In [1]:

import numpy as np
import pandas as pd

In [2]:

```
ft = pd.read_csv(
    'transcriptors_matrix.txt',
    sep=',', header=None, names=['id', 'family', 'A', 'C', 'G', 'T'])
print(type(ft), ft.shape)
ft
```

<class 'pandas.core.frame.DataFrame'> (489, 6)

Out[2]:

	id	family	А	С	G	Т
0	>MA0020.1	Dof2	A[212121 037]	C[000014 6]	G[000212 3]	T[00002 5]
1	>MA0021.1	Dof3	A[212121 006]	C[000010 6]	G[000213 9]	T[00008 0]
2	>MA0034.1	Gam1	A [4 10 3 23 25 1	C [6 5 13 1 0 24	G[117010 0	T[43900 0
3	>MA0044.1	HMG-1	A[00000 6	C[50421 1	G[830100 3	T [0 10 9 1 12 3
4	>MA0045.1	HMG-I/Y	A[379311 11	C[50160 0	G[43143 2	T[24310 1
5	>MA0053.1	MNB1A	A [15 15 15 0 3]	C[00009]	G[000150]	T[00003]
6	>MA0054.1	myb.Ph3	A [19 64 63 4 10 10	C [3 1 0 62 27 2	G [2 2 2 3 16 53	T [46 3 5 1 17 5
7	>MA0064.1	PBF	A[161616 01]	C[00009]	G[000161]	T[00005]
8	>MA0082.1	squamosa	A [11 0 24 16 23 17	C [14 22 0 2 0 0	G[10410 0	T[482117 13
9	>MA0096.1	bZIP910	A [15 0 0 35 0 0	C [15 0 0 0 35 0	G [5 0 35 0 0 35	T[035000 0
10	>MA0097.1	bZIP911	A [1 17 0 0 33 0	C[00000 33	G [31 16 0 33 0 0	T[103300 0
11	>MA0120.1	id1	A[13100 2	C[132913 11	G[24934 4	T [20 14 12 12 7 7
12	>MA0121.1	ARR10	A [14 0 14 0 4 1	C[10006 8	G[015100 0	T[000155 6
13	>MA0123.1	abi4	A[012001 0	C [49 0 20 23 3 45	G [0 37 29 1 45 4	T[000250 0
14	>MA0127.1	PEND	A [30 6 1 4 3 0	C [2 19 0 0 33 2	G[87242 2	T [2 10 39 34 4 38
15	>MA0128.1	EmBP-1	A[911300 0	C[070120 0	G[040013 0	T[41010 13
16	>MA0129.1	TGA1A	A [4 13 0 2 1 0	C [4 0 14 0 0 15	G[010130 0	T[711014 0
17	>MA0549.1	BZR2	A [23 41 0 89 0 0	C [45 8 99 0 99 0	G [10 50 0 4 0 99	T[210060 0
18	>MA0551.1	HY5	A [103 108 52 9 150 0	C [60 56 55 8 165 317	G [68 62 30 279 1 1	T [89 94 183 24 4 2

	id	family	Α	С	G	Т
19	>MA0552.1	PIF1	A [37 25 35 53 27 9	C [12 0 28 8 27 17	G [37 53 29 40 32 73	T [28 36 22 13 28 15
20	>MA0553.1	SMZ	A[00000 0	C [105 119 0 147 0 0	G [0 28 0 0 147 25	T [42 0 147 0 0 122
21	>MA0554.1	SOC1	A [288 172 65 62 117 20 	C [123 264 27 0 47 855 	G [99 113 34 51 89 13	T [378 339 762 775 635 0
22	>MA0555.1	SVP	A [28 28 28 33 7 3	C [24 11 1 16 85 79	G [21 8 10 11 0 1	T [19 45 53 32 0 9
23	>MA0556.1	AP3	A [92 84 48 0 185 217	C [36 33 241 232 61 22	G [20 39 0 1 11 11	T [143 135 2 58 34 41
24	>MA0557.1	FHY3	A [46 75 9 226 2 3	C [81 58 210 3 230 16	G [40 25 9 2 1 213	T [68 77 7 4 2 3
25	>MA0558.1	FLC	A [121 90 95 82 78 54	C [44 61 37 26 25 193	G [45 63 21 36 66 12	T [65 61 122 131 106 16
26	>MA0559.1	PI	A [156 57 357 448 551 498	C [335 394 79 12 1 0	G [48 44 67 83 4 12	T [19 63 55 15 2 48
27	>MA0560.1	PIF3	A [210 58 19 0 527 0	C [26 131 399 527 0 409	G [148 222 94 0 0 0	T [143 116 15 0 0 118
28	>MA0561.1	PIF4	A [0 335 0 49 0 0	C [335 0 335 0 0 0	G [0 0 0 286 0 335	T [0 0 0 0 335 0
29	>MA0562.1	PIF5	A [0 0 286 0 91 0	C [78 286 0 286 0 0	G[0000 1950	T [208 0 0 0 0 286
459	>MA1396.1	GATA6	A [45 391 0 417 0 0	C [159 0 0 0 0 417	G [37 22 417 0 0 0	T [176 4 0 0 417 0
460	>MA1397.1	At1g74840	A [300 255 46 39 0 600 	C [8 32 92 0 0 0	G [14 114 2 561 600 0	T [278 199 460 0 0 0
461	>MA1398.1	At5g47390	A [225 190 277 290 76 136	C [119 122 73 45 285 300	G [52 67 63 92 48 57	T [203 220 186 172 190 106
462	>MA1399.1	At5g08520	A [194 292 248 209 221 197	C [69 23 17 42 59 2	G [166 72 15 180 30 376	T [169 211 318 167 288 23
463	>MA1400.1	At1g19000	A [272 319 282 33 29 0 	C [22 9 27 81 0 0	G [73 28 136 1 570 599	T [232 243 154 484 0 0

