Genome wide analysis of Arabidopsis thaliana reveals high frequency of AAAG$_{N7}$CTTT motif

Rajesh Mehrotra$^{a,b,*}$, Vishesh Jain$^a$, Chandra Shekhar$^c$, Sandhya Mehrotra$^a$

$^a$ Department of Biological Sciences, Birla Institute of Technology & Sciences, Pilani, Rajasthan 333031, India
$^b$ Okinawa Institute of Science and Technology, Onna son, Okinawa, Japan
$^c$ Department of Mathematics, Birla Institute of Technology & Sciences, Pilani, Rajasthan 333031, India

ARTICLE INFO

Article history:
Received 4 March 2014
Accepted 14 May 2014
Available online 29 August 2014

Keywords:
Arabidopsis thaliana
Cis element
Designer promoter
Dof
Genome

ABSTRACT

Sequence specific elements in DNA regulate transcription by recruiting transcription factors. The Dof proteins are a large family of transcription factors that share a single highly conserved zinc finger. The core to which Dof proteins bind has a consensus AAAG or ACTTTA sequence. These motifs have been over represented in many promoters. We performed a genome wide analysis of AAAG repeat elements increasing the spacer length from 0 to 25. Similar analyses was done with AAAG-CTTT motifs. We report unusual high frequency of AAAG$_{N7}$CTTT in Arabidopsis thaliana genome. We also conclude that there is a preference for A/G nucleotides in spacer sequence between two AAAG repeats.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

Introduction

Promoters frequently contain multiple functional regulatory elements (Wray et al., 2003). This has an inherent question. How do redundancy and the evolution of cis element multiplicity take place. Cis elements are non coding DNA sequences present upstream of a gene and is required for proper spatio-temporal expression of the gene present downstream of it. It contains binding sites for transcription factors. The Dof domain proteins are typical example of plant specific transcription factors (Riechmann et al., 2000; Yanagisawa and Sheen, 1998; Yanagisawa, 2002). Dof transcription factor binds to a core sequence AAAG as shown by Vicente-Carbajosa et al. (1997) in a pull down assay. Dof domain proteins
have been shown to interact with another class of transcription factors (Zhang et al., 1995). We are very much interested in knowing how transcription factors select their target-like sequences which are scattered on the entire chromosome and how they function at this site. We have shown earlier that the minimal core sequences of the commonly occurring cis elements can enhance promoter expression, even when used out of their native contexts (Mehrotra and Mehrotra, 2010; Mehrotra and Panwar, 2009; Mehrotra et al., 2005; Sawant et al., 2005). Using ACGT core sequence we showed that probabilistic model is not followed when we look for the evolution of cis element multiplicity (Mehrotra et al., 2012, 2013). In this study we searched for the multiplicity of AAAG core sequence in the genome of Arabidopsis thaliana and reported that AAAGn7, CTTT is a preferred sequence in the genome. This information will be useful for designer promoters where specific interactions could be directed.

**Methodology**

The objective was to find out the frequency of the recurring sequences. Sequences of chromosomes were downloaded form the NCBI website (www.ncbi.nlm.nih.gov) and converted to a single line sequence using Notepad++. An ANSI C code was generated and later a code in Python 2.6.5 was used to find the results. The code written is as follows:

1. Code to find frequency of AAAG(A/G/C/T)AAAG as in Table 1

2. Python codes to find frequency of two AAAG/CTTTs separated by 0–25 nt. Spacer

```c
#include<stdio.h>
#include<stdlib.h>
int main ()
{
    char *p,*x;
    int j, singlemotif= 0;

    int nA = 0;
    int nG = 0;
    int nC = 0;
    int nT = 0;

    if(( x = (char*) malloc(25000000)) == NULL)
    {
        printf("No space available \n");
        exit(1);
    }
    FILE *fp=fopen("chr1_Ied.txt","r");
    if (fp == NULL)
    {
        puts ("Cannot open file");
        exit(1);
    }
    while(!feof(fp))
    {
        for(p=x;p<=x+25000000;p++)
        {
            scanf(fp,"%c",p);
        }
        for (p=x;p<=x+25000000;p++)
        {
            for (p=x;p<=x+25000000;p++)
            {
```

