

Current Trends in OMICS (CTO)

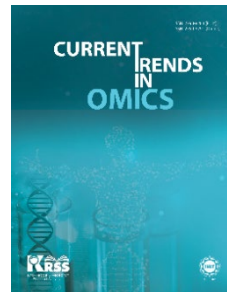
Volume 4 Issue 2, Fall 2024


ISSN(P): 2790-8283, ISSN(E): 2790-8291

Homepage: <https://journals.umt.edu.pk/index.php/cto>



Article QR



- Title:** Evolutionary Dynamics of Polintons: A Rarely Investigated Superfamily of DNA Transposons
- Author (s):** Muhammad Umair Khan¹, Faisal Nouroz¹, Amna Waheed¹, and Shumaila Noreen²
- Affiliation (s):** ¹Department of Bioinformatics, Hazara University Mansehra, Mansehra, Pakistan
²Department of Zoology, Hazara University Mansehra, Mansehra, Pakistan
- DOI:** <https://doi.org/10.32350/cto.42.02>
- History:** Received: June 06, 2023, Revised: August 23, 2024, Accepted: September 24, 2024, Published: October 10, 2024
- Citation:** Khan MU, Nouroz F, Waheed A, Noreen S. Evolutionary dynamics of polintons: a rarely investigated superfamily of DNA transposons. *Curr Trend OMICS*. 2024;4(2):21–35. <https://doi.org/10.32350/cto.42.02>
- Copyright:** © The Authors
- Licensing:**  This article is open access and is distributed under the terms of [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)
- Conflict of Interest:** Author(s) declared no conflict of interest



A publication of
The Department of Life Sciences, School of Science
University of Management and Technology, Lahore, Pakistan

Evolutionary Dynamics of Polintons: A Rarely Investigated Superfamily of DNA Transposons

Muhammad Umair Khan¹, Faisal Nouroz^{1*}, Amna Waheed¹, and Shumaila Noreen²

¹Department of Bioinformatics, Hazara University Mansehra, Pakistan

²Department of Zoology, Hazara University Mansehra, Pakistan

ABSTRACT

Polintons, also known as ‘Maverick’ is a rarely investigated superfamily of DNA transposons. Although, discovered in recent decades, these elements are now found in several eukaryotic genomes due to the advancements and expansions in genome sequencing projects. The current study was conducted to investigate the evolutionary genomics and distribution of Polintons in various genomes. Approximately, 102 Polinton elements from various species were collected from Replibase – GIRI database. They have shown a high structural diversity and ranged in sizes from 03 to 42kb. The phylogenetic relationship and evolutionary history of the Polinton superfamily was investigated using bioinformatics software. Phylogenetic trees were created based on two major protein coding domains, such as RVE integrase and DNA POL B. In both cases, the sequences were clustered in 2 clades, several sub-clades, and groups with variable number of elements in each group. The cladograms represented the extensive genetic diversity and evolutionary history of these elements and allowed to observe the intricate branching patterns and relationships within the Polinton superfamily. The cladogram provides a roadmap to explore the relationships and dynamics of the Polinton superfamily in detail. The current study aimed to enhance the understanding of evolutionary dynamics of Polintons found in different organisms.

Keywords: Evolution, phylogeny, Polinton, RVE integrase, transposons

1. INTRODUCTION

The transposons or the jumping genes are DNA sequences, which can move and change their position within the genomes. Due to their mobility within the genomes, they may result in genetic diversity, mutations, duplications, and alterations in the gene expression. Therefore, they are the

*Corresponding Author: faisalnouroz@gmail.com

main drivers of genome evolution and duplication. Moreover, they are the abundant elements proliferating in almost all genomes ranging from bacteria to human beings and constitute a major proportion of organism genomes. Due to their diversity and mobility, transposons are important factors which revolutionized the molecular and genetic fields. They are classified into two major classes based on their mode of transposition, such as Class I (Retrotransposons) and Class II (DNA transposons) having further sub-classes and superfamilies with variable structures (Fig. 1). In Class I elements, the Retrotransposons are further sub-divided into LTR Retrotransposons (Copia, Gypsy, Retriviruses superfamilies) and Non-LTR Retrotransposons (LINEs & SINEs). They follow the copy and paste mechanism of transposition, where DNA is first transcribed into RNA template and reverse transcribed into cDNA involving the reverse transcriptase (RT). In contrast, in Class II elements, the DNA transposons follow the cut and paste mechanism, where the DNA element integrates itself into a new site within the genome in presence of transposase without making their copy. DNA transposons are further classified into several superfamilies, which can be differentiated based on their target site duplications (TSDs), terminal inverted repeats (TIRs), transposase domain, and internal sequences. Among DNA transposons, the most active and investigated superfamilies include Ac/Ds or hAT, CACTA, TC1-Mariner, Mutator, Merlin, PIF-Harbinger, Transib, and PiggyBac. On the other hand, the rarely investigated superfamilies include the Helitron, Polinton, and Crypton. The DNA transposons have encoded a transposase domain necessary for their transposition, while the structure of Polintons/Maverick is different from the other DNA transposons. They encode various domains, where retroviral-like integrase (INT), putative ATPase, adenoviral-like protease, and DNA polymerase B (POLB) are important domains [1–6].

