



Sorghum bicolor L. CAMTA Transkripsiyon Faktörlerinin Genom Çaplı Analizi*

Damla KIZILKAYA^{1,a} Ayşe Gül KASAPOĞLU^{2,b} Arash HOSSEINPOUR^{3,4,c}

Kamil HALİLOĞLU^{3,d} Selman MUSLU^{2,e} Emre İLHAN^{2,f,**}

¹Ankara Yıldırım Beyazıt Üniversitesi, Geleneksel ve Tamamlayıcı Tıp Araştırma ve Uygulama Merkezi, Keçiören/Ankara, Türkiye

²Erzurum Teknik Üniversitesi, Fen Fakültesi, Moleküler Biyoloji ve Genetik Bölümü, Erzurum, Türkiye

³Atatürk Üniversitesi, Ziraat Fakültesi, Tarla Bitkileri Bölümü, Erzurum, Türkiye

⁴Crop and Horticultural Science Research Department, Ardabil Agricultural and Natural Resources Research and Education Center, AREEO, Ardabil (Moghan), Iran

**Sorumlu yazar e-mail: emre.ilhan@erzurum.edu.tr

doi: 10.17097/ataunizfd.690138

Geliş Tarihi (Received): 18.02.2020 Kabul Tarihi (Accepted): 29.05.2020 Yayın Tarihi (Published): 25.09.2020

ÖZ: Kalmmodulin bağlayıcı transkripsiyon aktivatörü olan CAMTA gen ailesi, bitki familyasında karakterize edilmiş kalmmodulin bağlayıcı transkripsiyon faktörleridir. CAMTA gen ailesi hastalıklara karşı direnç, biyotik ve abiyotik stres etmenlerine karşı yanıt gibi çeşitli biyolojik süreçlerde önemli roller üstlenmektedir. Bu çalışmada, sorgum (*Sorghum bicolor* L.) genomunda 7 CAMTA geni belirlendi ve kök ve sürgün dokularında Sobic-CAMTA genlerinin ifade profilleri analiz edildi. Sobic-CAMTA proteinlerinin moleküler ağırlıkları ve uzunlukları sırasıyla 95,22 kDa (Sobic-CAMTA-6) ile 114,86 kDa (Sobic-CAMTA-5) ve 845 (Sobic-CAMTA-6) ile 1030 (Sobic-CAMTA-5) amino asit arasındadır. Sobic-CAMTA genleri arasında tahmini olarak belirlenen ekzonların sayısı en düşük 10 en yüksek 13'tür. İzoelektrik noktaları ise 5,55 (Sobic-CAMTA-5.) ile 8,36 (Sobic-CAMTA-4) arasında değişmektedir. Sobic-CAMTA-2/Sobic-CAMTA-3 tandem duplike genler iken, Sobic-CAMTA-3/Sobic-CAMTA-5 ve Sobic-CAMTA-6/Sobic-CAMTA-7 ise segmental duplike genler olarak tespit edilmiştir. *S. bicolor* L., *Arabidopsis thaliana* (L.) Heynh ve *Zea mays* L. CAMTA proteinleri kullanılarak çizilen filogenetik ağaca göre 3 ana grup (A, B ve C) elde edilmiştir. Sobic-CAMTA genlerinin ifade profilleri, *S. bicolor* L. bitkisinin farklı dokularına farklı azot kaynağı uygulaması ile belirlenmiştir. Bu çalışmanın sonuçları, sorgum bitkisinde CAMTA transkripsiyon faktörü gen ailesinin moleküler yapısının anlaşılması için önemli bilgiler sağlayacaktır.

Anahtar Kelimeler: Ankrin, Kalmmodulin, CG-1 DNA bağlanma domaini, IQ motifi, *Sorghum bicolor* L.

Genome wide analysis of *Sorghum bicolor* L. CAMTA Transcription Factors

ABSTRACT: Calmodulin binding transcription activators (CAMTAs) are well-characterized in the plants. The CAMTA gene family plays an important role in a variety of biological processes, such as resistance to diseases or response to biotic and abiotic factors. In this study, 7 CAMTA genes were identified in the *Sorghum bicolor* L. genome and expression profiles of Sobic-CAMTA genes in root and shoot tissues were analyzed. The molecular weight and length of Sobic-CAMTA proteins ranged from 95.22 kDa (Sobic-CAMTA-6) to 114.86 kDa (Sobic-CAMTA-5) and 845 (Sobic-CAMTA-6) to 1030 (Sobic-CAMTA-5) amino acids, respectively. Estimated number of exons determined among Sobic-CAMTA genes was between 10 and 13. The isoelectric points ranged from 5.55 (Sobic-CAMTA-5.) to 8.36 (Sobic-CAMTA-4). Sobic-CAMTA-2/Sobic-CAMTA-3 tandem duplicated genes, while Sobic-CAMTA-3/Sobic-CAMTA-5 and Sobic-CAMTA-6/Sobic-CAMTA-7 were identified as segmentally-duplicated genes. According to the phylogenetic tree, which drawn using CAMTA proteins of *S. bicolor* L., *Arabidopsis thaliana* (L.) Heynh and *Zea mays* L., three main groups (A, B and C) were obtained. The expression profiles of Sobic-CAMTA genes were determined by applying different nitrogen sources to different tissues of *S. bicolor* L. The results of this study would provide important information to understand the molecular structure of the CAMTA transcription factor gene family in sorghum.

Keywords: Ankyrin, Calmodulin, CG-1 DNA binding domain, IQ motif, *Sorghum bicolor* L.

Bu makaleye atıfta bulunmak için / To cite this article: *Sorghum bicolor* L. CAMTA Transkripsiyon Faktörlerinin Genom Çaplı Analizi. Kızılkaya, D., Kasapoğlu, A.G., Hossein Pour, A., Haliloğlu, K., Muslu, S., İlhan, E., 2020. Atatürk Üniv. Ziraat Fak. Derg., 51 (3): 267-278. doi: 10.17097/ataunizfd.690138

^aORCID: <https://orcid.org/0000-0002-4967-1249> ^bORCID: <https://orcid.org/0000-0002-6447-4921>

^cORCID: <https://orcid.org/0000-0003-2611-8034> ^dORCID: <https://orcid.org/0000-0002-4014-491X>

^eORCID: <https://orcid.org/0000-0003-4777-0726> ^fORCID: <https://orcid.org/0000-0002-8404-7900>

*Bu çalışma "I. International Iğdır Congress on Multidisciplinary Studies (06-08 Kasım 2018, Iğdır, Türkiye)"de sunulmuş ve özet kitapçığında özeti basılmıştır.

GİRİŞ

Transkripsiyon faktörleri (TF'ler) çok spesifik bir durumda DNA'ya bağlanan proteinlerdir ve gen ifadesini baskılar veya iyileştirirler. Çok çeşitli uyarılara yanıtta, TF'leri hücre döngüsü ve fizyolojik yanıtlardan hücre farklılaşması ve gelişimine kadar olan pek çok önemli biyolojik işlevde görev alırlar (de Mendoza et al., 2013). İki değerlikli ve ökaryotik organizmalarda ikincil mesajcı olan Ca^{2+} , bitkiler tarafından kullanılır. Ca^{2+} sinyalleri, Ca^{2+} bağlama proteinlerinin aktivitesiyle algılanır (Wei et al., 2017). Bitkilerde son zamanlarda 3 sınıf Ca^{2+} mesajcısı belirlenmiştir. Bunlar kalmodulinler ve kalmodulin benzeri proteinler (CaM ve CML'ler), kalsiyum-bağlı protein kinazlar (CPDK) ve kalsinörin B-benzeri proteinler (CBL)'dir (Ludwig et al., 2004). Bugüne kadar CaM'ların, kalmodulin-bağlama transkripsiyon aktivatörleri (CAMTA) de dahil en az 90 transkripsiyon faktörünü düzenleyebildiği tespit edilmiştir (Du and Poovaiah, 2004).