{
    if ( (*p == 'A') && (* (p+1) == 'A') && (* (p+2) == 'A') && (* (p+3) == 'G') && (* (p+4) == 'A') && (* (p+5) == 'A') && (* (p+6) == 'A') && (* (p+7) == 'A') && (* (p+8) == 'G') ) nA = nA + 1;
    if ( (*p == 'A') && (* (p+1) == 'A') && (* (p+2) == 'A') && (* (p+3) == 'G') && (* (p+4) == 'G') && (* (p+5) == 'A') && (* (p+6) == 'A') && (* (p+7) == 'A') && (* (p+8) == 'G') ) nG = nG + 1;
    if ( (*p == 'A') && (* (p+1) == 'A') && (* (p+2) == 'A') && (* (p+3) == 'G') && (* (p+4) == 'C') && (* (p+5) == 'A') && (* (p+6) == 'A') && (* (p+7) == 'A') && (* (p+8) == 'G') ) nC = nC + 1;
    if ( (*p == 'A') && (* (p+1) == 'A') && (* (p+2) == 'A') && (* (p+3) == 'G') && (* (p+4) == 'T') && (* (p+5) == 'A') && (* (p+6) == 'A') && (* (p+7) == 'A') && (* (p+8) == 'G') ) nT = nT + 1;
}

printf("No. of matches of single motif are \%d\n",singlemotif);
printf("count of AAAGAAAAA = \%d\n",nA);
printf(" count of AAAGAAAG = \%d\n",nG);
printf(" count of AAAGCAAA = \%d\n",nC);
printf("count of AAAGTAAAG = \%d\n",nT);

    • f = open("C:\Users\Ujjwal\Downloads\chromosome1.txt","r")
    buff = f.read()
    values = []

    for i in range(0, 27):
        values[i] = 0
        for x in range(0, len(buff) - i):
            if (buff[x:x+4] == "AAAG") and (buff[x+4:x+4+i+4] == "AAAG"): values[i] = values[i] + 1

    print("printing frequency below\n")
    for i in range(0, 27):
        print( 'Nucleotides AAAGnAAAG separated by + str(values[i]) + ','' + str(values[i]) )

    • f = open("C:\Users\Ujjwal\Downloads\chromosome1.txt","r")
    buff = f.read()
    values = []
Results and discussions

AAAGn7CTTT sequence is highly preferred in A. thaliana genome

Dof proteins, which are typically composed of 200–400 amino acids, are defined as DNA-binding proteins that have a highly conserved Dof domain. The strong similarity among Dof DNA-binding domains suggested that all Dof proteins display similar DNA-binding specificity. Indeed, an AAAG sequence or its reversibly oriented sequence, CTTT, is always found in the binding sequences of individual Dof proteins (Chen et al., 1996; dePaolis et al., 1996; Kang and Singh, 2000; Mena et al., 1998; Plesch et al., 2001;
Washio, 2001; Yanagisawa and Izui, 1993) except a pumpkin Dof protein (AOBP) that recognizes an AGTA motif (Kisu et al., 1998). In *A. thaliana*, two AAAGs separated by one nucleotide is a known binding site for the OBP-1 protein (Yanagisawa, 2002). Similarly clusters of AAAG sites have been shown to additively contribute to guard cell-specificity of AtMYB60 promoter in guard cells (Cominelli et al., 2011).

With an intention to discover potential new DOF binding sites in *A. thaliana*, the frequency of two AAAG or CTTT motifs separated by an increasing distance was carried out. The frequency of AAAGAAAG without any spacer has a maximum occurrence of 12,738 as shown in Table 1 and Fig. 1. As we increase the spacer length, the frequency of occurrences started decreasing.