Of DNA transposons, Polinton or Maverick emerged within various genomes and were initially discovered in protist genomes, however, they were later identified in other organisms. They are not as abundant as other DNA transposons, however, still proliferating in several host genomes. They are the largest sized transposons among various DNA transposon superfamilies ranging from 3 to 42kb. They show structural complexity and sophistication as compared to other DNA transposon superfamilies. They exhibit various domains involved in various cellular processes, such as protein synthesis and DNA replication. They encode a protein primed type B DNA polymerase (POL B) and a retroviral like (RVE) family. Recent

investigations have shown that Polinton also harbors few virus-like domains, such as major capsid protein (MCP), minor capsid protein (mCP), and a DNA packaging protein ATPase, suggesting their evolutionary relationship with viruses. The structural diversity and presence of Polintons in various sequenced genomes have shown a significant role in shaping the development and evolution of eukaryotic genomes. Based on their sizes and structural architecture, the Polintons show structural diversity with the virophages, the viruses infecting the microbial eukaryotes [6–10].

Polintons are rarely investigated and were identified from few organism genomes. Their study and investigations have raised several questions regarding their origin, biological significance, and evolutionary history. Several hypotheses have been proposed to explain the origin of Polintons including a potential ancestral relationship with viruses or independent evolutionary events. Exploring the mysteries of Polintons would provide a major dynamic interplay between Polintons and their host genomes. Polintons have not been properly investigated, however, their presence in various genomes clarifies that they play a major role in host genome evolution and their plasticity [11–13].

The current study aimed to explore the structural diversity, evolutionary relationship, sizes, and identification of Polintons in various organisms. Notably, the study explored structural genomics and evolutionary dynamics of Polintons by investigating the DNA POL B and RVE Integrase domains.

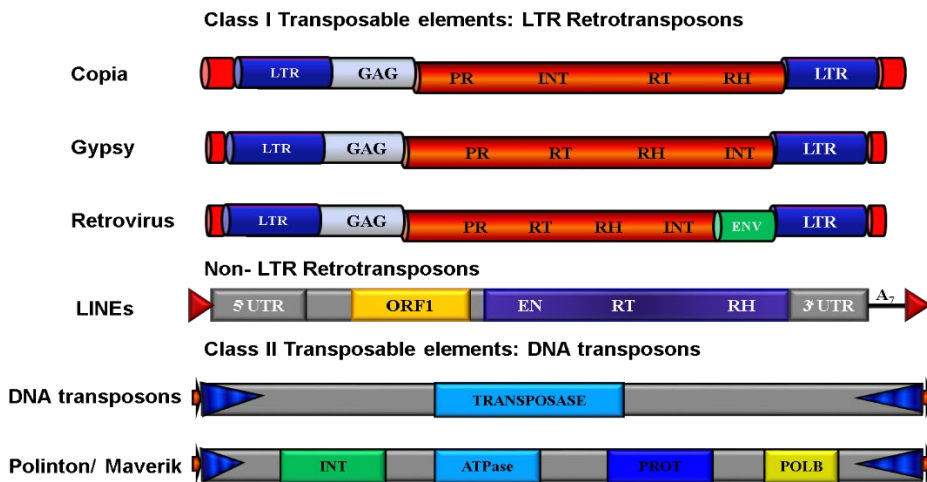


Figure 1. Schematic Representation of Various Types of Transposable Elements. Class 1 or Retrotransposons are Subdivided into LTR

Retrotransposons (Copia, Gypsy, Retroviruses) and Non-LTR Retrotransposons (LINEs) with Various Protein Coding Domains. Class II DNA Transposons have Transposase Domain for their Transposition. Polintons or Maverik Elements Show Different Structure Encoding Domains, Such as Integrase, ATPase, Protease, and DNA Polymerase B in their Structures

2. MATERIALS AND METHODS

2.1. Data Mining and Sequence Retrieval

The Repbase database of transposable elements was mined for Polinton sequences and various Polinton sequences deposited in Repbase database were retrieved for further analysis.