	id	family	А	С	G	Т
464	>MA1401.1	EPR1	A [252 421 594 597 597 0	C [93 9 0 0 0 0	G [164 16 0 0 0 0	T [88 151 3 0 0 597
465	>MA1402.1	BPC6	A [24 18 13 15 12 6	C [134 19 161 7 142 2 	G[201057 5	T [28 151 4 161 27 175
466	>MA1403.1	BPC5	A [95 3 102 2 100 1	C[00000 0	G [7 99 0 98 0 99	T[00022 2
467	>MA1404.1	BPC1	A [82 483 8 501 67 500	C [18 4 15 0 9 0	G [421 29 483 18 448 39	T [18 23 33 20 15 0
468	>MA0954.2	ATHB7	A [92 53 473 560 0 2	C [58 389 62 3 1 2	G [68 92 11 0 1 464	T [382 66 54 37 598 132
469	>MA1000.2	ERF105	A [144 149 168 189 50 0 	C [57 181 44 37 307 0	G [278 36 257 312 0 586	T [107 220 117 48 229 0
470	>MA0586.2	SPL14	A [125 23 0 0 0 583	C [7 534 583 0 0 0	G [74 0 0 583 0 0	T [377 26 0 0 583 0
471	>MA1026.2	ATHB15	A [265 363 268 212 199 13	C [43 31 35 63 46 97	G [158 84 60 131 329 0	T [124 112 227 184 16 480
472	>MA0548.2	AGL15	A [133 48 16 84 0 0	C [218 33 5 8 596 435	G [137 5 3 21 0 0	T [111 513 575 486 3 164
473	>MA0550.2	BZR1	A [95 4 351 0 573 0	C [260 519 34 575 1 588 	G [131 4 203 0 4 0	T [102 61 0 13 10 0
474	>MA1059.2	SPL5	A [244 273 102 25 0 0	C [107 73 150 153 0 0 	G [109 123 89 36 598 0 	T [138 129 257 384 0 598
475	>MA1405.1	SIZF2	A [16191 88380 264 3329 7440 63395	C [20703 627 78711 918 2953 4073	G [39343 8331 20496 1251 88717 4720	T [23762 2662 530 94502 889 27813
476	>MA1406.1	ATHB4	A [2392 82255 98348 297 3466 99128	C [40381 3298 1291 333 42828 66	G [1915 618 93 76 40460 333	T [55312 13829 268 99294 13246 473
477	>MA1407.1	bZIP14	A [40600 43800 1700 1700 95000 1700	C [3100 27100 1700 1700 1700 95000	G [28100 27100 1700 95000 1700 1700	T [28100 2100 95000 1700 1700 1700

	id	family	А	С	G	т
478	>MA1408.1	FaEOBII	A [45035 21068 1124 2313 77467 1296	C [6949 992 422 1649 6291 865	G [26996 73515 9987 767 2500 97032	T [21020 4425 88466 95271 13742 807
479	>MA1409.1	OsRR22	A [51257 86870 486 97906 495 72395	C [11381 3042 522 287 484 7724	G [11771 6041 97670 778 510 464 	T [25591 4047 1322 1028 98511 19416
480	>MA1410.1	StBRC1	A [11498 2437 419 4402 4602 925	C [23683 2082 447 14768 59710 97660	G [36598 93015 98500 75089 31742 962	T [28222 2466 634 5741 3946 453
481	>MA1411.1	TSAR1	A [31261 3300 12545 93404 284 1558	C [26502 5250 87319 16 98016 142	G [19831 91222 51 5527 142 98016	T [22405 228 84 1052 1558 284
482	>MA1412.1	TSAR2	A [17675 2687 1998 72496 392 4131	C [28815 2274 97506 128 95417 60	G [22859 94487 139 26792 60 95417	T [30652 552 358 585 4131 392
483	>MA1413.1	UIF1	A [42370 55605 86079 310 89162 21230	C [15280 8841 1019 188 201 96 	G [20606 19501 5263 86115 10588 172	T [21744 16053 7639 13387 50 78503
484	>MA1414.1	E2FA	A [394 355 23 21 35 7	C [203 339 62 1375 6 1415	G [184 579 1355 22 1391 18	T [670 178 11 33 19 11
485	>MA1415.1	REF6	A [1983 2634 3458 3511 1 3622 	C [209 137 27 34 3661 25	G [1144 577 24 118 1 30 	T [351 339 178 24 24 10
486	>MA1085.2	WRKY40	A [89 107 234 0 0 0	C [49 34 0 0 0 234	G [34 38 0 234 0 0	T [62 55 0 0 234 0
487	>MA1416.1	RAMOSA1	A [10200 56235 7487 61422 1100 74267	C [3932 3491 2714 2682 1543 348	G [56665 12116 61021 8810 70293 83	T [4047 3002 3622 1930 1908 146
488	>MA1417.1	O2	A [306 129 29 56 3 2021 	C [374 391 9 1959 2016 3 	G [1046 151 1895 12 17 12	T [312 1367 105 11 2 2

489 rows × 6 columns

```
In [3]:
```

```
print(ft['A'][0])
print(ft['C'][0])
print(ft['G'][0])
print (ft['T'][0])
Α
  [
        21
                21
                       21
                               0
                                       3
                                              7 ]
                                              6]
C
   [
         0
                 0
                        0
                               0
                                      14
G
  [
         0
                 0
                        0
                              21
                                       2
                                              3 ]
                                       2
                                              5 1
Т
   Γ
         0
                 0
                        0
                               0
In [4]:
def f_str_list (string):
    string = string.split()
    listed = []
    i = 2
    while(string[i] != ']'):
            listed.append(int (string[i]))
            i +=1
    return listed
In [5]:
ft['A'][0].split()
Out[5]:
['A', '[', '21', '21', '21', '0', '3', '7', ']']
In [6]:
np grupo a = ft['A']
In [7]:
np_grupo_a.count()
Out[7]:
489
In [8]:
i=0
while (i<489):
    np_grupo_a[i] = f_str_list(np_grupo_a[i])
    i += 1
```