Table 1

| chr1 | chr2 | chr3 | chr4 | chr5 | Total |
|------|------|------|------|------|-------|
| 0    | 3224 | 2171 | 2501 | 1934 | 2908  | 12,738 |
| 1    | 2951 | 1873 | 2282 | 1711 | 2499  | 11,316 |
| 2    | 3314 | 2088 | 2546 | 2302 | 3215  | 13,465 |
| 3    | 2635 | 1755 | 2112 | 1693 | 2390  | 10,585 |
| 4    | 2732 | 1751 | 2038 | 1594 | 2377  | 10,492 |
| 5    | 2577 | 1746 | 2076 | 1570 | 2256  | 10,225 |
| 6    | 2529 | 1792 | 2107 | 1541 | 2373  | 10,342 |
| 7    | 2407 | 1663 | 2278 | 1589 | 2278  | 10,215 |
| 8    | 2533 | 1644 | 2134 | 1548 | 2341  | 10,200 |
| 9    | 2201 | 1454 | 1720 | 1330 | 2026  | 8731  |
| 10   | 2184 | 1518 | 1737 | 1390 | 2067  | 8860  |
| 11   | 2308 | 1543 | 1719 | 1365 | 2073  | 9008  |
| 12   | 2169 | 1454 | 1763 | 1274 | 2021  | 8681  |
| 13   | 2194 | 1438 | 1671 | 1352 | 1939  | 8594  |
| 14   | 2497 | 1501 | 1909 | 1428 | 2080  | 9415  |
| 15   | 2172 | 1435 | 1738 | 1348 | 2022  | 8715  |
| 16   | 2556 | 1507 | 1789 | 1439 | 2142  | 9433  |
| 17   | 2482 | 1690 | 2028 | 1583 | 2331  | 10,114|
| 18   | 2154 | 1459 | 1888 | 1345 | 1925  | 8771  |
| 19   | 2230 | 1476 | 1819 | 1378 | 1917  | 8820  |
| 20   | 2338 | 1553 | 1776 | 1418 | 2105  | 9190  |
| 21   | 2144 | 1430 | 1646 | 1296 | 1939  | 8455  |
| 22   | 2129 | 1308 | 1609 | 1267 | 1881  | 8194  |
| 23   | 2159 | 1467 | 1733 | 1435 | 2015  | 8809  |
| 24   | 2254 | 1443 | 1689 | 1549 | 1997  | 8932  |
| 25   | 2147 | 1504 | 1721 | 1400 | 1923  | 8695  |

![Fig. 1](image-url) Frequency of two AAAG motifs separated by all possible distances till 25 bp across the five chromosomes of *Arabidopsis thaliana*. 

Washio, 2001; Yanagisawa and Izui, 1993) except a pumpkin Dof protein (AOBP) that recognizes an AGTA motif (Kisu et al., 1998). In *A. thaliana*, two AAAGs separated by one nucleotide is a known binding site for the OBP-1 protein (Yanagisawa, 2002). Similarly clusters of AAAG sites have been shown to additively contribute to guard cell-specificity of AtMYB60 promoter in guard cells (Cominelli et al., 2011). With an intention to discover potential new DOF binding sites in *A. thaliana*, the frequency of two AAAG or CTTT motifs separated by an increasing distance was carried out.

The frequency of AAAGAAAG without any spacer has a maximum occurrence of 12,738 as shown in Table 1 and Fig. 1. As we increase the spacer length, the frequency of occurrences started decreasing. There
was a slight increase in frequency for the spacer length 14–17. Statistical analyses (data not shown) indicated them to be non significant as the deviation was essentially within 10–15%. Similar trend was observed for (CTTT)nCTTT as shown in Table 2 and Fig. 2.

A very interesting observation was made when we looked for combination of AAAG and CTTT sequences. An unexpected high frequency was observed for AAAGn7CTTT. The frequency of occurrence was observed as

Table 2
Frequency of two CTTT motifs separated by all possible distances (till 25 bp), across the five chromosomes. 'n' represents the intervening distance between the motifs. The second column displays the value of 'n'.