2.2. Investigation of Polintons and their Domains

Polintons retrieved from Repbase database were further analyzed. The element's name, host organisms from where they were identified, full element sizes, and their domains with sizes were investigated and tabulated. The sequences of Polintons were investigated in conserved domain database (CDD) of NCBI to detect various conserved protein domains.

2.3. Multiple Sequence Alignment and Phylogenetic Analyses

Multiple sequence alignment was performed in CLASTALW implemented in BioEdit software. After the sequence alignment, small gaps were removed and frameshifts were introduced to bring sequences into the same frame for further inspection. Trees were constructed using Mega software with 1000 bootstrap values. To calculate the genetic distance, neighbor joining method was applied in Mega Software.

3. RESULTS

One hundred and two (102) Polinton sequences, deposited in Repbase database of transposable elements, were retrieved and analyzed further. The names of elements, the host organism from where they were detected, full lengths (sizes) of elements, identified domains as well as sizes of domains were identified and listed as shown in Table 1. The Polinton elements ranged in sizes from 1.5kb to 42.5kb with variations in sizes and of various Polintons proliferating in various genomes. The smallest, however, partial Polinton elements were 1589 nucleotide (nt) Polinton-1B_HM identified from host genome *Hydra vulgaris* and Polinton1B_CPB (1858 nt) from

Chrysemys picta bellii both with missing RVE Integrase and DNA POLB domains. The largest elements were Polinton-2_HM (42550) and Polinton-3_HM (38791 nt) from *Hydra vulgaris*. The important domains of Polinton sequences were also detected by running their sequences in conserved domain database (CDD) implemented in NCBI. The important domains detected from Polintons included AEP, archaeo-eukaryotic primase; MCP, major capsid protein; RVE Integrase; POLB, protein-primed polymerase of family B; Dcm, methyltransferase of the Dcm family; GIY, GIY-YIG family nuclease; S1H, superfamily 1 helicase; primpol, S3H, superfamily 3 helicase; primase-polymerase, and TVpol, transposon-viral polymerase.

Using the MEGA program, phylogenetic tree was constructed to visualize the relationships of Polinton sequences across various species. Of 102 Polinton sequences, the RVE Integrase domain was identified from 75 sequences, which was further used for phylogenetic analyses. The circular cladogram of these 75 Polinton sequences (based on RVE Integrase) was clustered into two distinct clades (Figure 2), with each clade further clustering in sub-clades, groups, and sub-groups. The first clade consists of 41 sequences, while the second exhibits 34 sequences. Clade 1 is further clustered in 8 groups with various numbers of sequences in each group. Polintons from *Nasonia vitripennis* and *Strongylocentrotus purpuratus* are mostly clustered together in one group of Clade 1. The Polinton sequences of *Nematostella vectensis*, *Danio rerio*, and *Xenopus tropicalis* are clustered together in one group. The homologous Polinton sequences of *Pogonomyrmex barbatus* and *Strongylocentrotus purpuratus* are grouped together, while few elements are clustered in sister group with *Nasonia vitripennis*. The Polinton sequences from *Schmidtea mediterranea* and *Hydra vulgaris* came close together in a group, showing evolutionary relationship with each other. Clade 2 is further clustered in 8 groups with variable numbers of elements in each group. The RVE Integrase-based Polinton sequences from *Trichomonas vaginalis* are clustered together in a group. Although, the sequences from *Nasonia vitripennis*, *Drosophila bipunctate*, and *Drosophila eugracilis* were found in various groups, indicating evolutionary phenomena (Fig.1). This evolutionary representation allows the observation of intricate branching patterns and relationships within the Polinton superfamily. The cladogram provides a roadmap for exploring the relationships and dynamics of the Polinton superfamily in detail.