In [9]:

np_grupo_a

```
Out[9]:
```

```
0
                                    [21, 21, 21, 0, 3, 7]
                                    [21, 21, 21, 0, 0, 6]
1
2
                       [4, 10, 3, 23, 25, 1, 3, 6, 10, 5]
3
                              [0, 0, 0, 0, 0, 6, 3, 0, 0]
4
       [3, 7, 9, 3, 11, 11, 11, 3, 4, 3, 8, 8, 9, 9, \dots]
5
                                       [15, 15, 15, 0, 3]
6
                       [19, 64, 63, 4, 10, 10, 13, 3, 28]
7
                                       [16, 16, 16, 0, 1]
8
       [11, 0, 24, 16, 23, 17, 24, 8, 14, 1, 14, 25, ...
9
                                  [15, 0, 0, 35, 0, 0, 0]
10
                 [1, 17, 0, 0, 33, 0, 0, 0, 0, 1, 11, 1]
                     [1, 3, 1, 0, 0, 2, 3, 1, 8, 0, 0,
11
12
                               [14, 0, 14, 0, 4, 1, 0, 3]
13
                          [0, 12, 0, 0, 1, 0, 5, 3, 3, 4]
                         [30, 6, 1, 4, 3, 0, 2, 32, 4, 2]
14
                                [9, 1, 13, 0, 0, 0, 0, 0]
15
                                  [4, 13, 0, 2, 1, 0, 13]
16
17
                 [23, 41, 0, 89, 0, 0, 0, 3, 16, 17, 36]
18
       [103, 108, 52, 9, 150, 0, 310, 3, 1, 0, 2, 4, ...
       [37, 25, 35, 53, 27, 9, 38, 0, 114, 0, 0, 0, 0...
19
20
                               [0, 0, 0, 0, 0, 0, 135, 0]
21
       [288, 172, 65, 62, 117, 20, 0, 459, 137, 130, ...
22
       [28, 28, 28, 33, 7, 3, 58, 67, 86, 77, 78, 34,...
       [92, 84, 48, 0, 185, 217, 283, 240, 156, 95, 9...
23
24
            [46, 75, 9, 226, 2, 3, 0, 4, 18, 16, 63, 82]
       [121, 90, 95, 82, 78, 54, 9, 135, 198, 261, 23...
25
       [156, 57, 357, 448, 551, 498, 301, 312, 231, 7...
26
                   [210, 58, 19, 0, 527, 0, 0, 0, 0, 259]
27
28
                             [0, 335, 0, 49, 0, 0, 0, 64]
29
                              [0, 0, 286, 0, 91, 0, 0, 0]
459
       [45, 391, 0, 417, 0, 0, 0, 345, 93, 298, 94, 9...
460
       [300, 255, 46, 39, 0, 600, 0, 578, 549, 168, 208]
       [225, 190, 277, 290, 76, 136, 11, 11, 599, 0, ...
461
       [194, 292, 248, 209, 221, 197, 0, 598, 0, 590,...
462
463
       [272, 319, 282, 33, 29, 0, 599, 0, 580, 554, 8...
       [252, 421, 594, 597, 597, 0, 597, 0, 0, 22, 15...
464
465
       [24, 18, 13, 15, 12, 6, 0, 2, 1, 0, 1, 2, 2, 0...
466
       [95, 3, 102, 2, 100, 1, 99, 0, 99, 1, 102, 0, ...
467
       [82, 483, 8, 501, 67, 500, 43, 478, 39, 520, 2...
468
       [92, 53, 473, 560, 0, 2, 600, 0, 0, 37, 324, 1...
469
       [144, 149, 168, 189, 50, 0, 21, 1, 0, 0, 21, 1...
470
          [125, 23, 0, 0, 0, 583, 0, 302, 239, 164, 148]
471
       [265, 363, 268, 212, 199, 13, 518, 590, 0, 27,...
       [133, 48, 16, 84, 0, 0, 200, 88, 162, 7, 89, 1...
472
473
       [95, 4, 351, 0, 573, 0, 0, 5, 0, 11, 213, 198,...
474
       [244, 273, 102, 25, 0, 0, 598, 0, 0, 0, 546, 170]
       [16191, 88380, 264, 3329, 7440, 63395, 889, 94...
475
       [2392, 82255, 98348, 297, 3466, 99128, 998, 55...
476
       [40600, 43800, 1700, 1700, 95000, 1700, 1700, ...
477
478
       [45035, 21068, 1124, 2313, 77467, 1296, 1374, ...
       [51257, 86870, 486, 97906, 495, 72395, 444, 16...
479
       [11498, 2437, 419, 4402, 4602, 925, 6571, 3063...
480
       [31261, 3300, 12545, 93404, 284, 1558, 1052, 8...
481
482
       [17675, 2687, 1998, 72496, 392, 4131, 585, 358...
       [42370, 55605, 86079, 310, 89162, 21230, 11260...
483
            [394, 355, 23, 21, 35, 7, 1, 1368, 574, 535]
484
485
       [1983, 2634, 3458, 3511, 1, 3622, 6, 3533, 45,...
486
               [89, 107, 234, 0, 0, 0, 230, 232, 83, 76]
```

```
487
       [10200, 56235, 7487, 61422, 1100, 74267, 9, 74...
       [306, 129, 29, 56, 3, 2021, 18, 354, 4, 19, 19...
Name: A, Length: 489, dtype: object
In [10]:
type(np_grupo_a[0][0])
Out[10]:
int
In [11]:
```

ft['A'] = np_grupo_a

In [12]:

ft['A']