| Chr | Chr1 | Chr2 | Chr3 | Chr4 | Chr5 | Total |
|-----|------|------|------|------|------|-------|
| 0   | 3195 | 2086 | 2447 | 1896 | 2764 | 12,388|
| 1   | 3192 | 1910 | 2262 | 1755 | 2627 | 11,746|
| 2   | 3269 | 2168 | 2465 | 2139 | 3060 | 13,101|
| 3   | 2648 | 1706 | 2150 | 1677 | 2497 | 10,678|
| 4   | 2616 | 1688 | 2081 | 1547 | 2374 | 10,306|
| 5   | 2582 | 1723 | 2078 | 1606 | 2452 | 10,441|
| 6   | 2591 | 1740 | 2014 | 1612 | 2344 | 10,301|
| 7   | 2402 | 1708 | 2183 | 1664 | 2292 | 10,249|
| 8   | 2416 | 1729 | 2097 | 1576 | 2287 | 10,105|
| 9   | 2286 | 1448 | 1836 | 1390 | 2053 | 9013  |
| 10  | 2260 | 1528 | 1698 | 1384 | 2108 | 8978  |
| 11  | 2326 | 1531 | 1770 | 1381 | 2212 | 9220  |
| 12  | 2231 | 1484 | 1683 | 1293 | 1939 | 8630  |
| 13  | 2143 | 1484 | 1683 | 1407 | 1896 | 8613  |
| 14  | 2360 | 1606 | 1837 | 1435 | 2136 | 9374  |
| 15  | 2493 | 1523 | 1656 | 1431 | 1978 | 9081  |
| 16  | 2227 | 1494 | 1829 | 1477 | 2327 | 9354  |
| 17  | 2402 | 1673 | 2043 | 1568 | 2320 | 10,006|
| 18  | 2237 | 1482 | 1797 | 1318 | 1985 | 8819  |
| 19  | 2240 | 1444 | 1657 | 1353 | 2001 | 8695  |
| 20  | 2305 | 1555 | 1746 | 1402 | 2101 | 9109  |
| 21  | 2180 | 1459 | 1610 | 1402 | 2045 | 8996  |
| 22  | 2124 | 1401 | 1578 | 1278 | 1946 | 8327  |
| 23  | 2218 | 1527 | 1747 | 1361 | 2014 | 8867  |
| 24  | 2156 | 1428 | 1625 | 1321 | 1916 | 8446  |
| 25  | 2219 | 1457 | 1703 | 1373 | 1986 | 8738  |

**Fig. 2.** Frequency of two CTTT motifs separated by all possible distances till 25 bp, across the five chromosomes of *Arabidopsis thaliana.*
14,977 which is more than two times the predecessor whose frequency is 7177 as shown in Table 3 and Fig. 3. However, when we change the orientation to CTTTn7 AAAG this tendency was not observed as shown in Table 4. The other implication of this is that transcriptional factor binding is direction specific. Not all AAAG motifs in plant promoters are targets of the Dof domain proteins. However, since an AAAG and a CTTT motif separated by a distance of 7 bp is present in an exceptionally high frequency, we think it is highly likely that this sequence combination may have a functional significance yet to be discovered.

**Table 3**

Frequency of a AAAG and a CTTT motif separated by all possible distances (till 25 bp), across the five chromosomes. ‘n’ represents the intervening distance between the motifs. The second column displays the value of ‘n’.

| AAAGnCTTT | chr1 | chr2 | chr3 | chr4 | chr5 | Total |
|-----------|------|------|------|------|------|-------|
| 0         | 2379 | 1352 | 1417 | 1570 | 2320 | 9058  |
| 1         | 1504 | 910  | 1236 | 905  | 1384 | 5939  |
| 2         | 1187 | 792  | 921  | 736  | 1018 | 4654  |
| 3         | 1398 | 903  | 993  | 792  | 1205 | 5291  |
| 4         | 1199 | 842  | 992  | 957  | 1190 | 5180  |
| 5         | 1308 | 863  | 995  | 795  | 1221 | 5182  |
| 6         | 1853 | 1205 | 1396 | 1069 | 1654 | 7177  |
| 7         | 3827 | 2482 | 2854 | 2358 | 3456 | 14,577|
| 8         | 1546 | 990  | 1201 | 922  | 1350 | 6009  |
| 9         | 1534 | 1026 | 1197 | 994  | 1405 | 6156  |
| 10        | 1674 | 1050 | 1183 | 968  | 1366 | 6241  |
| 11        | 1544 | 1006 | 1218 | 1083 | 1633 | 6484  |
| 12        | 1620 | 976  | 1201 | 977  | 1470 | 6244  |
| 13        | 1557 | 1033 | 1180 | 974  | 1358 | 6102  |
| 14        | 1660 | 1081 | 1245 | 1012 | 1385 | 6383  |
| 15        | 1687 | 1119 | 1309 | 1032 | 1479 | 6626  |
| 16        | 1664 | 1107 | 1335 | 1016 | 1575 | 6697  |
| 17        | 1715 | 1092 | 1251 | 1135 | 1871 | 7064  |
| 18        | 1685 | 1119 | 1508 | 970  | 1454 | 6736  |
| 19        | 1518 | 992  | 1189 | 1014 | 1388 | 6101  |
| 20        | 1649 | 1040 | 1231 | 925  | 1354 | 6199  |
| 21        | 1673 | 1111 | 1269 | 951  | 1412 | 6416  |
| 22        | 1635 | 1046 | 1298 | 955  | 1454 | 6388  |
| 23        | 1548 | 1066 | 1196 | 937  | 1485 | 6232  |
| 24        | 1631 | 1059 | 1243 | 969  | 1461 | 6363  |
| 25        | 1655 | 1081 | 1278 | 977  | 1521 | 6512  |