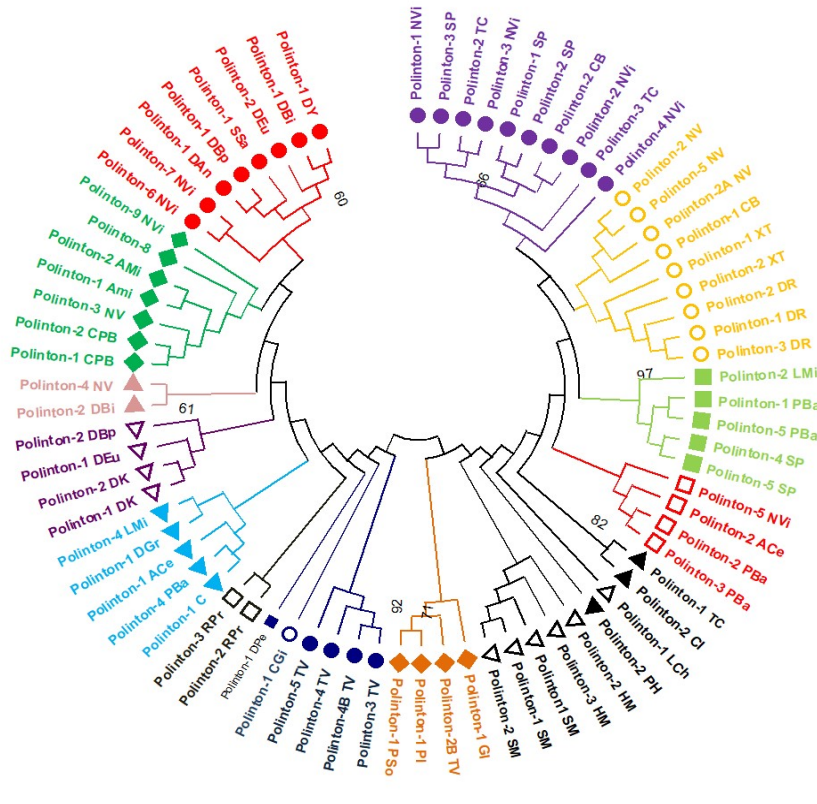


Figure 2. Seventy-five (75) Polinton Sequences having RVE Integrase Domain were Subjected to Phylogenetic Analysis. The Phylogenetic Tree was Constructed in Mega X using the 1000 Bootstrap Values. The Genetic Distance was Calculated through Neighbor-Joining Method. The Tree was Resolved in 2 Major Clades with Several Sub-Clades and Groups.

The POLB domain was also brought under evolutionary investigations, where of 102 Polinton sequences, 63 showed the POLB domain. Using the neighbor-joining method in Mega, a circular cladogram was created of these 63 Polinton sequences (Fig. 3). The tree revealed two major clades with distinct subclades and groups. The first clade consists of 52 elements, while the second clade is represented by 11 elements. Notably, the second clade encompassed a wide range of data from Polinton 1 TC to Polinton 2 TC. Although, there were some similarities, the elements in the first and second

clades exhibited distinct characteristics. The Polinton sequences from *Nasonia vitripennis* and *Nematostella vectensis* were found distributed in various groups. Polinton 5TV and Polinton 1 XT are closely related within the first clade, forming a sister group. However, Polinton 2NV, being a part of the same group, showed a more distant relationship with Polinton 5TV and Polinton 1 XT (Figure 3). This cladogram visually represents the genetic relationships among Polinton sequences, focusing on the DNA POLB domain. The presence of distinct clades and subclades indicates evolutionary divergences within the Polinton superfamily.

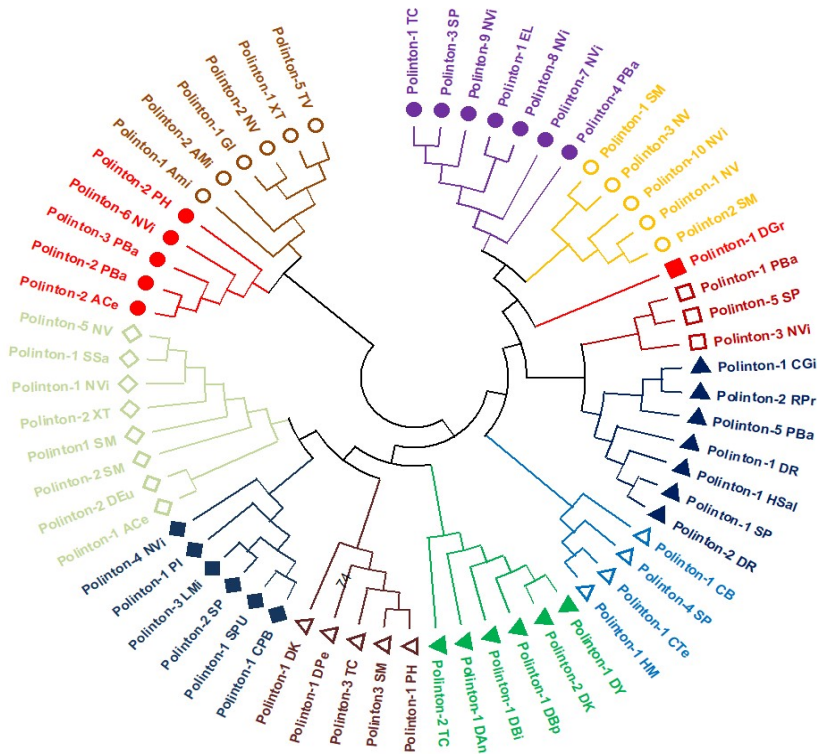


Figure 3. Sixty-three (63) Polinton Sequences having POLB Domain were Subjected to Phylogenetic Analysis. The Phylogenetic Tree was Constructed in Mega X using the 1000 Bootstrap Values. The Neighbor-Joining Method was Used to Calculate the Genetic Distance for Tree Construction

The current study highlighted the significance of phylogenetic analysis in understanding the evolutionary connections between transposons and other organisms, shedding light on the origins of eukaryotes and the role of Polinton in eukaryotic genome evolution.