Out[12]:

```
0
                                    [21, 21, 21, 0, 3, 7]
                                    [21, 21, 21, 0, 0, 6]
1
2
                       [4, 10, 3, 23, 25, 1, 3, 6, 10, 5]
3
                              [0, 0, 0, 0, 0, 6, 3, 0, 0]
4
       [3, 7, 9, 3, 11, 11, 11, 3, 4, 3, 8, 8, 9, 9, \dots]
5
                                       [15, 15, 15, 0, 3]
6
                       [19, 64, 63, 4, 10, 10, 13, 3, 28]
7
                                       [16, 16, 16, 0, 1]
8
       [11, 0, 24, 16, 23, 17, 24, 8, 14, 1, 14, 25, ...
9
                                  [15, 0, 0, 35, 0, 0, 0]
10
                 [1, 17, 0, 0, 33, 0, 0, 0, 0, 1, 11, 1]
11
                     [1, 3, 1, 0, 0, 2, 3, 1, 8, 0, 0,
12
                               [14, 0, 14, 0, 4, 1, 0, 3]
13
                          [0, 12, 0, 0, 1, 0, 5, 3, 3, 4]
                         [30, 6, 1, 4, 3, 0, 2, 32, 4, 2]
14
                                [9, 1, 13, 0, 0, 0, 0, 0]
15
                                  [4, 13, 0, 2, 1, 0, 13]
16
17
                 [23, 41, 0, 89, 0, 0, 0, 3, 16, 17, 36]
18
       [103, 108, 52, 9, 150, 0, 310, 3, 1, 0, 2, 4, ...
       [37, 25, 35, 53, 27, 9, 38, 0, 114, 0, 0, 0, 0...
19
20
                               [0, 0, 0, 0, 0, 0, 135, 0]
21
       [288, 172, 65, 62, 117, 20, 0, 459, 137, 130, ...
22
       [28, 28, 28, 33, 7, 3, 58, 67, 86, 77, 78, 34,...
       [92, 84, 48, 0, 185, 217, 283, 240, 156, 95, 9...
23
24
            [46, 75, 9, 226, 2, 3, 0, 4, 18, 16, 63, 82]
25
       [121, 90, 95, 82, 78, 54, 9, 135, 198, 261, 23...
       [156, 57, 357, 448, 551, 498, 301, 312, 231, 7...
26
                   [210, 58, 19, 0, 527, 0, 0, 0, 0, 259]
27
28
                             [0, 335, 0, 49, 0, 0, 0, 64]
29
                              [0, 0, 286, 0, 91, 0, 0, 0]
459
       [45, 391, 0, 417, 0, 0, 0, 345, 93, 298, 94, 9...
460
       [300, 255, 46, 39, 0, 600, 0, 578, 549, 168, 208]
       [225, 190, 277, 290, 76, 136, 11, 11, 599, 0, ...
461
       [194, 292, 248, 209, 221, 197, 0, 598, 0, 590,...
462
463
       [272, 319, 282, 33, 29, 0, 599, 0, 580, 554, 8...
       [252, 421, 594, 597, 597, 0, 597, 0, 0, 22, 15...
464
465
       [24, 18, 13, 15, 12, 6, 0, 2, 1, 0, 1, 2, 2, 0...
466
       [95, 3, 102, 2, 100, 1, 99, 0, 99, 1, 102, 0, ...
467
       [82, 483, 8, 501, 67, 500, 43, 478, 39, 520, 2...
468
       [92, 53, 473, 560, 0, 2, 600, 0, 0, 37, 324, 1...
469
       [144, 149, 168, 189, 50, 0, 21, 1, 0, 0, 21, 1...
470
          [125, 23, 0, 0, 0, 583, 0, 302, 239, 164, 148]
471
       [265, 363, 268, 212, 199, 13, 518, 590, 0, 27,...
       [133, 48, 16, 84, 0, 0, 200, 88, 162, 7, 89, 1...
472
473
       [95, 4, 351, 0, 573, 0, 0, 5, 0, 11, 213, 198,...
474
       [244, 273, 102, 25, 0, 0, 598, 0, 0, 0, 546, 170]
       [16191, 88380, 264, 3329, 7440, 63395, 889, 94...
475
       [2392, 82255, 98348, 297, 3466, 99128, 998, 55...
476
       [40600, 43800, 1700, 1700, 95000, 1700, 1700, ...
477
478
       [45035, 21068, 1124, 2313, 77467, 1296, 1374, ...
479
       [51257, 86870, 486, 97906, 495, 72395, 444, 16...
       [11498, 2437, 419, 4402, 4602, 925, 6571, 3063...
480
       [31261, 3300, 12545, 93404, 284, 1558, 1052, 8...
481
482
       [17675, 2687, 1998, 72496, 392, 4131, 585, 358...
       [42370, 55605, 86079, 310, 89162, 21230, 11260...
483
            [394, 355, 23, 21, 35, 7, 1, 1368, 574, 535]
484
485
       [1983, 2634, 3458, 3511, 1, 3622, 6, 3533, 45,...
486
               [89, 107, 234, 0, 0, 0, 230, 232, 83, 76]
```

```
487 [10200, 56235, 7487, 61422, 1100, 74267, 9, 74...
488 [306, 129, 29, 56, 3, 2021, 18, 354, 4, 19, 19...
Name: A, Length: 489, dtype: object
```

In [13]:

```
np_grupo_c = ft['C']
```

In [14]:

```
i=0
while (i<489):
    np_grupo_c[i] = f_str_list(np_grupo_c[i])
    i += 1</pre>
```

```
In [15]:
```

```
ft['C'] = np_grupo_c
```

In [16]:

ft['C']

```
Out[16]:
```

```
0
                                      [0, 0, 0, 0, 14, 6]
1
                                      [0, 0, 0, 0, 10, 6]
2
                     [6, 5, 13, 1, 0, 24, 14, 0, 11, 19]
3
                              [5, 0, 4, 2, 1, 1, 5, 1, 8]
4
        [5, 0, 1, 6, 0, 0, 0, 3, 1, 4, 5, 1, 0, 5, 0,
5
                                          [0, 0, 0, 0, 9]
6
                           [3, 1, 0, 62, 27, 2, 8, 17, 1]
7
                                          [0, 0, 0, 0, 9]
8
            [14, 22, 0, 2, 0, 0, 1, 0, 1, 1, 5, 5, 2, 5]
9
                                  [15, 0, 0, 0, 35, 0, 0]
10
                 [0, 0, 0, 0, 0, 33, 0, 0, 1, 0, 20, 32]
11
                 [1, 3, 2, 9, 13, 11, 8, 1, 5, 6, 17, 0]
12
                                 [1, 0, 0, 0, 6, 8, 0, 6]
13
                   [49, 0, 20, 23, 3, 45, 28, 25, 31, 26]
14
                         [2, 19, 0, 0, 33, 2, 0, 7, 3, 5]
                                [0, 7, 0, 12, 0, 0, 1, 1]
15
16
                                  [4, 0, 14, 0, 0, 15, 0]
17
                   [45, 8, 99, 0, 99, 0, 0, 0, 5, 33, 29]
18
       [60, 56, 55, 8, 165, 317, 1, 311, 5, 9, 1, 1, ...
       [12, 0, 28, 8, 27, 17, 14, 103, 0, 113, 0, 2, ...
19
20
                        [105, 119, 0, 147, 0, 0, 12, 147]
21
       [123, 264, 27, 0, 47, 855, 617, 37, 78, 6, 8, ...
       [24, 11, 1, 16, 85, 79, 16, 5, 0, 1, 8, 10, 0,...
22
23
       [36, 33, 241, 232, 61, 22, 0, 1, 10, 0, 2, 0, ...
24
       [81, 58, 210, 3, 230, 16, 220, 0, 208, 40, 68,...
25
       [44, 61, 37, 26, 25, 193, 185, 70, 10, 0, 3, 3...
       [335, 394, 79, 12, 1, 0, 8, 14, 0, 6, 54, 7, 2...
26
27
                [26, 131, 399, 527, 0, 409, 0, 0, 0, 33]
28
                          [335, 0, 335, 0, 0, 0, 99, 206]
29
                           [78, 286, 0, 286, 0, 0, 0, 95]
459
       [159, 0, 0, 0, 0, 417, 69, 0, 57, 50, 47, 162,...
460
                    [8, 32, 92, 0, 0, 0, 0, 6, 0, 60, 72]
       [119, 122, 73, 45, 285, 300, 44, 6, 0, 0, 599,...
461
          [69, 23, 17, 42, 59, 2, 0, 0, 0, 0, 0, 53, 40]
462
463
       [22, 9, 27, 81, 0, 0, 0, 0, 8, 0, 56, 61, 86, 71]
            [93, 9, 0, 0, 0, 0, 0, 597, 46, 100, 102]
464
       [134, 19, 161, 7, 142, 2, 160, 6, 177, 1, 168,...
465
       [0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, ...
466
467
       [18, 4, 15, 0, 9, 0, 17, 3, 0, 6, 2, 11, 9, 0, ...
468
          [58, 389, 62, 3, 1, 2, 0, 0, 0, 3, 46, 57, 91]
       [57, 181, 44, 37, 307, 0, 0, 569, 0, 1, 544, 0...
469
470
           [7, 534, 583, 0, 0, 0, 581, 87, 105, 104, 86]
471
       [43, 31, 35, 63, 46, 97, 1, 0, 0, 82, 0, 0, 0, ...
       [218, 33, 5, 8, 596, 435, 80, 85, 0, 8, 45, 42...
472
473
       [260, 519, 34, 575, 1, 588, 0, 0, 0, 277, 47, ...
474
         [107, 73, 150, 153, 0, 0, 0, 598, 0, 0, 8, 259]
475
       [20703, 627, 78711, 918, 2953, 4073, 88717, 12...
       [40381, 3298, 1291, 333, 42828, 66, 228, 764, ...
476
477
       [3100, 27100, 1700, 1700, 1700, 95000, 1700, 1...
478
       [6949, 992, 422, 1649, 6291, 865, 865, 21093, ...
479
       [11381, 3042, 522, 287, 484, 7724, 91648, 2838...
       [23683, 2082, 447, 14768, 59710, 97660, 91694,...
480
       [26502, 5250, 87319, 16, 98016, 142, 5527, 51,...
481
482
       [28815, 2274, 97506, 128, 95417, 60, 26792, 13...
       [15280, 8841, 1019, 188, 201, 96, 8610, 98826,...
483
       [203, 339, 62, 1375, 6, 1415, 1376, 10, 345, 293]
484
       [209, 137, 27, 34, 3661, 25, 21, 21, 10, 1063,...
485
486
                     [49, 34, 0, 0, 0, 234, 0, 0, 54, 48]
```

```
26/04/2018
                                             Pré-processamento
          [3932, 3491, 2714, 2682, 1543, 348, 99, 279, 1...
  487
          [374, 391, 9, 1959, 2016, 3, 1485, 103, 13, 19...
  Name: C, Length: 489, dtype: object
  In [17]:
   np grupo g = ft['G']
  In [18]:
   i=0
   while (i<489):
       np_grupo_g[i] = f_str_list(np_grupo_g[i])
  In [19]:
   ft['G'] = np_grupo_g
  In [20]:
   np_grupo_t = ft['T']
  In [21]:
   i=0
   while (i<489):
       np_grupo_t[i] = f_str_list(np_grupo_t[i])
```

```
i += 1
```

In [22]:

```
ft['T'] = np_grupo_t
```

In [23]:

ft

Out[23]:

	id	family	А	С	G	Т
0	>MA0020.1	Dof2	[21, 21, 21, 0, 3, 7]	[0, 0, 0, 0, 14, 6]	[0, 0, 0, 21, 2, 3]	[0, 0, 0, 0, 2, 5]
1	>MA0021.1	Dof3	[21, 21, 21, 0, 0, 6]	[0, 0, 0, 0, 10, 6]	[0, 0, 0, 21, 3, 9]	[0, 0, 0, 0, 8, 0]
2	>MA0034.1	Gam1	[4, 10, 3, 23, 25, 1, 3, 6, 10, 5]	[6, 5, 13, 1, 0, 24, 14, 0, 11, 19]	[11, 7, 0, 1, 0, 0, 6, 19, 0, 1]	
3	>MA0044.1	HMG-1	[0, 0, 0, 0, 0, 6, 3, 0, 0]	[5, 0, 4, 2, 1, 1, 5, 1, 8]	[8, 3, 0, 10, 0, 3, 0, 2, 1]	[0, 10, 9, 1, 12, 3, 5, 10, 4]
4	>MA0045.1	HMG-I/Y	[3, 7, 9, 3, 11, 11, 11, 3, 4, 3, 8, 8, 9, 9, 	[5, 0, 1, 6, 0, 0, 0, 3, 1, 4, 5, 1, 0, 5, 0, 7]	[4, 3, 1, 4, 3, 2, 2, 2, 8, 6, 1, 4, 2, 0, 3, 0]	[2, 4, 3, 1, 0, 1, 1, 6, 1, 1, 0, 1, 3, 0, 0, 5]
5	>MA0053.1	MNB1A	[15, 15, 15, 0, 3]	[0, 0, 0, 0, 9]	[0, 0, 0, 15, 0]	[0, 0, 0, 0, 3]
6	>MA0054.1	myb.Ph3	[19, 64, 63, 4, 10, 10, 13, 3, 28]	[3, 1, 0, 62, 27, 2, 8, 17, 1]	[2, 2, 2, 3, 16, 53, 0, 1, 0]	[46, 3, 5, 1, 17, 5, 49, 49, 41]
7	>MA0064.1	PBF	[16, 16, 16, 0, 1]	[0, 0, 0, 0, 9]	[0, 0, 0, 16, 1]	[0, 0, 0, 0, 5]
8	>MA0082.1	squamosa	[11, 0, 24, 16, 23, 17, 24, 8, 14, 1, 14, 25,	[14, 22, 0, 2, 0, 0, 1, 0, 1, 1, 5, 5, 2, 5]	[1, 0, 4, 1, 0, 0, 0, 1, 15, 28, 1, 0, 3, 8]	[4, 8, 2, 11, 7, 13, 5, 21, 0, 0, 10, 0, 4, 10]
9	>MA0096.1	bZIP910	[15, 0, 0, 35, 0, 0, 0]	[15, 0, 0, 0, 35, 0, 0]	[5, 0, 35, 0, 0, 35, 0]	[0, 35, 0, 0, 0, 0, 35]
10	>MA0097.1	bZIP911	[1, 17, 0, 0, 33, 0, 0, 0, 0, 1, 11, 1]	[0, 0, 0, 0, 0, 0, 33, 0, 0, 1, 0, 20, 32]	[31, 16, 0, 33, 0, 0, 33, 0, 32, 22, 1, 0]	[1, 0, 33, 0, 0, 0, 0, 33, 0, 10, 1, 0]
11	>MA0120.1	id1	[1, 3, 1, 0, 0, 2, 3, 1, 8, 0, 0, 2]	[1, 3, 2, 9, 13, 11, 8, 1, 5, 6, 17, 0]	=	[20, 14, 12, 12, 7, 7, 11, 21, 10, 11, 7, 4]
12	>MA0121.1	ARR10	[14, 0, 14, 0, 4, 1, 0, 3]	[1, 0, 0, 0, 6, 8, 0, 6]	[0, 15, 1, 0, 0, 0, 0, 9, 4]	[0, 0, 0, 15, 5, 6, 6, 2]
13	>MA0123.1	abi4	[0, 12, 0, 0, 1, 0, 5, 3, 3, 4]	[49, 0, 20, 23, 3, 45, 28, 25, 31, 26]	_	[0, 0, 0, 25, 0, 0, 10, 10, 10, 12]