CAMTA'lar bitkilerde ve hayvanlarda evrimsel olarak korunduğu bildirilen ve iyi karakterize edilmiş kalmodulin bağlama transkripsiyon faktörlerinin bir ailesidir (Pant et al., 2018). CAMTA ailesi ilk olarak tütün (NtER1) bitkisinde tespit edilmiştir (Yang and Poovaiah, 2000). CAMTA'lar; amino ucunda CG-1 DNA bağlama domaini, devamında spesifik olmayan DNA bağlanmasında rol oynayan bir TIG domaini, protein-protein etkileşimlerine aracılık eden bir Ankrin (ANK) tekrar domaini ve proteinin amino ucu ve karboksil ucu arasında bulunan Ca^{2+} bağımlı CaM bağlama domaini (CaCMD) ve Ca^{2+} 'dan bağımsız bir yolda CaM ile bağlanan çeşitli sayılarda IQ motifleri (IQXXRGXXR) içerir (Hu et al., 2015; Wei et al., 2017; Pant et al., 2018). CAMTA, hedef genlerin promotör bölgesindeki spesifik cis elementlerini [(A/C)CGCG(C/G/T) ve (A/C)CGTGT = absisik asit (ABA) – yanıt elementi (ABRE)] tanıyarak ve bağlanarak transkripsiyonel düzenlemeyi gerçekleştirir ve böylece hedef genlerin ifadesini düzenler (Choi et al., 2005; Finkler et al., 2007; Shanguan et al., 2014; Pant et al., 2018). Şimdiye kadar *Arabidopsis*, *Medicago truncatula*, soya fasulyesi ve mısır gibi farklı bitki türlerinden bazı CAMTA genleri tanımlanmıştır. *Arabidopsis*'te tanımlanan 6 CAMTA proteini bulunmakta olup, Ia, IIa, IIIc ve IIIe olmak üzere dört sınıfa ayrılmaktadır (Hu et al., 2015; Wei et al., 2017). *Arabidopsis*'te AtCAMTA3 bitki bağımsızlığı bakımından negatif bir düzenleyicidir (Du et al., 2009). Diğer taraftan 6 *Arabidopsis* CAMTA geni aşırı sıcaklık, tuz ve yaralanma gibi çoklu çevresel sinyallere yanıt verir. Ayrıca bu genler, etilen ve ABA gibi hormonlar ve metil jasmonat, H_2O_2 ve salisilik asit gibi sinyal molekülleriyle teşvik edilebilirler (Yang and Poovaiah, 2002). AtCAMTA3 üzerine yapılan

çalışmalar soğuk stresine yanıtta görev alan CBF2 ifadesinin pozitif bir düzenleyicisi olduğunu göstermektedir. AtCAMTA1 ve AtCAMTA3'ün mutantları ile yapılan çalışmada ise donmaya karşı düşük tolerans gösterdikleri tespit edilmiştir (Kim et al., 2013).

Sorgum [*Sorghum bicolor* (L.) Moench], C4 fotosentez yolunu kullanmakta olup, marjinal alanlara yüksek adaptasyon göstermesi, azotu kullanması ve aşırı büyüyen fazla miktarda biyokütleyle sahip olması nedeniyle dünyada en fazla ekimi yapılan tahıl bitkileri arasındadır. Ayrıca geniş bir germplasmanın oluşu geniş bir genetik çeşitliliğe sahip olmasına sebep olmaktadır. Büyüme özellikleri ve kullanımına göre tane tahılı, tatlı sorgum, yem bitkisi ve yakıt gibi sınıflara ayrılmaktadır (Xu et al., 2018). Sorgum, kendine döllen bir bitki olup genetik ve morfolojik çeşitliliği olan önemli bir diploid bitkidir (Menz et al., 2002).

Bitkilerde azot; amino asit, protein, nükleik asit, amid, klorofil gibi önemli fonksiyonlarda yer alan organik bileşiklerin yapısına katılması nedeniyle, topraktan alınan besinlerden en fazla ihtiyaç duyulan elementtir (Miller and Cramer, 2005).

Yapılan literatür araştırmalarında sorgum bitkisinde CAMTA gen ailesinin genom çaplı belirlenmesi üzerine herhangi bir bilgiye rastlanmamıştır. İlk kez yapılan bu çalışmada CAMTA gen ailesi üyelerinin sorgum genomunda in siliko yöntemler kullanılarak genom çaplı olarak belirlenmesi amaçlanmıştır. Sorgum genomunda biyotik ve abiyotik streslere toleransın genetik mekanizmasının daha iyi anlaşılmasını sağlamak amacıyla hem yapısal hem de filogenetik düzeyde CAMTA proteinlerini kodlayan diziler belirlenerek, karakterize edilmiştir.

MATERYAL VE METOT

Sorghum bicolor L. genomunda CAMTA proteinlerinin belirlenmesi

Sorghum bicolor L. genomundaki CAMTA gen ailesinin protein dizilerine Pfam veritabanından elde edilen Pfam Erişim Numarası (PF00612) kullanılarak Phytozome v12.1 veritabanından ulaşılmıştır (Phytozome, 2020). Sorgum genomundaki (McCormick et al., 2018) tüm muhtemel CAMTA proteinlerini belirlemek için, hem Phytozome v12.1 veritabanındaki blastp hem de gizli Markov model (HMM) (HMMER, 2020) araması varsayılan parametrelerle kullanılmıştır. İlişkili olmayan diziler "decrease redundancy aracı" (ExPASy, 2020a) kullanılarak elde edilmiştir. İlişkili dizilerdeki CAMTA domaininin varlığı HMMER (HMMER, 2020) veritabanı kullanılarak araştırılmıştır. Elde edilen CAMTA proteinlerinin moleküler ağırlığı ve

teorik izoelektrik noktası (pI) “ProtParam aracı” (ExPASy, 2020b) kullanılarak tespit edilmiştir.

CAMTA genlerinin yapısı, fiziksel yerleri, gen duplikasyonu ve korunmuş motiflerinin belirlenmesi

CAMTA proteinlerinin ekzon-intron bölgeleri hakkında bilgi edinmek için, Gene Structure Display Server v2.0 kullanılmıştır (Guo et al., 2007). Hem genom dizileri hem de kodlanan DNA (CDS) dizileri CAMTA genlerinin pozisyon bilgisini tahmin etmede kullanılmıştır. Phytozome v12.1 veritabanı kullanılarak, CAMTA genlerinin kromozomal yerleri ve büyüklükleri tespit edilmiştir.

CAMTA genleri tüm *S. bicolor* L. kromozomlarına işaretlenmiş ve MapChart programı ile çizilmiştir (Voorrips, 2002). Gen duplikasyon olayları “Bitki Genom Katlanma Veritabanı, BGKV (Plant Genome Duplication Database)” sağlayıcısı (PGDD, 2020) kullanılarak belirlenmiştir. CAMTA genlerinin ikili çiftleri arasındaki Non-sinonim oranları (Ka), sinonim oranları (Ks) ve evrimsel zorlamalar (Ka/Ks) PAL2NAL (Suyama et al., 2006) PAML ara yüz aracında (Yang, 2007) hesaplanmıştır.

CAMTA proteinlerinin ilave korunmuş motiflerini belirlemek için, “Multiple EM for Motif Elicitation (MEME) Aracı” kullanılmıştır (Bailey et al., 2006). Minimum/maksimum genişlik ve Motiflerin maksimum sayısı için sınırlar sırasıyla 2, 50 ve 15 olarak ayarlanmıştır. Motif bölgeleri 2 ile 300 arasındadır. Bölge dağılımı tekrarların herhangi bir sayısı olarak ayarlanmıştır. Belirlenen motifler InterPro veritabanının varsayılan ayarları kullanılarak taranmıştır (Quevillon et al., 2005). Ayrıca korunmuş bölge dizi analizleri için CAMTA domainlerinin dizi logo analizleri WEBLOGO online web aracı kullanılarak çizilmiştir (Crooks et al., 2004).