Fig. 3. Frequency of CTTT and AAAG motifs separated by all possible distances till 25 bp, across the five chromosomes of Arabidopsis thaliana.
A and G are preferred as flanking nucleotides

We were interested to know which residues predominate in the flanking of AAAG sequence. Such studies are very important because many studies indicate that flanking sequences are very important for binding specificity (Foster et al., 1994; Izawa et al., 1993). We changed one nucleotide at a time following AAAG. As shown in Table 5, A and G predominate as flanking residues although there is an exception when (AAAG)–(AAAG) is separated by one nucleotide where the frequency of G flanking is 1918 which is less than C which is 2057. In all other cases G dominates as a flanking sequence over C and T.

Conclusions

The promoter region of many genes contain multiple binding sites for the same transcription factor. One possibility is that individuals with multiple, redundant binding sites have higher fitness. Cis regulatory element multiplicity has been correlated with several gene functionalities like Promoters containing multiple sites evolve more slowly. In this paper we focused on the multiplicity of AAAG sequence with varied spacer lengths and also in combination with CTTT sequence. We report that AAAGn7 CTTT is a preferred sequence in the genome of A. thaliana. This information will be useful for designer promoters where specific interactions could be directed.

Acknowledgment

This work is part of UGC funded project of Dr Sandhya Mehrotra and Dr Rajesh Mehrotra. The authors are grateful to the administration of Birla Institute of Technology and Sciences, Pilani, Rajasthan for providing logistic support.
Table 5
Frequency of flanking nucleotide between two AAAG motifs separated by an increasing sequence length across five chromosomes.

| AAAG?_AAAG | chr1 | chr2 | chr3 | chr4 | chr5 | Total |
|------------|------|------|------|------|------|-------|
| A          | 1480 | 962  | 1143 | 885  | 1264 | 5734  |
| G          | 487  | 306  | 399  | 296  | 430  | 1918  |
| C          | 551  | 330  | 413  | 286  | 477  | 2057  |
| T          | 433  | 275  | 327  | 244  | 328  | 1607  |
| AA         | 725  | 441  | 574  | 403  | 646  | 2789  |
| AG         | 402  | 248  | 286  | 234  | 327  | 1479  |
| AC         | 189  | 124  | 135  | 115  | 366  | 929   |
| AT         | 196  | 118  | 178  | 123  | 173  | 788   |
| AAA        | 286  | 170  | 189  | 156  | 250  | 1051  |
| AAG        | 152  | 105  | 143  | 96   | 136  | 632   |
| AAC        | 74   | 41   | 59   | 48   | 55   | 277   |
| AAT        | 65   | 37   | 28   | 36   | 43   | 209   |
| AAAA       | 187  | 98   | 109  | 81   | 141  | 616   |
| AAAG       | 122  | 79   | 95   | 65   | 118  | 479   |
| AAAC       | 52   | 35   | 43   | 23   | 31   | 184   |
| AAAT       | 33   | 18   | 23   | 19   | 38   | 131   |
| AAAA       | 76   | 42   | 69   | 49   | 66   | 302   |
| AAAGG      | 33   | 26   | 45   | 30   | 37   | 171   |
| AAACG      | 22   | 21   | 13   | 14   | 15   | 85    |
| AAATG      | 15   | 13   | 11   | 6    | 18   | 63    |
| AAAAAG     | 55   | 37   | 45   | 28   | 43   | 208   |
| AAAGAC     | 20   | 13   | 18   | 9    | 27   | 87    |
| AAAGCG     | 4    | 6    | 5    | 6    | 14   | 35    |
| AAAAAC     | 8    | 1    | 6    | 6    | 2    | 22    |
| AGAGA      | 102  | 85   | 87   | 72   | 108  | 454   |
| AGGGG      | 47   | 32   | 39   | 25   | 49   | 192   |
| AGGCC      | 32   | 30   | 30   | 27   | 26   | 145   |
| AGTGT      | 55   | 34   | 47   | 41   | 59   | 236   |

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.mgene.2014.05.003.