	id	family	А	С	G	Т
14	>MA0127.1	PEND	[30, 6, 1, 4, 3, 0, 2, 32, 4, 2]	[2, 19, 0, 0, 33, 2, 0, 7, 3, 5]	[8, 7, 2, 4, 2, 2, 1, 2, 0, 14]	[2, 10, 39, 34, 4, 38, 39, 1, 35, 21]
15	>MA0128.1	EmBP-1	[9, 1, 13, 0, 0, 0, 0, 0, 0]	[0, 7, 0, 12, 0, 0, 1, 1]	[0, 4, 0, 0, 13, 0, 12, 12]	[4, 1, 0, 1, 0, 13, 0, 0]
16	>MA0129.1	TGA1A	[4, 13, 0, 2, 1, 0, 13]	[4, 0, 14, 0, 0, 15, 0]	[0, 1, 0, 13, 0, 0, 1]	[7, 1, 1, 0, 14, 0, 1]
17	>MA0549.1	BZR2	[23, 41, 0, 89, 0, 0, 0, 3, 16, 17, 36]	[45, 8, 99, 0, 99, 0, 0, 5, 33, 29]	[10, 50, 0, 4, 0, 99, 0, 96, 36, 46, 33]	[21, 0, 0, 6, 0, 0, 99, 0, 42, 3, 1]
18	>MA0551.1	HY5	[103, 108, 52, 9, 150, 0, 310, 3, 1, 0, 2, 4,	[60, 56, 55, 8, 165, 317, 1, 311, 5, 9, 1, 1,	[68, 62, 30, 279, 1, 1, 9, 5, 311, 1, 317, 165	[89, 94, 183, 24, 4, 2, 0, 1, 3, 310, 0, 150,
19	>MA0552.1	PIF1	[37, 25, 35, 53, 27, 9, 38, 0, 114, 0, 0, 0, 0	[12, 0, 28, 8, 27, 17, 14, 103, 0, 113, 0, 2,	[37, 53, 29, 40, 32, 73, 14, 11, 0, 1, 114, 0,	[28, 36, 22, 13, 28, 15, 48, 0, 0, 0, 0, 112,
20	>MA0553.1	SMZ	[0, 0, 0, 0, 0, 0, 135, 0]	[105, 119, 0, 147, 0, 0, 12, 147]	[0, 28, 0, 0, 147, 25, 0, 0]	[42, 0, 147, 0, 0, 122, 0, 0]
21	>MA0554.1	SOC1	[288, 172, 65, 62, 117, 20, 0, 459, 137, 130,	[123, 264, 27, 0, 47, 855, 617, 37, 78, 6, 8,	[99, 113, 34, 51, 89, 13, 9, 56, 0, 0, 0, 0, 0, 8	[378, 339, 762, 775, 635, 0, 262, 336, 673, 75
22	>MA0555.1	SVP	[28, 28, 28, 33, 7, 3, 58, 67, 86, 77, 78, 34,	[24, 11, 1, 16, 85, 79, 16, 5, 0, 1, 8, 10, 0,	[21, 8, 10, 11, 0, 1, 7, 8, 0, 2, 3, 14, 74, 9	[19, 45, 53, 32, 0, 9, 11, 12, 6, 12, 3, 34, 2
23	>MA0556.1	AP3	[92, 84, 48, 0, 185, 217, 283, 240, 156, 95, 9	[36, 33, 241, 232, 61, 22, 0, 1, 10, 0, 2, 0,	[20, 39, 0, 1, 11, 11, 0, 0, 104, 30, 191, 284	[143, 135, 2, 58, 34, 41, 8, 50, 21, 166, 0, 0
24	>MA0557.1	FHY3	[46, 75, 9, 226, 2, 3, 0, 4, 18, 16, 63, 82]	[81, 58, 210, 3, 230, 16, 220, 0, 208, 40, 68,	[40, 25, 9, 2, 1, 213, 0, 228, 6, 17, 40, 50]	[68, 77, 7, 4, 2, 3, 15, 3, 3, 162, 64, 59]
25	>MA0558.1	FLC	[121, 90, 95, 82, 78, 54, 9, 135, 198, 261, 23	[44, 61, 37, 26, 25, 193, 185, 70, 10, 0, 3, 3	[45, 63, 21, 36, 66, 12, 4, 37, 38, 3, 3, 18,	[65, 61, 122, 131, 106, 16, 77, 33, 29, 11, 36