Table 1. List of Polintons/Maverick collected from various host organisms with their sizes and domains. NI; Not identified

No.	Element name	Host organism	Size (nt)	RVE Integrase size (nt)	POLB size (nt)
1	Mavirus Spezl Polinton	<i>Cafeteria roenbergensis</i>	19063	314	NI
2	Polinton10 Nvi	<i>Nasonia vitripennis</i>	5601	NI	1085
3	Polinton1B CPB	<i>Chrysemys picta bellii</i>	1858	NI	NI
4	Polinton-1B HM	<i>Hydra vulgaris</i>	1589	NI	NI
5	Polinton-1B TV	<i>Trichomonas vaginalis</i>	20519	NI	NI
6	Polinton-1 Ace	<i>Atta cephalotes</i>	14410	281	869
7	Polinton-1 Ami	<i>Crocodylidae</i>	14974	320	2333
8	Polinton-1 CB	<i>Caenorhabditis briggsae</i>	16633	326	1454
9	Polinton-1 CGi	<i>Crassostrea gigas</i>	20773	323	1462
10	Polinton-1 CI	<i>Ciona intestinalis</i>	15061	326	NI
11	Polinton-1 CPB	<i>Chrysemys picta bellii</i>	13369	326	1415
12	Polinton-1 CTe	<i>Capitella teleta</i>	5187	NI	1217
13	Polinton-1 Dan	<i>Drosophila ananassae</i>	17900	326	1217
14	Polinton-1 DBi	<i>Drosophila biarmipes</i>	13688	335	857
15	Polinton-1 DBp	<i>Drosophila bipectinata</i>	7049	326	1451
16	Polinton-1 Del	<i>Drosophila elegans</i>	10296	NI	563
17	Polinton-1 DEu	<i>Drosophila eugracilis</i>	10967	332	NI
18	Polinton-1 DGr	<i>Drosophila grimshawi</i>	15303	326	1268
19	Polinton-1 DK	<i>Drosophila kikkawai</i>	7252	326	863
20	Polinton-1 Dpe	<i>Drosophila persimilis</i>	15797	335	1460
21	Polinton-1 DR	<i>Danio rerio</i>	18485	320	1883
22	Polinton-1 DY	<i>Drosophila yakuba</i>	14782	335	1451
23	Polinton-1 EI	<i>Entamoeba invadens</i>	16504	NI	473
24	Polinton-1 EL	<i>Esox Lucius</i>	14284	NI	1883
25	Polinton-1 GI	<i>Rhizophagus intraradices</i>	11954	347	1583
26	Polinton-1 HM	<i>Hydra vulgaris</i>	20689	NI	1436
27	Polinton-1 HSal	<i>Harpegnathos saltator</i>	3554	NI	2267
28	Polinton-1 LCh	<i>Latimeria chalumnae</i>	17525	218	535
29	Polinton-1 LMi	<i>Locusta migratoria</i>	8133	NI	NI
30	Polinton-1 NV	<i>Nematostella vectensis</i>	17653	NI	1295