Filogenetik analizler ve dizi hizalama

Filogenetik analizler, 1000 tekrarlı bootstrap değeri ile Neighbor-joining (NJ; Komşu katılım) metoduna göre yapılmıştır. CAMTA protein dizileri ClustalW kullanılarak hizalanmıştır (Thompson et al., 1997). MEGA v7 programı kullanılarak filogenetik ağaç elde edilmiştir (Tamura et al., 2011). Ağaç, İnteraktif Yaşam Ağacı (iTOL) ara yüzü kullanılarak şekillendirilmiştir (Letunic and Bork, 2011)

Sinteni analizi

S. bicolor ile *A. thaliana* ve *Z. mays* CAMTA genlerinin ortologları Plant Genome Duplication Veritabanı (PGDD, 2018) kullanılarak belirlenmiştir (Lee et al., 2013). Ortolog genlerin protein dizileri Phytozome v12.1 veritabanından elde edilmiştir. Sinteni haritası “iTAK- Plant Transcription factor

&Protein Kinase Identifier and Classifier” kullanılarak çizilmiştir (Zheng et al., 2016)

İN SILIKO GEN İFADE ANALİZİ

Illumina RNAseq veri setleri Phytozome v12.1 veritabanı kullanılarak elde edilmiştir. Sobic-CAMTA genlerinin *S. bicolor* bitkisinin farklı dokularında farklı azot kaynağı uygulaması ile CAMTA transkripsiyon faktörünün ifade profili belirlenmiştir. İn siliko ifade profilleri FPKM (Dizilenmiş her milyon baz çifti transkript dizisinin her kilobaz fragmanlarının beklenen sayısı) birimlerinde Cufflinks ile hesaplanmıştır (Trapnell et al., 2013). FPKM değerleri log₂'ye dönüştürülerek CIMMiner (CIMMiner, 2020) algoritması ile heatmap elde edilmiştir.

BULGULAR VE TARTIŞMA

***S. bicolor* L. genomundaki CAMTA gen ailesinin belirlenmesi**

Phytozome v12.1 veritabanında bulunan *S. bicolor* genomunda, Pfam veritabanından elde edilen Pfam erişim numarası kullanılarak anahtar kelime taraması (keyword search) uygulanmıştır. Bu tarama sonucunda CAMTA protein homolojisine sahip yedi gen tespit edilmiştir. CAMTA domainlerinin varlığı HMMER veritabanı kullanılarak doğrulanmıştır. Ayrıca varsayılan CAMTA domaini bulunan genlerden ilişkili olmayan diziler uzaklaştırıldıktan sonra yeniden HMMER veritabanı ile doğrulanmıştır. *S. bicolor* genomundan elde edilen yedi aday CAMTA geni Çizelge 1’de listelenmiştir. Yapılan çalışmalarda *Manihot esculenta* (5), *Ricinus communis* (4), *Linum usitatissimum* (8), *Populus trichocarpa* (7), *Medicago truncatula* (4), *Phaseolus vulgaris* (8), *Glycine max* (14), *Cucumis sativus* (3), *Prunus persica* (4), *Malus domestica* (8), *Fragaria vesca* (4), *Arabidopsis thaliana* (6), *Arabidopsis lyrata* (6), *Capsella rubella* (6), *Brassica rapa* (9), *Thellungiella halophila* (6), *Carica papaya* (2), *Gossypium raimondii* (7), *Theobroma cacao* (5), *Citrus sinensis* (5), *Citrus clementina* (4), *Eucalyptus grandis* (4), *Vitis vinifera* (4), *Solanum tuberosum* (7), *Solanum lycopersicum* (7), *Mimulus guttatus* (4), *Aquilegia coerulea* (5), *Zea mays* (9), *Setaria italica* (7), *Panicum virgatum* (6), *Oryza sativa* (6), *Brachypodium distachyon* (7), *Selaginella moellendorffii* (1), *Physcomitrella patens* (3) ve *Phaseolus vulgaris* (8) gibi bitki genomlarında CAMTA gen ailesinin tespit edildiği bildirilmiştir. Ayrıca su yosunu, *Volvox carteri*, *Coccomyxa subellipsoidea* C-169, *Micromonas pusilla* CCMP1545, *Micromonas pusilla* RCC299, *Ostreococcus hicmarinus* üzerinde yapılan çalışmalarda CAMTA gen ailesi üyeleri belirlenmemiştir (Rahman et al., 2016). Bitkilerde artan sayıda CaM-bağlayıcı protein rapor edilmiştir.

Ayrıca, sinyaline yanıt veren proteinler veya etilenin neden olduğu CaM-bağlayıcı proteinler olarak da adlandırılan Calmodulin binding transkripsiyon aktivatörleri, ilk olarak, spesifik olmayan DNA

bağlama aktivitesi olan CaM-bağlayıcı proteinler için bir taramada bitkilerde keşfedilmiştir (Bouché et al., 2002; Yang and Poovaiah, 2002; Reddy et al., 2011, Büyük et al., 2019).

Çizelge 1. Sorgum genomunda tespit edilen CAMTA proteinleri hakkında bilgi tablosu

Table 1. The table of the information on CAMTA proteins identified in sorghum genome

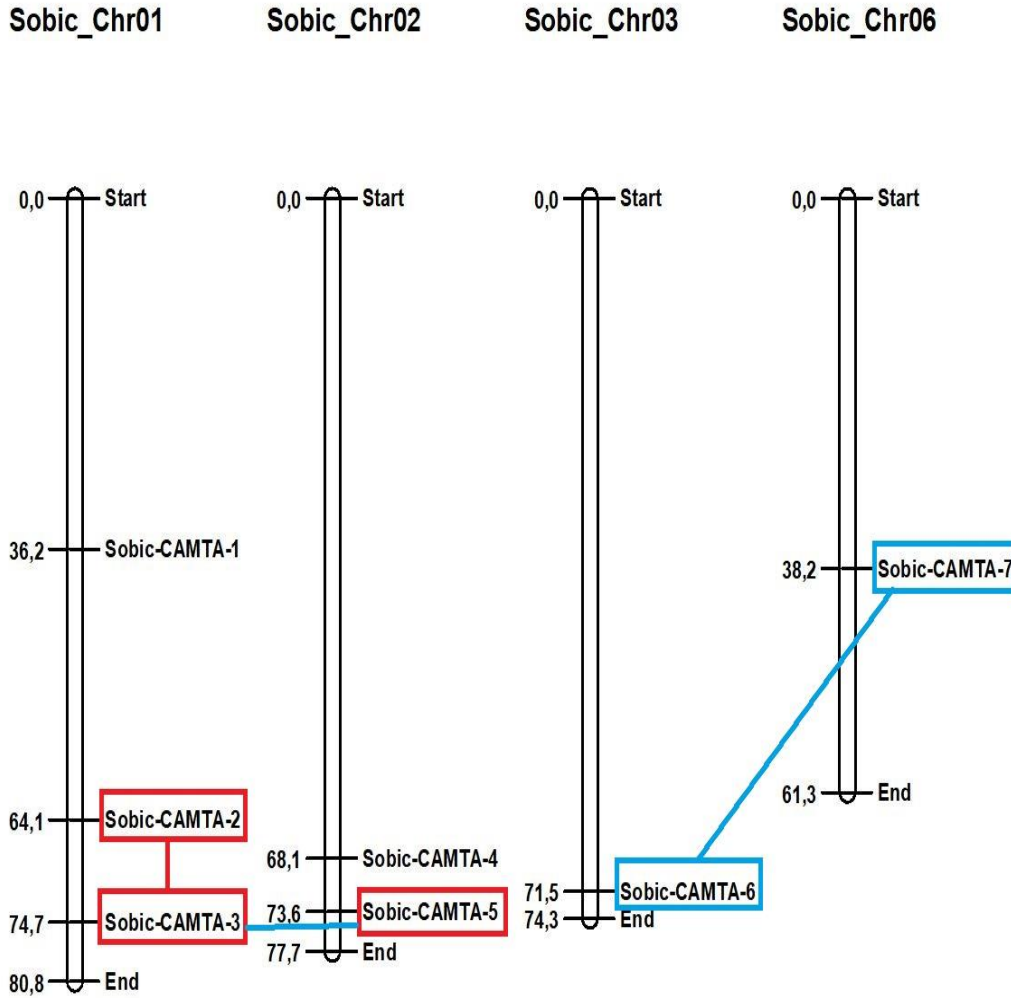
Gen No	Pytozome adı	Kromozom lokasyonu	aa	Moleküler ağırlığı (kDa)	pI	CDS uzunluğu
Sobic-CAMTA-1	Sobic.001G261501	Chr01:36154209-36162915	1022	113,53	5,63	3069
Sobic-CAMTA-2	Sobic.001G351700	Chr01:64108477-64117857 reverse strand	1024	114,4	5,86	3075
Sobic-CAMTA-3	Sobic.001G474600	Chr01:74671806-74678991	994	110,36	5,71	2985
Sobic-CAMTA-4	Sobic.002G306000	Chr02:68083385-68092596 reverse strand	914	102,44	8,36	2745
Sobic-CAMTA-5	Sobic.002G380200	Chr02:73626556-73633704 reverse strand	1030	114,86	5,55	3093
Sobic-CAMTA-6	Sobic.003G407200	Chr03:71482355-71489364	845	95,22	6,97	2538
Sobic-CAMTA-7	Sobic.006G051800	Chr06:38176208-38184325	995	111,80	5,95	2985