	id	family	А	С	G	т
26	>MA0559.1	PI	[156, 57, 357, 448, 551, 498, 301, 312, 231, 7	[335, 394, 79, 12, 1, 0, 8, 14, 0, 6, 54, 7, 2	[48, 44, 67, 83, 4, 12, 244, 69, 327, 545, 7,	[19, 63, 55, 15, 2, 48, 5, 163, 0, 0, 51, 32,
27	>MA0560.1	PIF3	[210, 58, 19, 0, 527, 0, 0, 0, 0, 259]	[26, 131, 399, 527, 0, 409, 0, 0, 0, 33]	[148, 222, 94, 0, 0, 0, 527, 137]	[143, 116, 15, 0, 0, 118, 0, 527, 0, 98]
28	>MA0561.1	PIF4	[0, 335, 0, 49, 0, 0, 0, 64]	[335, 0, 335, 0, 0, 0, 99, 206]	[0, 0, 0, 286, 0, 335, 183, 65]	[0, 0, 0, 0, 335, 0, 53, 0]
29	>MA0562.1	PIF5	[0, 0, 286, 0, 91, 0, 0, 0]	[78, 286, 0, 286, 0, 0, 0, 95]	[0, 0, 0, 0, 195, 0, 286, 161]	[208, 0, 0, 0, 0, 286, 0, 30]
459	>MA1396.1	GATA6	[45, 391, 0, 417, 0, 0, 0, 345, 93, 298, 94, 9	[159, 0, 0, 0, 0, 0, 417, 69, 0, 57, 50, 47, 162,	[37, 22, 417, 0, 0, 0, 6, 72, 244, 30, 40, 53,	[176, 4, 0, 0, 417, 0, 342, 0, 23, 39, 236, 11
460	>MA1397.1	At1g74840	[300, 255, 46, 39, 0, 600, 0, 578, 549, 168, 208]	[8, 32, 92, 0, 0, 0, 0, 6, 0, 60, 72]	[14, 114, 2, 561, 600, 0, 0, 5, 38, 235, 228]	[278, 199, 460, 0, 0, 0, 600, 11, 13, 137, 92]
461	>MA1398.1	At5g47390	[225, 190, 277, 290, 76, 136, 11, 11, 599, 0,	[119, 122, 73, 45, 285, 300, 44, 6, 0, 0, 599,	=	[203, 220, 186, 172, 190, 106, 544, 567, 0, 59
462	>MA1399.1	At5g08520	[194, 292, 248, 209, 221, 197, 0, 598, 0, 590,	[69, 23, 17, 42, 59, 2, 0, 0, 0, 0, 0, 53, 40]	[166, 72, 15, 180, 30, 376, 596, 0, 0, 8, 0, 4	[169, 211, 318, 167, 288, 23, 2, 0, 598, 0, 0,
463	>MA1400.1	At1g19000	[272, 319, 282, 33, 29, 0, 599, 0, 580, 554, 8	[22, 9, 27, 81, 0, 0, 0, 0, 0, 8, 0, 56, 61, 86, 71]	[73, 28, 136, 1, 570, 599, 0, 0, 5, 33, 338, 3	[232, 243, 154, 484, 0, 0, 0, 599, 6, 12, 123,
464	>MA1401.1	EPR1	[252, 421, 594, 597, 597, 0, 597, 0, 0, 22, 15	[93, 9, 0, 0, 0, 0, 0, 0, 0, 0, 597, 46, 100, 102]	[164, 16, 0, 0, 0, 0, 0, 0, 0, 0, 0, 3, 66, 108]	[88, 151, 3, 0, 0, 597, 0, 526, 277, 90]
465	>MA1402.1	BPC6	[24, 18, 13, 15, 12, 6, 0, 2, 1, 0, 1, 2, 2, 0	[134, 19, 161, 7, 142, 2, 160, 6, 177, 1, 168,	[2, 0, 10, 5, 7, 5, 3, 1, 2, 0, 0, 0, 0, 0, 0,	[28, 151, 4, 161, 27, 175, 25, 179, 8, 187, 19

	id	family	А	С	G	Т
466	>MA1403.1	BPC5	[95, 3, 102, 2, 100, 1, 99, 0, 99, 1, 102, 0, 	[0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0,	[7, 99, 0, 98, 0, 99, 0, 100, 2, 100, 0, 102,	[0, 0, 0, 2, 2, 2, 2, 3, 1, 1, 0, 0, 0, 0, 1, 0,
467	>MA1404.1	BPC1	[82, 483, 8, 501, 67, 500, 43, 478, 39, 520, 2	[18, 4, 15, 0, 9, 0, 17, 3, 0, 6, 2, 11, 9, 0,	[421, 29, 483, 18, 448, 39, 461, 44, 482, 2, 5	[18, 23, 33, 20, 15, 0, 18, 14, 18, 11, 2, 7,
468	>MA0954.2	АТНВ7	[92, 53, 473, 560, 0, 2, 600, 0, 0, 37, 324, 1	[58, 389, 62, 3, 1, 2, 0, 0, 0, 3, 46, 57, 91]	[68, 92, 11, 0, 1, 464, 0, 0, 2, 551, 181, 98,	[382, 66, 54, 37, 598, 132, 0, 600, 598, 9, 49
469	>MA1000.2	ERF105	[144, 149, 168, 189, 50, 0, 21, 1, 0, 0, 21, 1	[57, 181, 44, 37, 307, 0, 0, 569, 0, 1, 544, 0	[278, 36, 257, 312, 0, 586, 565, 0, 586,	[107, 220, 117, 48, 229, 0, 0, 16, 0, 0, 21, 6
470	>MA0586.2	SPL14	[125, 23, 0, 0, 0, 583, 0, 302, 239, 164, 148]	[7, 534, 583, 0, 0, 0, 581, 87, 105, 104, 86]	[74, 0, 0, 583, 0, 0, 0, 129, 113, 83, 104]	[377, 26, 0, 0, 583, 0, 2, 65, 126, 232, 245]
471	>MA1026.2	ATHB15	[265, 363, 268, 212, 199, 13, 518, 590, 0, 27,	[43, 31, 35, 63, 46, 97, 1, 0, 0, 82, 0, 0, 0,	[158, 84, 60, 131, 329, 0, 68, 0, 0, 462, 0, 0	[124, 112, 227, 184, 16, 480, 3, 0, 590, 19, 0
472	>MA0548.2	AGL15	[133, 48, 16, 84, 0, 0, 200, 88, 162, 7, 89, 1	[218, 33, 5, 8, 596, 435, 80, 85, 0, 8, 45, 42	[137, 5, 3, 21, 0, 0, 50, 53, 4, 0, 31, 168, 4	[111, 513, 575, 486, 3, 164, 269, 373, 433, 58
473	>MA0550.2	BZR1	[95, 4, 351, 0, 573, 0, 0, 5, 0, 11, 213, 198,	[260, 519, 34, 575, 1, 588, 0, 0, 0, 277, 47,	[131, 4, 203, 0, 4, 0, 588, 0, 588, 39, 280, 9	[102, 61, 0, 13, 10, 0, 0, 583, 0, 261, 48, 15
474	>MA1059.2	SPL5	[244, 273, 102, 25, 0, 0, 598, 0, 0, 0, 546, 170]	[107, 73, 150, 153, 0, 0, 0, 598, 0, 0, 8, 259]	[109, 123, 89, 36, 598, 0, 0, 598, 598, 1, 19]	[138, 129, 257, 384, 0, 598, 0, 0, 0, 0, 43, 150]
475	>MA1405.1	SIZF2	[16191, 88380, 264, 3329, 7440, 63395, 889, 94	[20703, 627, 78711, 918, 2953, 4073, 88717, 12	[39343, 8331, 20496, 1251, 88717, 4720, 2953,	[23762, 2662, 530, 94502, 889, 27813, 7440, 33