No.	Element name	Host organism	Size (nt)	RVE Integrase size (nt)	POLB size (nt)
31	Polinton-1 NVi	<i>Nasonia vitripennis</i>	14499	329	1408
32	Polinton-1 PBa	<i>Pogonomyrmex barbatus</i>	16368	329	1247
33	Polinton-1 PH	<i>Parhyale hawaiiensis</i>	10772	NI	692
34	Polinton-1 PI	<i>Phytophthora infestans</i>	18398	314	755
35	Polinton-1 PSo	<i>Phytophthora sojae</i>	19227	314	NI
36	Polinton-1 SM	<i>Schmidtea mediterranea</i>	14867	206	1178
37	Polinton-1_SP	<i>Strongylocentrotus purpuratus</i>	16918	329	1970
38	Polinton-1 SPU	<i>Sphenodon punctatus</i>	11940	NI	1567
39	Polinton-1 SSa	<i>Salmo salar</i>	16270	335	1886
40	Polinton-1 TC	<i>Tribolium castaneum</i>	13486	332	1919
41	Polinton-1 TV	<i>Trichomonas vaginalis</i>	20724	NI	NI
42	Polinton-1 XT	<i>Xenopus tropicalis</i>	13692	329	1922
43	Polinton-2A NV	<i>Nematostella vectensis</i>	20836	329	733
44	Polinton-2B TV	<i>Trichomonas vaginalis</i>	25444	329	NI
45	Polinton-2 Ace	<i>Atta cephalotes</i>	14167	335	1958
46	Polinton-2 Ami	<i>Alligator mississippiensis</i>	15555	320	1916
47	Polinton-2 CB	<i>Caenorhabditis briggsae</i>	15471	329	923
48	Polinton-2 CI	<i>Ciona intestinalis</i>	13695	326	NI
49	Polinton-2 CPB	<i>Chrysemys picta bellii</i>	6310	320	NI
50	Polinton-2 DBi	<i>Drosophila biarmipes</i>	8566	314	NI
51	Polinton-2 DBp	<i>Drosophila bipectinata</i>	3416	302	NI
52	Polinton-2 DEu	<i>Drosophila eugracilis</i>	11524	335	1451
53	Polinton-2 DK	<i>Drosophila kikkawai</i>	6981	326	1226
54	Polinton-2 DR	<i>Danio rerio</i>	16276	323	1655
55	Polinton-2 HM	<i>Hydra vulgaris</i>	42550	347	NI
56	Polinton-2 Lmi	<i>Locusta migratoria</i>	7239	272	NI
57	Polinton-2 NV	<i>Nematostella vectensis</i>	21301	329	1892
58	Polinton-2 NVi	<i>Nasonia vitripennis</i>	8996	338	NI
59	Polinton-2 PBa	<i>Pogonomyrmex barbatus</i>	18080	335	1958
60	Polinton-2 PH	<i>Parhyale hawaiiensis</i>	16250	323	1475
61	Polinton-2 RPr	<i>Rhodnius prolixus</i>	17029	323	1412
62	Polinton-2 SM	<i>Schmidtea mediterranea</i>	15944	326	1762
63	Polinton-2_SP	<i>Strongylocentrotus purpuratus</i>	14353	329	1952
64	Polinton-2 TC	<i>Tribolium castaneum</i>	16981	329	1490
65	Polinton-2 TV	<i>Trichomonas vaginalis</i>	23222	NI	NI
66	Polinton-2 XT	<i>Xenopus tropicalis</i>	14828	320	1910

No.	Element name	Host organism	Size (nt)	RVE Integrase size (nt)	POLB size (nt)
67	Polinton-3 CI	<i>Ciona intestinalis</i>	6621	NI	NI
68	Polinton-3 DR	<i>Danio rerio</i>	18618	320	365
69	Polinton-3 HM	<i>Hydra vulgaris</i>	38791	344	NI
70	Polinton-3 Lmi	<i>Locusta migratoria</i>	5107	NI	794
71	Polinton-3 NV	<i>Nematostella vectensis</i>	16575	344	1451
72	Polinton-3 Nvi	<i>Nasonia vitripennis</i>	12588	323	1193
73	Polinton-3 Pba	<i>Pogonomyrmex barbatus</i>	18021	335	1949
74	Polinton-3 RPr	<i>Rhodnius prolixus</i>	10522	332	NI
75	Polinton-3_SP	<i>Strongylocentrotus purpuratus</i>	16510	332	1922
76	Polinton-3 TC	<i>Tribolium castaneum</i>	17681	335	1508
77	Polinton-3 TV	<i>Trichomonas vaginalis</i>	22038	284	1244
78	Polinton-4B TV	<i>Trichomonas vaginalis</i>	21735	281	1244
79	Polinton-4N1 DR	<i>Danio rerio</i>	10436	NI	NI
80	Polinton-4 DR	<i>Danio rerio</i>	20287	NI	769
81	Polinton-4 Hma	<i>Hydra vulgaris</i>	8516	NI	NI
82	Polinton-4 LMi	<i>Locusta migratoria</i>	14821	188	NI
83	Polinton-4 LVa	<i>Litopenaeus vannamei</i>	16994	NI	NI
84	Polinton-4 NV	<i>Nematostella vectensis</i>	13070	317	NI
85	Polinton-4 NV	<i>Nasonia vitripennis</i>	12608	323	1007
86	Polinton-4 Pba	<i>Pogonomyrmex barbatus</i>	15965	329	1247
87	Polinton-4_SP	<i>Strongylocentrotus purpuratus</i>	15575	326	1892
88	Polinton-4 TV	<i>Trichomonas vaginalis</i>	21843	221	1244
89	Polinton-5B TV	<i>Trichomonas vaginalis</i>	3963	NI	NI
90	Polinton-5 NV	<i>Nematostella vectensis</i>	19102	329	1895
91	Polinton-5 NVi	<i>Nasonia vitripennis</i>	18379	335	1222
92	Polinton-5 Pba	<i>Pogonomyrmex barbatus</i>	19981	326	1244
93	Polinton-5_SP	<i>Strongylocentrotus purpuratus</i>	16525	323	1922
94	Polinton-5 TV	<i>Trichomonas vaginalis</i>	21759	332	1373
95	Polinton-6 NV	<i>Nasonia vitripennis</i>	10272	323	716
96	Polinton-7 Nvi	<i>Nasonia vitripennis</i>	8494	326	1229
97	Polinton-8 NVi	<i>Nasonia vitripennis</i>	16996	338	1979
98	Polinton-9 N	<i>Nasonia vitripennis</i>	17121	329	1454
99	Polinton-N1A NV	<i>Nematostella vectensis</i>	14335	NI	NI
100	Polinton1 SM	<i>Schmidtea mediterranea</i>	12786	431	743
101	Polinton2 SM	<i>Schmidtea mediterranea</i>	11578	NI	1114