pI: Teorik İzoelektrik Noktası

Ek olarak, Sobic-CAMTA genlerinin teorik izoelektrik noktaları, moleküler ağırlıkları, aminoasit ve CDS dizi uzunlukları da verilmiştir (Çizelge 1). Sobic-CAMTA genleri sorgum bitkisinin 1, 2, 3, ve 6 numaralı kromozomlarında dağılmıştır. En fazla gen, kromozom 1 üzerinde Sobic-CAMTA-1 Sobic-CAMTA-2 ve Sobic-CAMTA-3 genleri olarak bulunmuştur (Şekil 1). Sobic-CAMTA proteinleri büyüklüklerinin 845–1030 amino asit arasında değişiklik gösterdiği tespit edilmiştir. Bunlar arasında en uzun aminoasit dizisine sahip olan protein, 1030 aminoasit ile Sobic-CAMTA-5 olarak belirlenmiştir. En kısa protein ise 845 aminoasit dizisiyle Sobic-CAMTA-6 olarak tespit edilmiştir.

Sobic-CAMTA proteinlerinin asitlikten alkaliye doğru değişen teorik izoelektrik noktası 5,55 ile 8,36 arasında değişkenlik göstermiştir. En yüksek izoelektrik noktası 8,36 ile Sobic-CAMTA-4'e ait olup, en düşük izoelektrik noktası ise 5,55 ile Sobic-CAMTA-5'e aittir. Bu proteinlerin moleküler ağırlıkları ise 95,22 kDa (Sobic-CAMTA-6) ile 114,86 kDa (Sobic-CAMTA-5) arasında

değişmektedir. Gen duplikasyon analizi sonucunda Sobic-CAMTA-2/Sobic-CAMTA-3 tandem duplike genler olarak belirlenmiş iken, Sobic-CAMTA-3/Sobic-CAMTA-5 ve Sobic-CAMTA-6/Sobic-CAMTA-7 ise segmentel duplike genler olarak tespit edilmiştir. Sobic-CAMTA-2 ile Sobic-CAMTA-3 genlerinin Ka değeri 0,3792, Ks değeri 2,5660 ve Ka/Ks değeri ise 0,1478, Sobic-CAMTA-3 ile Sobic-CAMTA-5 genlerinin Ka değeri 0,3423, Ks değeri 3,3026 ve Ka/Ks değeri ise 0,1037 olarak tespit edilmiştir. Sobic-CAMTA-6 ile Sobic-CAMTA-7 genlerinin Ka değeri 0,4611, Ks değeri 3,7670 ve Ka/Ks değeri ise 0,1224 olarak belirlenmiştir. Ka/Ks oranı 1'den büyükse gen dizisinin evrimi süresince pozitif seleksiyon olduğu, 1'den küçük olması arındırıcı seleksiyon ve 1'e eşit olması duplikasyon olaylarında doğal seleksiyonu göstermektedir (Juretic et al., 2005). Çalışmadan elde edilen bulgulara göre ardışık ve bölgesel genler arasındaki Ka/Ks değerlerinin 1'den küçük çıkmasından dolayı sorgum bitkisinin evrimsel sürecinde yüksek oranda arındırıcı seleksiyon olduğu anlaşılmaktadır.



Şekil 1. Sobic-CAMTA genlerinin *Sorghum bicolor* L. kromozomundaki dağılışı (Genler arasındaki kırmızı çizgi tandem duplikasyonu, mavi çizgi ise segmental duplike genleri göstermektedir)

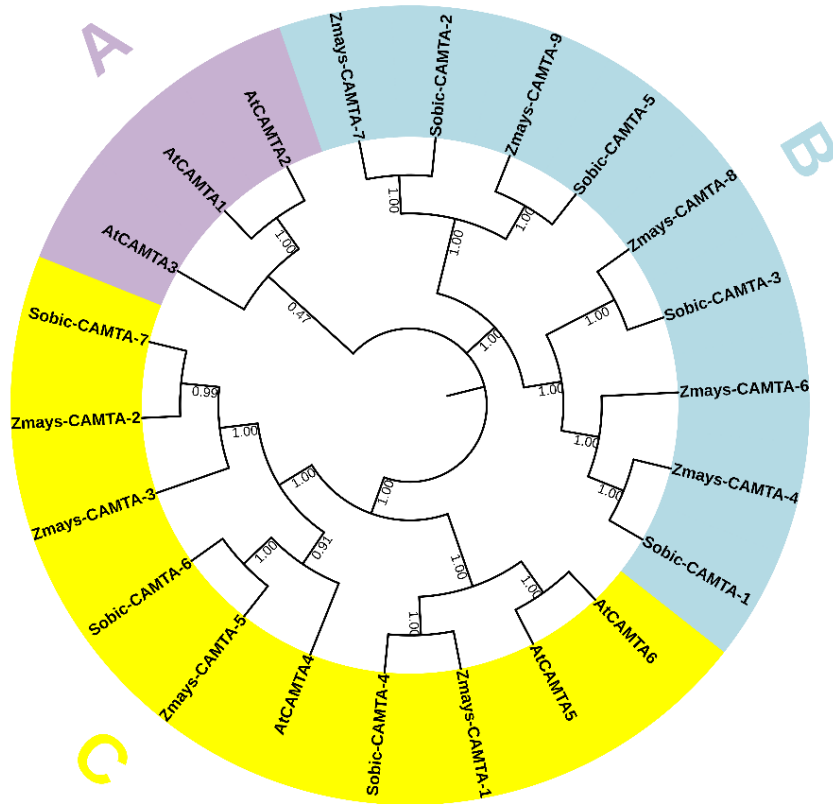
Figure 1. The distribution in Sobic-CAMTA genes in *Sorghum bicolor* L. chromosomes (red line among genes shows tandem duplication, blue line shows segmental duplicate genes)

Sobic-CAMTA'ların filogenetik analizleri, korunmuş motifleri ve gen yapısı

Sobic-CAMTA proteinleri arasındaki ilişkiyi belirlemek için *A. thaliana* ve *Z. mays* bitkilerinin CAMTA genleri kullanılarak filogenetik ağaç oluşturulmuştur. Filogenetik ağaç CAMTA proteinlerinin aminoasit dizisine bağlı olarak 1000 tekrarlı bootstrap değeri ile komşu katılım metodu kullanılarak çizilmiştir (Şekil 2).

CAMTA proteinleri A, B ve C olmak üzere 3 ana grupta kümelenebilir. A grubunda AtCAMTA1, 2 ve 3 proteinleri, B grubunda Sobic-CAMTA-1, Zmays-CAMTA-4 (GRMZM2G153594), Zmays-

CAMTA-6 (GRMZM2G341747), Sobic-CAMTA-3, Zmays-CAMTA-8 (GRMZM2G447551), Sobic-CAMTA-5, Zmays-CAMTA-9 (GRMZM2G017368), Sobic-CAMTA-2 ve Zmays-CAMTA-7 (GRMZM2G431243) proteinleri ve C grubunda ise AtCAMTA6, AtCAMTA5, Zmays-CAMTA-1 (GRMZM2G032336), Sobic-CAMTA-4, AtCAMTA4, Zmays-CAMTA-5 (GRMZM2G171600), Sobic-CAMTA-6, Zmays-CAMTA-3 (GRMZM2G152661), Zmays-CAMTA-2 (GRMZM2G143205) ve Sobic-CAMTA-7 proteinleriyle filogenetik bir ilişki göstermiştir.