References

Chen, W., et al., 1996. The promoter of an H2O2-inducible, Arabidopsis glutathione-S-transferase gene contains closely linked OBFBinding sites. Plant J. 10, 955–966.

Cominelli, E., et al., 2011. DOF-binding sites additively contribute to guard cell-specificity of AtMYB60 promoter. BMC Plant Biol. 11, 162.

dePaolis, A., et al., 1996. A rolB regulatory to a new class of single zinc finger plant proteins. Plant J. 10, 215–223.

Foster, R., Izawa, T., Chua, N.H., 1994. Plant bZIP proteins gather at ACGT elements. FASEB J. 8, 192–200.

Izawa, T., Foster, R., Chua, N.H., 1993. Plant bZIP protein binding specificity. J. Mol. Biol. 230 (4), 1131–1144 (20).

Kang, H.-G., Singh, K., 2000. Characterization of salicylic acid-responsive, Arabidopsis Dof domain proteins: over expression of OBP3 leads to growth defects. Plant J. 21, 329–339.

Kisu, Y., et al., 1998. Characterization and expression of a new class of zinc finger protein that binds to silencer region of ascorbate oxidase gene. Plant Cell Physiol. 39, 1054–1064.

Mehrotra, R., Mehrotra, S., 2010. Promoter activation by ACCT in response to salicylic and abscisic acids is differentially regulated by the spacing between two copies of the motif. J. Plant Physiol. 167, 1214–1218.

Mehrotra, Rajesh, Panwar, Jitendra, 2009. Dimerisation of GT element interferes negatively with gene activation. J. Genet. 88, 257–260.

Mehrotra, Rajesh, Yadav, Amit, Bhalaria, Purva, et al., 2012. Evidence for directed evolution of larger size motif in Arabidopsis thaliana genome. Sci. World J. http://dx.doi.org/10.1101/2012/983528.2012 (Article ID 983528).

Mehrotra, Rajesh, Zutsi, Ipshita, Sethi, Sachin, et al., 2013. Patterns and evolution of ACCT repeat cis element landscape across four plant genomes. BMC Genomics 14, 203 (http://www.biomedcentral.com/1471-2164/14/203/abstract).
Mena, M., et al., 1998. An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin-box of a native b-hordein promoter in barley endosperm. Plant J. 16, 53–62.

Plesch, G., et al., 2001. Involvement of TAAAG elements suggests a role for Dof transcription factors in guard cell-specific gene expression. Plant J. 28, 455–464.

Riechmann, J.L., Heard, J., Martin, G., et al., 2000. Arabidopsis transcription factors: genome wide comparative analysis among eukaryotes. Science 290, 2105–2110.

Sawant, S.V., Kiran, K., Mehrotra, R., et al., 2005. A variety of synergistic and antagonistic interactions mediated by cis-acting DNA motifs regulate gene expression in plant cells and modulate stability of the transcription complex formed on a basal promoter. J. Exp. Bot. 56, 2345–2353.

Vicente-Carbajosa, J., et al., 1997. A maize zinc-finger protein binds the prolamin box in zeingene promoters and interacts with the basicleucine zipper transcriptional activator. Proc. Natl. Acad. Sci. U. S. A. 94, 7685–7690.

Washio, K., 2001. Identification of Dof proteins with implication in the gibberellin-regulated expression of a peptidase gene following the germination of rice grains. Biochim. Biophys. Acta 1520, 54–62.

Wray, G.A., Hahn, M.W., Abouheif, E., et al., 2003. The evolution of transcriptional regulation in eukaryotes. Mol. Biol. Evol. 20, 1377–1419.

Yanagisawa, S., 2002. The Dof family of plant transcription factors. Trends Plant Sci. 7 (12).

Yanagisawa, S., Izui, K., 1993. Molecular cloning of two DNA-binding proteins of maize that are structurally different but interact with the same sequence motif. J. Biol. Chem. 268, 16028–16036.

Yanagisawa, S., Sheen, J., 1998. Involvement of maize Dof zinc finger proteins in tissue-specific and light-regulated gene expression. Plant Cell 10, 75–89.

Zhang, B., et al., 1995. Interactions between distinct types of DNA binding proteins enhance binding to ocs element promoter sequences. Plant Cell 7, 2241–2252.