteinase, suggestively experienced significant positive selection on *N. sumatrana* (table 4), a species inhabits the lowest altitude habitats (0-800 m) amongst Sumatran endemic species, which most of them are highland species [4]. Adaptation to lowland habitats in Sumatra would be influenced by multiple physiological factors and genetic factors. For instance, at the physiological level, an obvious adaptive phenotype is the development of 2 distinct types of lower pitchers of *N. sumatrana*. Lower pitchers of the first type are beared from seedlings and juvenile plants with wholly or partially ovate form. While, lower pitchers of the second type are produced by a basal offshoot developing from the rootstock and have many squat lower pitcher forms than the first type [4].

The function of both types of lower pitchers are to trap creeping insects [4]. Most of the nitrogen sources of some *Nepenthes* species inhabit lowland habitats were provided by ants [12, 36-38]. In our previous study [39], *N. sumatrana* was in the same subclade with 13 other Sumatran endemic species and had the closest relationship with *N. spathulata* and *N. tobaica*. *N. spathulata* is known as a highland species, and *N. tobaica* has distribution area from lowland to highland (table 1). Interestingly, both *N. spathulata* and *N. tobaica* have narrow peristomes on their upper pitchers, whereas *N. sumatrana* has broad peristome [4]. This broad peristome on upper pitchers of *N. sumatrana* may help the species to get more nutrients from flying insect than the two closest species, since it has wider slippery surface [6]. Therefore, the 2 distinct types of lower pitchers of *N. sumatrana* and the broad peristome of their upper pitchers are mirrored by the strategy employed to obtain nutrients from the trapped prey, and the abundance of nutrient uptake should be correlated to the fitness of the species in the lowland habitats. The 2 distinct types of lower pitchers of *N. sumatrana* are not found among other lowland *Nepenthes* species as well as the highland species [4]. These 2 distinct types of lower pitchers reveal that *N. sumatrana* has developed specific adaptation in response to nutrient stress that characterize the habitat where it grows. Thus, they may demonstrate that *N. sumatrana* is under selective pressure of prey and environment.

5. Conclusion

Molecular and adaptive evolutions of *Nep2* gene characterized *Nepenthes* as highly adaptable plants that actively respond to the environmental conditions and availability of prey in their habitats.

Acknowledgements

We wish to thank the Indonesian Carnivorous Plants Society (Komunitas Tanaman Karnivora Indonesia/KTKI) for kindly providing all *Nepenthes* materials used in this study. We also thank the LP2M of University of Al Azhar Indonesia for providing financial assistance for the presentation of this study at the lnternational Conference on Biosciences (ICoBio) 2019, in Bogor, Indonesia. We wish to express our gratitude to Shin-ichi Morinaga, Seishiro Aoki, Toshihide Kato, and Nozomu Utsuki for technical assistance during this study. This study was partly supperted by the Environment Research and Technology Development Fund (D-1008) from MOE, Natonal Bioresouce Project and GRENE environmental information from MEXT, and a Grant-in-Aid for scientific research from JSPS (22405012) to Ml.

References

[1] Ellison A M 2006 *Plant. Biol.* **8** 740
[2] Osunkoya O O *et al*. 2007 *Ann. Bot.* **99** 895

[3] Schultze W *et al*. 1997 *Oecologia* **112** 464
[4] McPherson S 2009 *Pitcher plants of the old world 1 and 2* (Redfern Natural History Productions Poole)
[5] Bauer U and Federle W 2009 *Plant Signal. Behav.* **4** 1019
[6] Bauer U *et al*. 2012 *J. Evol. Biol.* **25** 90
[7] Bohn H F and Federle W 2004 *Proc. Natl. Acad. Sci.* **101** 14138
[8] Gaume L and Di Giusto B 2009 *Ann. Bot.* **104** 1281
[9] Di Giusto B *et al*. 2008 *J. Biosci.* **33** 121
[10] Gaume L and Forterre Y 2007 *PLoS One* **2** 1
[11] Tokes Z A *et al*. 1974 *Planta* **119** 39
[12] Moran J A *et al*. 2001 *Ann. Bot.* **88** 307
[13] Athauda S B P *et al*. 2004 *Biochem. J.* **381** 295
[14] Hatano N and Hamada T 2008 *J. Proteome Res.* **7** 809
[15] Kervinen J *et al*. 1999 *The EMBO J.* **18** 3947
[16] Simoes I and Faro C 2004 *Eur. J. Biochem.* **271** 2067
[17] Davies D R 1990 *Annu. Rev. Biophys. Biophys. Chem.* **19** 189
[18] Rawlings N D and Barret A J 1995 *Methods Enzymol.* **248** 105
[19] Mutlu A and Gal S 1999 *Physiol. Plantarum* **105** 569
[20] An C I *et al*. 2002 *Planta* **214** 661
[21] Roos P R *et al*. 1991 *Eur. J. Biochem.* **202** 1021
[22] Asakura T *et al*. 1995 *Eur. J. Biochem.* **232** 77
[23] Redecker B *et al*. 1991 *DNA Cell Biol.* **10** 423
[24] Fukamizu A *et al*. 1988 *J. Mol. Biol.* **201** 443
[25] Hidaka M *et al*. 1986 *Gene* **43** 197
[26] Woolford C A *et al*. 1986 *Mol. Cell Biol.* **6** 2500
[27] Sakakibara Y and Yanagisawa H 2007 *Bull. Aichi Univ. Edu.* **56** 45
[28] Tamura K *et al*. 2011 *Mol. Biol. Evol.* **28** 2731
[29] Nei M and Gojobori T 1986 *Mol. Biol. Evol.* **3** 418
[30] Jeffares D C *et al*. 2008 *Trends Genet.* **24** 375
[31] Izban M G and Luse D S 1992 *J. Biol. Chem.* **267** 13647
[32] Audibert A *et al*. 2002 *Mol. Cell. Biol.* **22** 6706
[33] Brooker R J, Widmaier E P, Graham L E and Stilling P D 2008 *Biology* (McGraw Hill New York) pp. 77-120
[34] Selitrennikoff C P 2001 *Appl. Environ.Microbiol.* **67** 2883
[35] Hughes A L 2002 *Trends Genet.* **18** 433
[36] Adam J H 1997 *Pertanika J. Trop. Sci.* **20** 121
[37] Bazile V *et al*. 2012 *PLoS One.* **7** 1
[38] Moran J A and Moran A J 1998 *Int. J. Plant Sci.* **159** 996
[39] Alamsyah F and Ito M 2013 *Acta Phytotax. Geobot.* **64** 113