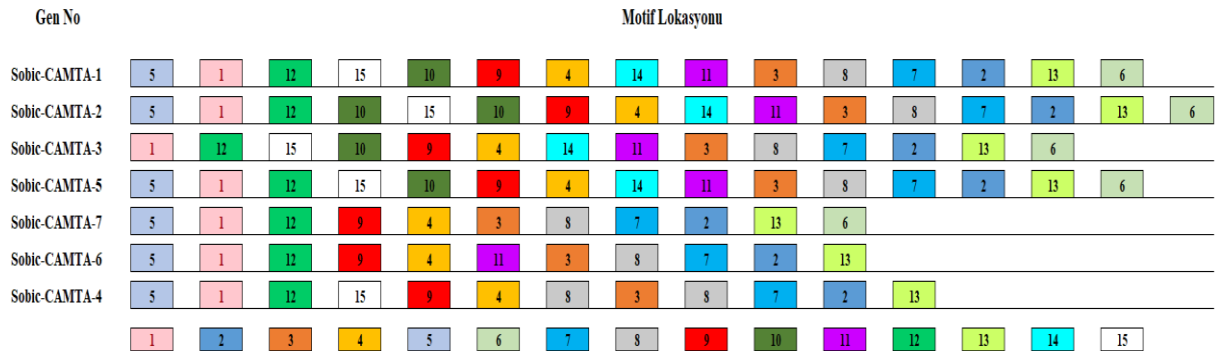


Şekil 2. *Sorghum bicolor*, *Arabidosis thaliana* ve *Zea mays* bitkilerinin CAMTA proteinleri ile çizilen filogenetik ağaç

Figure 2. Phylogenetic tree of CAMTA proteins of *Sorghum bicolor*, *Arabidosis thaliana* and *Zea mays*

MEME (v4.12.1) (Bailey et al., 2006) programı kullanılarak Sobic-CAMTA proteinlerinde yapılan korunmuş motif analizlerinde 15 korunmuş motif tespit edilmiştir. Belirlenen motiflerin uzunluğu 17-50 aminoasit aralığında olmuştur (Çizelge 2). En fazla motif Sobic-CAMTA-2’de (16), en az motif ise Sobic-CAMTA-7 ve Sobic-CAMTA-6’da (11) tespit edilmiştir. Tüm Sobic-CAMTA proteinlerinin Motif

1, Motif 2, Motif 3, Motif 4, Motif 7, Motif 8, Motif 9, Motif 12, Motif 13 içerdiği görülmüştür (Şekil 3). Ayrıca yapılan Interproscan analizi sonucunda Motif 1, Motif 5 ve Motif 12’nin CG-1 DNA-bağlanma domaini, Motif 3’ün Ankrin tekrarı-içeren domain süper ailesi, Motif 2 IQ motifi ve EF-hand bağlanma bölgesi ve Motif 4 Immunoglobulin-like fold domainlerini içerdikleri tespit edilmiştir.



Şekil 3. Sobic-CAMTA genlerinde tahmin edilen motif dağılımı

Figure 3. Distribution of estimated motif in Sobic-CAMTA genes

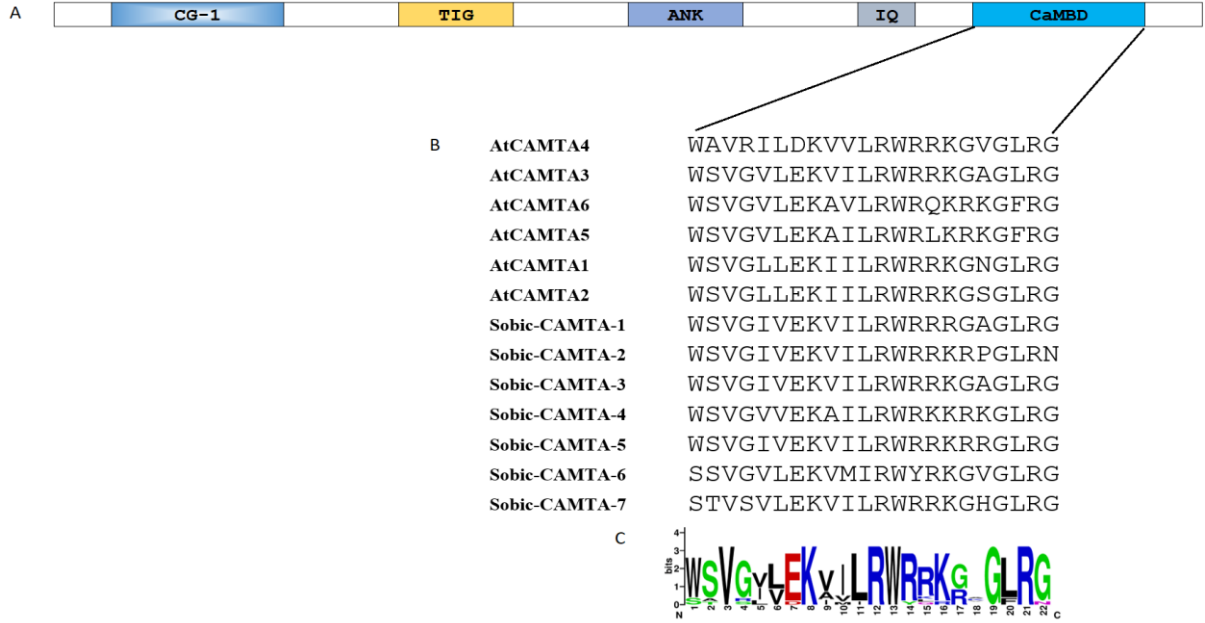
Çizelge 2. Sobic-CAMTA genlerinde muhtemel motiflerin dizi bilgileri
Table 2. Sequence information of the best possible motif in Sobic-CAMTA genes

Motif No	Motif Genişliği	Muhtemel En İyi Eşleşme	Domain
MOTIF-1	50	DGHNWRKKKDGKTVKEAHERLKVGNVDVL HCYYAHGEEENPNFQRRTYWML	CG-1 DNA-binding domain
MOTIF-2	50	FRGWKGRKEFMIRQIRVQIAHVRGHQVRK HYRKIVWSVGIVEKVILRW	IQ motif, EF-hand binding site
MOTIF-3	50	IHLCAALGYDWAIRPILAAGVNIINFRDAHGW ALHWAA YCGRERMVVALI	Ankyrin repeat-containing domain superfamily
MOTIF-4	50	WSCMFGDVEVP AEILQDGTLCRYTPCHRPGR VPFYITCSNRVACSEVREF	IPT domain
MOTIF-5	50	EAQTRWLRPTEICEILQNYKRFSIAPEPPNKPPS GSLFLFDRKVLRYFRK	CG-1 DNA-binding domain
MOTIF-6	50	GRKQTEERLQKALARVHSMAYPEARDQYQ RILTCVTKMQESQAMQDKML	N/A
MOTIF-7	37	KDSLGA VRNAAQAAARIYQAFRVFSFHRKKV VQYEDD	N/A
MOTIF-8	42	THEDPTGKTPADLASANGHKGIAFLAESALT SHLQALTLKE	N/A
MOTIF-9	39	PMLSQDQLFSIHDISPEWAYCGEKT KVLITGTF LCNPEH	N/A
MOTIF-10	50	YPLLKTQSSDLFTMEKDGFKKHDSFTRWMSK ELPEVDDLQIKSSSGAFWS	N/A
MOTIF-11	38	RDQFIEKLIKEKLHIWLIHKVHDGGKGP NVLD DEGQGV	N/A
MOTIF-12	19	EEDYMHIVLVHYLEVKGGK	CG-1 DNA-binding domain
MOTIF-13	17	RRKGHGLRGFRPEMQPM	N/A
MOTIF-14	31	NKMHLHIRLDKLLSLGPDDYQMTVINPGNEM	N/A
MOTIF-15	19	GYHPFTWMQQHVDGIGPVN	N/A

CAMTA gen ailesi ile daha önceki yapılan çalışmalarda CaMBD korunmuş bölgesini içerdikleri bildirilmiştir (Pant et al., 2018). WEBLOGO ara yüzü kullanılarak *A. thaliana* ve *S. bicolor* bitkilerinin CAMTA gen ailelerine ait proteinlerin karboksil ucuna doğru korunmuş olan CaMBD bölgesi tespit edilmiştir (Şekil 4). Önceki çalışmalarla benzer bulgulara ulaşılmıştır.