No.	Element name	Host organism	Size (nt)	RVE Integrase size (nt)	POLB size (nt)
102	Polinton3 SM	<i>Schmidtea mediterranea</i>	2836	NI	1379

4. DISCUSSION

The exploration of different databases has revolutionized the field of genomics and has led to the discovery of numerous new groups of organisms. One such group is the small viral genomes. This group has been identified through the detection of novel viruses. These small viral genomes pose unique challenges in terms of sequence analysis due to their rapid evolution and high mutation rates [12]. Moreover, another intriguing group of organisms which has garnered attention in recent years is the Polinton or Maverick transposons. Initially, classified as a type of DNA transposon, Mavericks have since been recognized as a viral group due to their distinct features and similarities to viral elements. They encode a unique class of integrases called c-integrases, which play essential roles in their replication and integration into host genomes [14].

The discovery of c-integrases has opened new avenues to understand the evolutionary relationships and functional significance of these elements. Comparative analyses of c-integrases have revealed striking similarities to retroviral integrases (RVE) in terms of the catalytic core domain and chromodomain [14, 15]. The presence of these conserved domains suggests a shared ancestry and functional convergence between c-integrases and RVE. By analyzing the positioning of segments in the current study, common ancestry, shared domains, and genetic similarities may be identified among Polinton elements. They also serve as a guide for further investigations into the functional implications and evolutionary significance of specific Polinton subgroups. The similarity of these Polintons was observed in various specie genomes and variations within same or closely-related genomes, revealing their diverse evolutionary nature.

To elucidate the evolutionary history of c-integrases, phylogenetic analyses have been performed. The RVE domain of Polintons was aligned with integrases from diverse elements, such as Tlr1 (ciliate), Tdd4 (slime mold), Ty1 (budding yeast), and Copia (Drosophila). The resulting phylogenetic tree revealed a monophyletic clade within the Polinton c-integrases group, with Tlr1 and Tdd-4 integrases forming a deep-branching position alongside the c-integrase of Mav_Tv1.1 from the parabasalid

Trichomonas vaginalis. These findings suggest a common ancestry and potential functional similarities between Tlr1, Tdd-4, and the c-integrase of *T. vaginalis*. Moreover, the presence of c-integrases in parabasalids, is considered as one of the ancient groups of eukaryotes, highlighting their deep-rooted evolutionary origins. Notably, the c-integrases found in *Phytophthora infestans* and *Tetrahymena thermophila* demonstrate a monophyletic grouping with the other protozoan eukaryotes, underscoring their basal position within the c-integrases phylogeny [14–16]. This provides valuable insights into the evolutionary history and diversification of c-integrases across different eukaryotic lineages.

Polintons, characterized by their large genome size and vertical inheritance, represent an ancient group of transposable elements that have diversified over billions of years. They have played a significant role in shaping the genomes of various eukaryotic organisms. Through the analysis of conserved domains, such as the DNA POLB region and the identification of conserved motifs, researchers have been able to discover new sequences and unveil the diversity of these elements within eukaryotic genomes [15–18].

4.1. Conclusion

The current study aimed to shed light on the diversity, activity, and evolution of the Polinton superfamily across various organisms. This study has the potential to impact disciplines, such as genetics, medicine, and evolutionary biology, contributing to a broader understanding of nature. By examining the RVE Integrase and POLB domains of Polinton, valuable insights were gained into the genetic diversity and evolutionary history of Polintons, since these 2 domains were major and the most conserved domains. The investigation of Polintons and their role in evolution remains a critical field that calls for further exploration.