	id	family	А	С	G	Т
476	>MA1406.1	ATHB4	[2392, 82255, 98348, 297, 3466, 99128, 998, 55	[40381, 3298, 1291, 333, 42828, 66, 228, 764,	[1915, 618, 93, 76, 40460, 333, 422, 2806, 684	[55312, 13829, 268, 99294, 13246, 473, 98352,
477	>MA1407.1	bZIP14	[40600, 43800, 1700, 1700, 95000, 1700, 1700, 	[3100, 27100, 1700, 1700, 1700, 95000, 1700, 1	[28100, 27100, 1700, 95000, 1700, 1700, 95000,	[28100, 2100, 95000, 1700, 1700, 1700, 9
478	>MA1408.1	FaEOBII	[45035, 21068, 1124, 2313, 77467, 1296, 1374, 	[6949, 992, 422, 1649, 6291, 865, 865, 21093, 	[26996, 73515, 9987, 767, 2500, 97032, 91609,	[21020, 4425, 88466, 95271, 13742, 807, 6153,
479	>MA1409.1	OsRR22	[51257, 86870, 486, 97906, 495, 72395, 444, 16	[11381, 3042, 522, 287, 484, 7724, 91648, 2838	[11771, 6041, 97670, 778, 510, 464, 611, 92633	[25591, 4047, 1322, 1028, 98511, 19416, 7297,
480	>MA1410.1	StBRC1	[11498, 2437, 419, 4402, 4602, 925, 6571, 3063	[23683, 2082, 447, 14768, 59710, 97660, 91694,	[36598, 93015, 98500, 75089, 31742, 962, 1031,	[28222, 2466, 634, 5741, 3946, 453, 705, 5158,
481	>MA1411.1	TSAR1	[31261, 3300, 12545, 93404, 284, 1558, 1052, 8	[26502, 5250, 87319, 16, 98016, 142, 5527, 51,	[19831, 91222, 51, 5527, 142, 98016, 16, 87319	[22405, 228, 84, 1052, 1558, 284, 93404, 12545
482	>MA1412.1	TSAR2	[17675, 2687, 1998, 72496, 392, 4131, 585, 358	[28815, 2274, 97506, 128, 95417, 60, 26792, 13	[22859, 94487, 139, 26792, 60, 95417, 128, 975	[30652, 552, 358, 585, 4131, 392, 72496, 1998,
483	>MA1413.1	UIF1	[42370, 55605, 86079, 310, 89162, 21230, 11260	[15280, 8841, 1019, 188, 201, 96, 8610, 98826,	[20606, 19501, 5263, 86115, 10588, 172, 5361,	[21744, 16053, 7639, 13387, 50, 78503, 74769,
484	>MA1414.1	E2FA	[394, 355, 23, 21, 35, 7, 1, 1368, 574, 535]	[203, 339, 62, 1375, 6, 1415, 1376, 10, 345, 293]	[184, 579, 1355, 22, 1391, 18, 1, 21, 121, 334]	[670, 178, 11, 33, 19, 11, 73, 52, 411, 289]

	id	family	А	С	G	Т
485	>MA1415.1	REF6	[1983, 2634, 3458, 3511, 1, 3622, 6, 3533, 45,	[209, 137, 27, 34, 3661, 25, 21, 21, 10, 1063,	[1144, 577, 24, 118, 1, 30, 3625, 105, 3581, 8	[351, 339, 178, 24, 24, 10, 35, 28, 51, 988, 524]
486	>MA1085.2	WRKY40	[89, 107, 234, 0, 0, 0, 0, 230, 232, 83, 76]	[49, 34, 0, 0, 0, 0, 234, 0, 0, 54, 48]	[34, 38, 0, 234, 0, 0, 0, 0, 46, 50]	[62, 55, 0, 0, 234, 0, 4, 2, 51, 60]
487	>MA1416.1	RAMOSA1	[10200, 56235, 7487, 61422, 1100, 74267, 9, 74	[3932, 3491, 2714, 2682, 1543, 348, 99, 279, 1	[56665, 12116, 61021, 8810, 70293, 83, 74518,	[4047, 3002, 3622, 1930, 1908, 146, 218, 165,
488	>MA1417.1	O2	[306, 129, 29, 56, 3, 2021, 18, 354, 4, 19, 19	[374, 391, 9, 1959, 2016, 3, 1485, 103, 13, 19	[1046, 151, 1895, 12, 17, 12, 7, 1571, 2, 28,	[312, 1367, 105, 11, 2, 2, 528, 10, 2019, 0, 4

489 rows × 6 columns

In [24]:

ft['A'][0][0]

Out[24]:

21