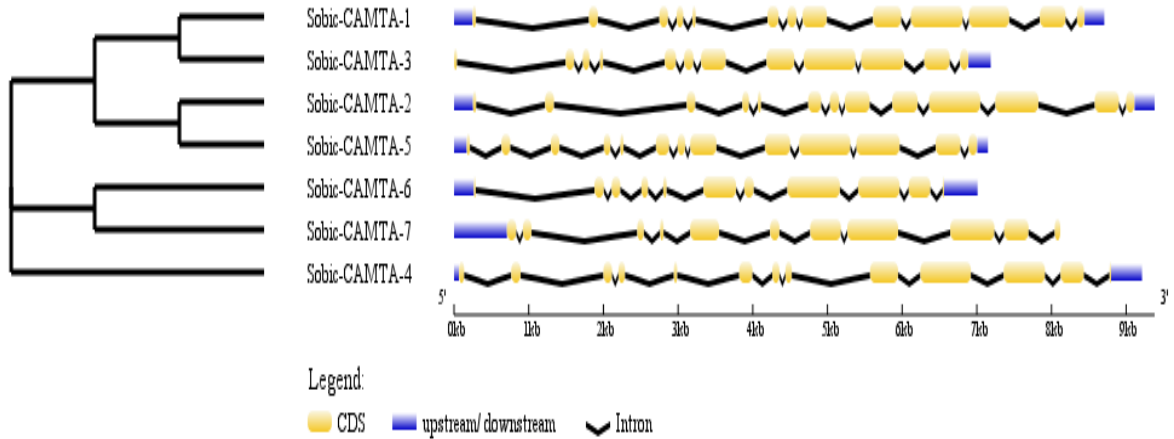
Gene Structure Display Server v2.0 ile Sobic-CAMTA genlerinde yapılan yapısal analiz sonucunda ekzon ve intron sayıları tespit edilmiştir. Bu analiz

sonucunda Sobic-CAMTA genleri arasında ekzon sayıları 10 ile 13 arasında değişmektedir. İtron sayıları da 10-12 arasında belirlenmiştir. En fazla sayıda ekzon içeren genler (13 ekzon) Sobic-CAMTA-1, -2 ve -5'dir. En az sayıda ekzon içeren gen (10 ekzon) Sobic-CAMTA-6'dır. En yüksek sayıda intron içeren genler (12 intron) Sobic-CAMTA-1,-2,-5 ve -4 olarak belirlenmiştir. En az sayıda intron içeren genler (10 intron) Sobic-CAMTA-6 ve Sobic-CAMTA-7 olarak tespit edilmiştir (Şekil 5).



Şekil 4. Tüm sorgum CAMTA proteinlerindeki korunmuş CaMBD bölgesi. (A) *Arabidopsis* CAMTA genlerindeki fonksiyonel olarak tespit edilmiş motif. (B) Altı *AtCAMTA* ve sorgum CAMTA proteinlerindeki korunmuş CaMBD'nin hizalanması. (C) Altı *AtCAMTA* ve 7 sorgum CAMTA proteinindeki CaMBD'nin dizi logosu.

Figure 4. The conserved CaMBD region in all sorghum CAMTA proteins. (A) Functionally identified motif in *Arabidopsis* CAMTA genes. (B) Alignment of the conserved CaMBD in six *AtCAMTA* and sorghum CAMTA proteins. (C) Sequence logo of CaMBD in six *AtCAMTA* and 7 sorghum CAMTA proteins



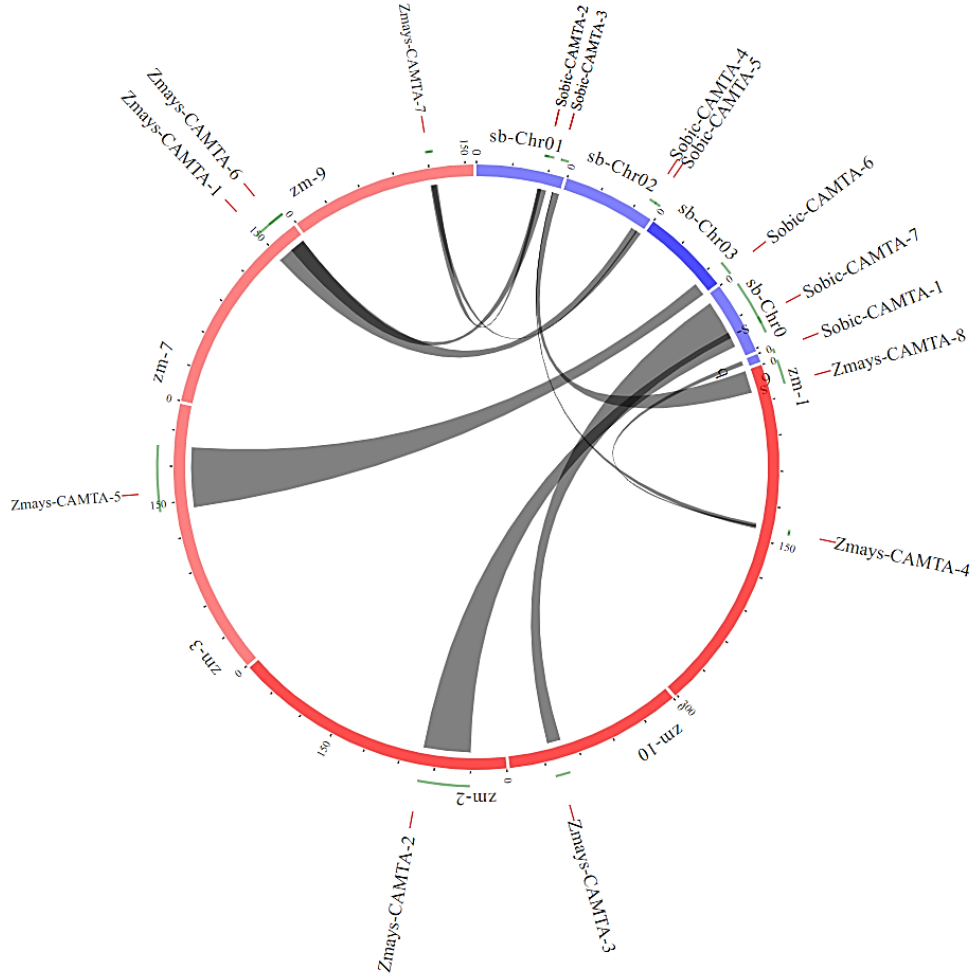
Şekil 5. Sobic-CAMTA genlerindeki ekzon ve intronların sayısı, uzunluğu ve pozisyonu

Figure 5. Number, length and position of exons and introns in Sobic-CAMTA genes

S. bicolor, *A. thaliana* ve *Z. mays* bitkilerindeki CAMTA genlerinin karşılaştırmalı ve sinteni analizler

S. bicolor, *A. thaliana* ve *Z. mays* bitkilerine ait CAMTA proteinleri arasında sinteni haritası çizilmiştir (Şekil 6). Sobic-CAMTA-2, -3, -4, -5, -6, -

7 genleri ile *Zmays*-CAMTA-1, *Zmays*-CAMTA-2, *Zmays*-CAMTA-3, *Zmays*-CAMTA-4, *Zmays*-CAMTA-5, *Zmays*-CAMTA-7, *Zmays*-CAMTA-8 ve *Zmays*-CAMTA-9 genleri arasında sinteni ilişkisi tespit edilmiştir. *S. bicolor* ile *A. thaliana* arasında sinteni ilişkisi bulunamamıştır.