CONFLICT OF INTEREST

The authors of the manuscript have no financial or non-financial conflict of interest in the subject matter or materials discussed in this manuscript.

DATA AVAILABILITY STATEMENT

The data associated with this study will be provided by the corresponding author upon request.

FUNDING DETAILS

No funding has been received for this research.

REFERENCES

1. Kapitonov VV, Jurka J. A universal classification of eukaryotic transposable elements implemented in Repbase. *Nat Rev Genet.* 2008;9:411–412. <https://doi.org/10.1038/nrg2165-c2>
2. Wicker T, Sabot F, Hua-Van A, et al. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet.* 2007;8:973–982. <https://doi.org/10.1038/nrg2165>
3. Nouroz F, Noreen S, Heslop-Harrison JS. Identification and characterization of LTR Retrotransposons in *Brassica*. *Turk J Biol.* 2015;39(5):740–757. <https://doi.org/10.3906/biy-1501-77>
4. Nouroz F, Noreen S, Heslop-Harrison JS. Characterization and diversity of novel *PIF/Harbinger* DNA transposons in *Brassica* genomes. *Pak J Bot.* 2016;48(1):167–178.
5. Nouroz F, Noreen S, Heslop-Harrison JS. Identification and evolutionary dynamics of CACTA DNA transposons in *Brassica*. *Pak J Bot.* 2017;49(2):789–798.
6. Pritham EJ. Transposable elements and factors influencing their success in eukaryotes. *J Hered.* 2009;100(5):648–655. <https://doi.org/10.1093/jhered/esp065>
7. Krupovic M, Koonin EV. Self-synthesizing transposons: unexpected key players in the evolution of viruses and defense systems. *Curr Opin Microbiol.* 2016;31:25–33. <https://doi.org/10.1016/j.mib.2016.01.006>
8. Arensburger P, Piégu B, Bigot Y. The future of transposable element annotation and their classification in the light of functional genomics - what we can learn from the fables of Jean de la Fontaine? *Mob Genet Elements.* 2016;6(6):e1256852. <https://doi.org/10.1080/2159256X.2016.1256852>
9. Blanc-Mathieu R, Ogata H. DNA repair genes in the *Megavirales* pangenome. *Curr Opin Microbiol.* 2016;31:94–100. <https://doi.org/10.1016/j.mib.2016.03.011>
10. Brum JR, Sullivan MB. Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nat Rev Microbiol.* 2015;13(3):147–159. <https://doi.org/10.1038/nrmicro3404>
11. Bouvet GF, Jacobi V, Plourde KV, Bernier L. Stress-induced mobility of OPHIO1 and OPHIO2, DNA transposons of the Dutch elm disease

- fungi. *Fungal Genet Biol.* 2008;45(4):565–578. <https://doi.org/10.1016/j.fgb.2007.12.007>
12. Fischer MG. The virophage family Lavidaviridae. *Curr Issues Mol Biol.* 2021;40:1–24. <https://doi.org/10.21775/cimb.040.001>
13. Cosby RL, Chang NC, Feschotte C. Host-transposon interactions: conflict, cooperation, and cooption. *Genes Dev.* 2019;33(17-18):1098–1116. <https://doi.org/10.1101/gad.327312.119>
14. Pritham EJ, Putliwala T, Feschotte C. Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses. *Gene.* 2007;390(1-2):3–17. <https://doi.org/10.1016/j.gene.2006.08.008>
15. Haapa-Paananen S, Wahlberg N, Savilahti H. Phylogenetic analysis of Maverick/Polinton giant transposons across organisms. *Mol Phylogenet Evol.* 2014;78:271–274. <https://doi.org/10.1016/j.ympev.2014.05.024>
16. Beck CR, Garcia-Perez JL, Badge RM, Moran JV. LINE-1 elements in structural variation and disease. *Annu Rev Genomics Hum. Genet.* 2011;12:187–215. <https://doi.org/10.1146/annurev-genom-082509-141802>
17. da Cruz MHP, Domingues DS, Saito PTM, Paschoal AR, Bugatti PH. TERL: classification of transposable elements by convolutional neural networks. *Brief Bioinform.* 2021;22(3):ebbaa185. <https://doi.org/10.1093/bib/bbaa185>
18. Haapa-Paananen A, Wahlberg N, Savilahti H. Phylogenetic analysis of Maverick/Polinton giant transposons across organisms. *Mol Phylogenet Evol.* 2014;78:271–274. <https://doi.org/10.1016/j.ympev.2014.05.024>