Şekil 6. *Sorghum bicolor* ve *Zea mays* arasındaki genlerinin geniş çaplı sinteni analizi
Figure 6. Large-scale synteny analysis of genes between *Sorghum bicolor* and *Zea mays*

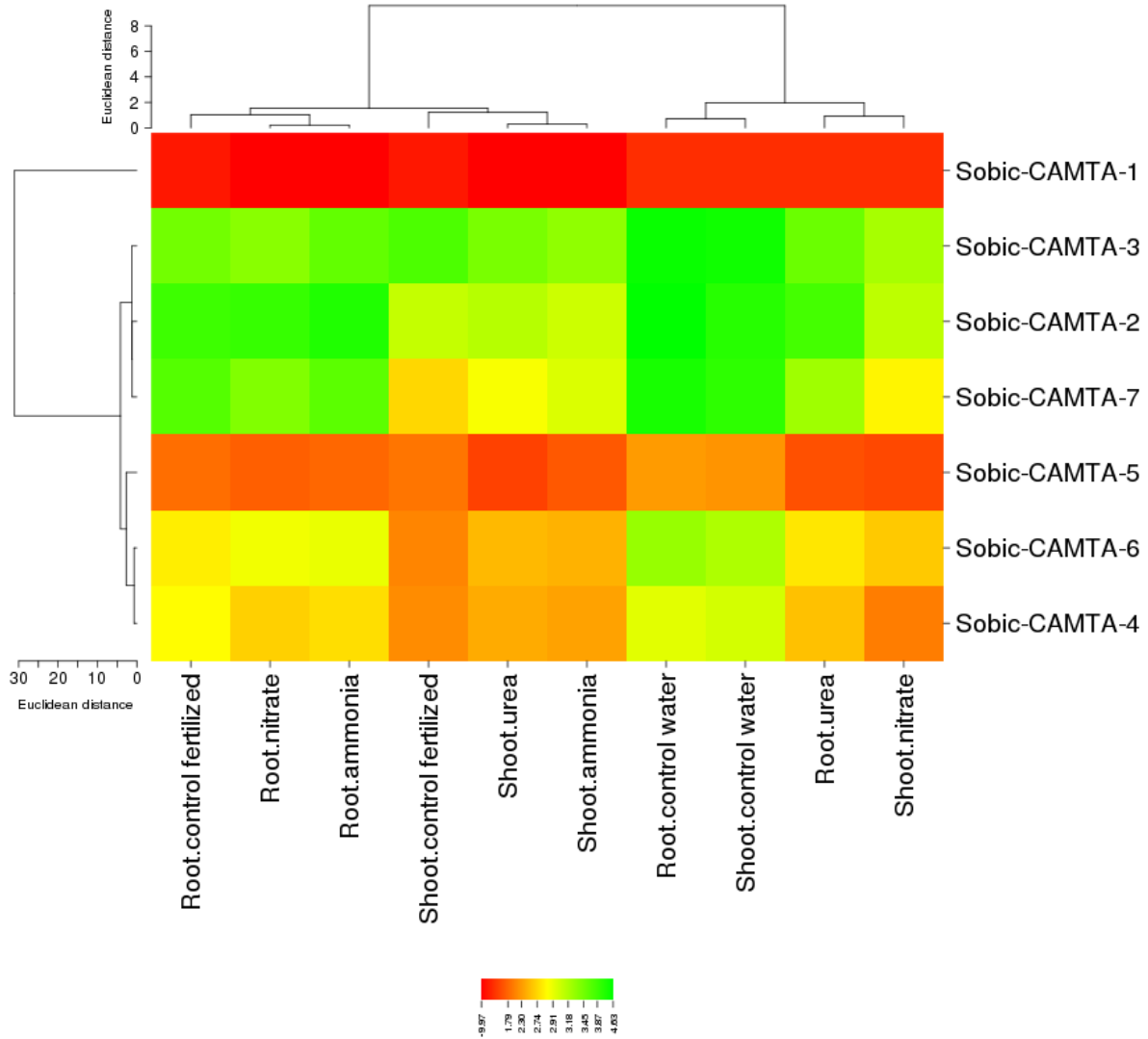
Sobic-CAMTA genlerinin genom çaplı ifadeleri

Sobic-CAMTA genlerinin *S. bicolor* bitkisinin kök ve sürgün dokularında farklı azot kaynağı uygulaması ile CAMTA transkripsiyon faktörünün ifade profili belirlenmiştir (Şekil 7). İfade düzeyleri incelendiğinde en fazla ifade olan genler sırasıyla Sobic-CAMTA-3, Sobic-CAMTA-2 ve Sobic-CAMTA-7 iken ifade seviyesi en az olan gen Sobic-CAMTA-5 olarak tespit edilmiştir. Sobic-CAMTA-1'in ise hiç ifade olmadığı görülmüştür. Kök ve sürgün dokularında su ile yapılan sulamada Sobic-CAMTA-2, -3 ve -7 genlerinin ifade seviyelerinin arttığı gözlemlenmiştir. Bunun yanında sürgün dokularında bu genlerde kullanılan azot kaynağına (üre, amonyak ve nitrat), su kontrolüne ve gübre kontrolüne göre ifade seviyeleri azalmıştır. Diğer taraftan kök dokularında ise bu genlerin ifade seviyeleri sürgüne göre artmıştır. Sobic-CAMTA-5'in ifade seviyesi ise kök ve sürgün dokularında özellikle üre ve nitratla yapılan uygulamalarda

azalma göstermiştir. Sobic-CAMTA-4 ve -6 genlerinde kök dokularında üre, amonyak ve nitratla yapılan uygulamalarda ifade seviyeleri artarken sürgün dokularında artış fazla olmamıştır. Yine her iki genin kök ve sürgün dokularında su kontrolü ile yapılan uygulamalarda ifade seviyelerinde artışlar gözlenmiştir. Hu et al. (2015) Çin lahanasında yaptıkları çalışmalarında nitrat uygulamasında BcCAMTA1 ve BcCAMTA5 dışındaki BcCAMTA genlerinin ifade seviyelerinin 12 saate kadar kademeli olarak azaldığı 24 saatte ise yeniden arttığı bildirilmiştir. Ayrıca aynı çalışmada genlerin 12. saatteki ifade seviyelerinin 4. saatteki ifade seviyelerinden daha yüksek olduğu belirtilmiştir. Aynı çalışmada soğuk, kuraklık, bakır iyonu ve nitrat stresi gibi çevresel streslere maruz bırakılan çin lahanası fidelerinde BcCAMTA genlerinin ifade seviyeleri araştırılmıştır. BcCAMTA2 dışındaki bütün BcCAMTA genlerinin soğuk stresinde ifade seviyelerinin arttığı tespit edilmiştir. Yapılan diğer

çalışmalarda ise pamuk bitkisinde lif gelişimiyle bazı CAMTA genlerinin ifade seviyelerinin pozitif korelasyon gösterdiği rapor edilmiştir (Pant et al., 2018). Biyotik ve abiyotik streslere maruz bırakılan kavak bitkisinde CAMTA gen ailesi üyelerinin ifade

seviyelerinde artışlar gözlenmiştir (Wei et al., 2017). Fasulye bitkisinde tuz stresi koşulları altında yapılan çalışmada CAMTA gen ailesi üyelerinin dokuya bağlı olarak ifade seviyelerinde artış ve azalışlar gösterdiği bildirilmiştir (Büyük et al., 2019)



Şekil 7. Sorgum bitkisinin farklı dokularında yapılan ifade analizi (Heatmap)
Figure 7. Expression analysis in different tissues of sorghum plant (Heatmap)

SONUÇ

Sorghum bitkisinde bulunan CAMTA gen ailesi ile in siliko olarak yapılan bu çalışmada 7 tane CAMTA üyesi belirlenmiştir. Bunlar 4 farklı kromozoma dağılmıştır. Sorgum bitkisinin in siliko olarak azotun farklı dokulara uygulanmasıyla gen ifadesi profilinde azot uygulanması gen ifade düzeyini düşürdüğü tespit edilmiştir. Bu çalışmanın sonuçları *Sorghum bicolor* bitkisinde CAMTA transkripsiyon gen ailesinin moleküler yapısının anlaşılması için önemli bilgiler sağlayacaktır.

Çıkar Çatışması Beyanı

Yazarlar, çalışma dahilinde herhangi bir kişisel ve/veya finansal çıkar çatışmasının olmadığını beyan eder.

Yazar Katkıları

Eİ, DK ve AGK araştırmayı tasarladı ve analizlerini yaptı, AH, KH ve SM çalışmanın düzenlenmesi, resimlerin, tabloların hazırlanması çalışmalarını yürüttüler. Tüm yazarlar makalenin

yazımına katkı yaptı ve makalenin yayın aşamasındaki süreçte görev alarak okuyup onayladılar.

KAYNAKLAR

- Bailey, T.L., Williams, N., Misleh, C., Li, W.W., 2006. MEME: Discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.*, 34: W369-W373.
- Bouché, N., Scharlat, A., Snedden, W., Bouchez, D., Fromm, H., 2002. A novel family of calmodulin-binding transcription activators in multicellular organisms. *J. Biol. Chem.*, 277 (24): 21851-21861.
- Büyük, İ., İlhan, E., Şener, D., Özsoy, A. U., Aras, S., 2019. Genome-wide identification of CAMTA gene family members in *Phaseolus vulgaris* L. and their expression profiling during salt stress. *Mol. Biol. Rep.*, 46 (3): 2721-2732.
- Choi, M.S., Kim, M.C., Yoo, J.H., Moon, B.C., Koo, S.C., Park, B.O., Lee, J.H., Koo, Y.D., Han, H.J., Lee, S.Y., Chung, W.S., Lim, C.O., Cho, M.J., 2005. Isolation of a calmodulin-binding transcription factor from rice (*Oryza sativa* L.). *J. Biol. Chem.*, 280: 40820-40831.
- CIMMiner, 2020. CIMMiner. <http://discover.nci.nih.gov/cimminer> (Erişim Tarihi: 20 Ocak 2020).
- Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: A sequence logo generator. *Genome Res.*, 14 (6): 1188-1190.
- de Mendoza, A., Sebé-Pedrós, A., Šestak, M. S., Matejčić, M., Torruella, G., Domazet-Lošo, T., Ruiz-Trillo, I., 2013. Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. *PNAS*, 110 (50): E4858-E4866.
- Du, L., Ali, G.S., Simons, K.A., Hou, J., Yang, T., Reddy, A.S.N., Poovaiah, B.W., 2009. Ca²⁺/calmodulin regulates salicylic-acid-mediated plant immunity. *Nature*, 457: 1154-1158.
- Du, L., Poovaiah, B. W., 2004. A novel family of Ca²⁺/calmodulin-binding proteins involved in transcriptional regulation: Interaction with fsh/Ring3 class transcription activators. *Plant. Mol. Biol.*, 54 (4): 549-569.
- ExPASy, 2020a. Decrease Redundancy Tool. http://web.expasy.org/decrease_redundancy/ (Erişim Tarihi: 20 Ocak 2020).
- ExPASy, 2020b. ProtParam. <http://web.expasy.org/protparam> (Erişim Tarihi: 20 Ocak 2020).
- Finkler, A., Kaplan, B., Fromm, H., 2007. Ca²⁺-responsive cis-elements in plants. *Plant Signal Behav.*, 2 (1): 17-19.
- Guo, A., Zhu, Q., Chen, X., Luo, J., 2007. GSDS: a gene structure display server. *Yi Chuan= Hereditas*, 29 (8): 1023-1026.
- HMMER, 2020. HMMER: biosequence analysis using profile hidden Markov models. <http://hmmer.org/> (Erişim Tarihi: 20 Ocak 2020).
- Hu, R., Wang, Z., Wu, P., Tang, J., Hou, X., 2015. Identification and abiotic stress analysis of calmodulin-binding transcription activator/signal responsive genes in non-heading Chinese cabbage (*Brassica campestris* ssp. *chinensis* Makino). *Plant Omics J.*, 8 (2): 141-147.
- Juretic, N., Hoen, D.R., Huynh, M.L., Harrison, P.M., Bureau, T.E., 2005. The evolutionary fate of MULE-mediated duplications of host gene fragments in rice. *Genome Res.*, 15 (9): 1292-1297.
- Kim, Y., Park, S., Gilmour, S.J., Thomashow, M.F., 2013. Roles of CAMTA transcription factors and salicylic acid in configuring the low temperature transcriptome and freezing tolerance of *Arabidopsis*. *Plant J.*, 75 (3): 364-376.
- Lee, T.H., Tang, H.B., Wang, X.Y., Paterson, A.H., 2013. PGDD: A database of gene and genome duplication in plants. *Nucleic Acids Res.*, 41 (D1): 1152-1158.
- Letunic, I., Bork, P., 2011. Interactive tree of life v2: Online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res.*, 39: W475-W478.
- Ludwig, A.A., Romeis, T., Jones, J., 2004. CDPK-mediated signalling pathways: specificity and cross-talk. *J. Exp. Bot.*, 55 (395): 181-188.
- McCormick, R.F., Truong, S.K., Sreedasyam, A., Jenkins, J., Shu, S., Sims, D., Kennedy, M., Amirebrahimi, M., Weers, B.D., McKinley, B., Mattison, A., Morishige, D.T., Grimwood, J., Schmutz, J., Mullet, J.E., 2018. The *Sorghum bicolor* reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *Plant J*, 93 (2): 338-354.
- Menz, M.A., Klein, R.R., Mullet, J.E., Obert, J.A., Unruh, N.C., Klein, P.E., 2002. A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP[®], RFLP and SSR markers. *Plant Mol. Biol.*, 48 (5-6): 483-499.
- Miller, A.J., Cramer, M.D., 2005. Root nitrogen acquisition and assimilation. *Plant Soil*, 274: 1-36.
- Pant, P., Iqbal, Z., Pandey, B. K., Sawant, S. V., 2018. Genome-wide comparative and evolutionary analysis of calmodulin-binding

- transcription activator (CAMTA) family in *Gossypium* species. *Sci. Rep.*, 8 (1): 5573.
- PGDD, 2020. Plant Genome Duplication Database. <http://chibba.agtec.uga.edu/duplication/index/locus> (Erişim Tarihi: 20 Ocak 2020).
- Phytozome, 2020. Phytozome Database. <https://phytozome.jgi.doe.gov/pz/portal.html> (Erişim Tarihi: 20 Ocak 2020).
- Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., Lopez, R., 2005. InterProScan: Protein domains identifier. *Nucleic Acids Res.*, 33 (suppl_2): W116-W120.
- Rahman, H., Yang, J., Xu, Y. P., Munyampundu, J. P., Cai, X. Z., 2016. Phylogeny of plant CAMTAs and role of AtCAMTAs in nonhost resistance to *Xanthomonas oryzae* pv. *oryzae*. *Front. Plant Sci.*, 7: 177.
- Reddy, A.S., Ali, G.S., Celesnik, H., Day, I.S., 2011. Coping with stresses: Roles of calcium- and calcium / calmodulin-regulated gene expression. *Plant Cell*, 23 (6): 2010-2032.
- Shangguan, L., Wang, X., Leng, X., Liu, D., Ren, G., Tao, R., Zhang, C., Fang, J., 2014. Identification and bioinformatic analysis of signal responsive/calmodulin-binding transcription activators gene models in *Vitis vinifera*. *Mol. Biol. Rep.*, 41: 1573-4978.
- Suyama, M., Torrents, D., Bork, P., 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.*, 34: W609-W612.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony methods. *Mol. Biol. Evol.*, 28 (10): 2731-2739.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, 25 (24): 4876-4882.
- Trapnell, C., Roberts, A., Goffl, O., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., Pachter, L., 2013. Differential gene and transcript expression analysis of RNAseq experiments with TopHat and Cufflinks. *Nat. Protoc.*, 7 (3): 562-578.
- Voorrips, R.E., 2002. MapChart: Software for the graphical presentation of linkage maps and QTLs. *J. Hered.*, 93 (1): 77-78.
- Wei, M., Xu, X., Li, C., 2017. Identification and expression of CAMTA genes in *Populus trichocarpa* under biotic and abiotic stress. *Sci. Rep.*, 7 (1): 17910.
- Xu, H., Ding, A., Chen, S., Marowa, P., Wang, D., Chen, M., Hu, R., Kong, Y., O'Neill, M., Chai, G., Zhou, G., 2018. Genome-Wide Analysis of Sorghum GT47 Family Reveals Functional Divergences of MUR3-Like Genes. *Front. Plant Sci.*, 9: 1773.
- Yang, T., Poovaiah, B.W., 2000. An early ethylene up regulated gene encoding a calmodulin binding protein involved in plant senescence and death. *J. Biol. Chem.*, 275 (49): 38467-38473.
- Yang, T., Poovaiah, B.W., 2002. A calmodulin-binding/CGCG box DNA-binding protein family involved in multiple signalling pathways in plants. *J. Biol. Chem.*, 277 (47): 45049-45058.
- Yang, Z.H., 2007. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.*, 24 (8): 1586-1591.
- Zheng, Y., Jiao, C., Sun, H.H., Rosli, H.G., Pombo, M.A., Zhang, P.F., Banf, M., Dai, X.B., Martin, G.B., Giovannoni, J.J., Zhao, P.X., Rhee, S.Y., Fei, Z.J., 2016. iTAK: A program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol. Plant.*, 9 (12): 1667-1670.