

**A búzaszem fejlődésében szerepet játszó gének
azonosítása bioinformatikai és molekuláris
módszerekkel**

Doktori (Ph.D.) értekezés

SZÜCS ATTILA

**MTA Szegedi Biológiai Központ,
Növénybiológiai Intézet**

Témavezető: Dr. Fehér Attila

Szeged

2007

A növényi petesejt transzkripció profilja

A növények életciklusa egy diploid sporofitikus és egy haploid gametofitikus nemzedék váltakozásaként zajlik. Jelenlegi ismereteink nagy része a domináns sporofitikus életszakaszra vonatkozik, annál is inkább, mert a növényi gametofiton csak néhány sejtből áll és azok is a sporofita szövetekbe mélyen beágyazva helyezkednek el. Napjaink modern kutatási eszközei azonban lehetővé teszik, hogy a növényeknek azt a rejtett életszakaszát is alaposabban megvizsgáljuk. T-DNS inszerciós mutagenézis segítségével lúdfűben több, mint száz olyan mutációt azonosítottak, amelyek a gametofiton fejlődését érintik (Pagnussat és mtsai. 2005). DNS mikroarray (DNS-chip) segítségével feltérképezték a gametofiton génkifejeződési mintázatát és 225 gametofiton specifikus gént azonosítottak (Yu és mtsai. 2005).

A petesejt speciális szerepet tölt be a gametofitonon belül, hiszen a megtermékenyítést követően belőle fejlődik az új sporofiton nemzedék. Nem tudjuk azonban, hogy ebben a képességében mennyi az anyai hatás és mennyi a megtermékenyítés szerepe. Arról is kevés információ áll rendelkezésünkre, hogy a növényi petesejt mennyire aktív és milyen folyamatok zajlanak benne. Mindezekről átfogó képet nyújthat a petesejt génkifejeződési mintázatának vizsgálata és annak összehasonlítása más növényi sejtekével.

A lúdfű petesejt vizsgálata ilyen szempontból nem lehetséges, kis mérete és izolálásának nehézsége miatt. A kukorica és a búza esetében azonban már évek óta rendelkezésünkre állnak azok a technikák, amelyek lehetővé teszik a petesejtek rutinszerű izolálását. Ezeknek a technikáknak a kidolgozásában Barnabás Beáta kutatócsoportjának úttörő szerepe volt. Erre a tapasztalatra és technikai lehetőségekre alapozva, közös kutatások keretében, elkészítettünk egy búza petesejt cDNS könyvtárat és megvizsgáltuk, hogy milyen gének fejeződnek ki ebben a speciális sejttypusban. Mint kiderült, ez a megközelítés nem csak bennünk merült fel, és Sprunck és munkatársai 2005-ben leközölték 404 búza petesejtben expresszáldó gén azonosítását és elsődleges jellemzését. Ezek a cDNS-ek a szekvencia adatbázisokban elérhetők, így összehasonlíthatóak voltak a laboratóriumunkban azonosított 251 búza petesejt cDNS szekvenciával. Az azonos szekvenciák száma nagyon kevés volt, így 237 petesejtben

expresszállódó új gént azonosítottunk. Ezek közül 73 még teljesen ismeretlen búza szekvencia, míg 188 korábban már más cDNS könyvtárakból előkerült. Mindkét cDNS szekvencia készlet arra enged következtetni, hogy a petesejtben jelentős anyagcsere aktivitás és intenzív fehérjeszintű változások zajlanak, azaz a növényi petesejt koránt sem annyira nyugvó, mint azt korábban feltételezték (Russell 1993).

Kumlehn és munkatársai (Kumlehn és mtsai. 2001) szintén beszámoltak már búza petesejtből, mégpedig a parthenogenetikus fejlődésre képes „Salmon” genotípus petesejtjeiből, készített cDNS bankról. A mintegy ötszáz azonosított cDNS szekvencia részletet azonban nem tették nyilvánossá. Azóta még több petesejt cDNS (EST) szekvenciát azonosítottak, de ezek sem kerültek még be a nyilvános szekvencia adatbázisokba. Elmondhatjuk azonban, hogy a három laboratórium összességében mintegy ezer-ezerötszáz különböző búza petesejtben expresszállódó gént azonosított, ezeknek az átfogó összehasonlító vizsgálata azonban még várat magára.

A fent említett búza petesejt cDNS könyvtárak kapcsán még meg kell említenünk, hogy egyedül az általunk készített könyvtár készült phagemid vektorban és olyan módon, hogy a petesejtben expresszállódó gének legalább felét ($1,5 \times 10^6$ gént) reprezentálja. Az amplifikált cDNS könyvtár szinte korlátlan mennyiségben áll rendelkezésre és lehetővé teszi hagyományos molekuláris szűrési technikák alkalmazását. A másik két cDNS könyvtár lényegesen kisebb számú, néhány ezer, plazmid DNS-be klónozott cDNS fragmentumot tartalmaz.

A zigóta és a kétsejtes embrió transzkripció profilja

Sprunck és munkatársai (Sprunck és mtsai. 2005) kétsejtes búza embrióból is készítettek cDNS könyvtárat és ennek felhasználásával 289 EST szekvenciát határoztak meg. A petesejt és a kétsejtes zigóta transzkript profiljának összehasonlítása azt jelezte, hogy az embrióban intenzívebb transzkripció és transzláció folyik és természetesen a sejtosztódáshoz kapcsolódó gének kifejeződése is kimutatható. Kukorica kétsejtes embriókból is azonosítottak cDNS szekvenciákat (Okamoto és mtsai. 2005), mégpedig ebben az esetben a bazális és apikális sejtből külön-külön, és megállapították, hogy a két sejt között is jelentős génkifejeződési különbség van.

Ezek a vizsgálatok azonban nem nyújtottak információt arról, hogy mikor következik be jelentős változás a génaktivításban a petesejt megtermékenyítését követően. Ennek a kérdésnek a megválaszolásában segíthet az általunk a megtermékenyítést követően hét órával begyűjtött zigótákból készített cDNS könyvtár, és az ebből származó cDNS-ek szekvenálása. Az eddig azonosított 369 zigóta EST szekvencia alapján azt mondhatjuk, hogy a megtermékenyítést követő hetedik órában már jelátvitelben résztvevő fehérjék és több transzkripciós faktor is megjelenik, ami arra utal, hogy a zigóta saját génkészlete már aktiválódott. Ennek az egyértelmű igazolása azonban még további kísérleteket igényel.

Búza, és általában növényi, zigóta cDNS könyvtárral, tudomásunk szerint csak a mi laboratóriumunk rendelkezik. Az általunk azonosított 369 zigóta szekvencia közül 81 eddig ismeretlen búza gént vagy allélt reprezentál. A 288 már ismert szekvencia közül közül 23 a petesejt, 22 pedig a kétsejtes embrió cDNS könyvtárban is előfordul. Természetesen a rendelkezésre álló szekvencia mennyiség nem elégséges átfogó összehasonlító elemzés elvégzéséhez. További szekvenálásra és egyedi gének kifejeződésének vizsgálatára van szükség ahhoz, hogy a petesejt megtermékenyítését követő transzkripciós változásokról valós képet kaphassunk.

Genomikai megközelítési lehetőségek a búza kutatásában

A búza (*Triticum aestivum* L.) az egyik legfontosabb gazdasági növény, amelynek kutatása kiemelt jelentőséggel bír. Ugyanakkor a modern genomikai kutatási módszerek alkalmazása a búza esetében nem könnyű feladat. Ezt a problémát a búza genom hatalmas mérete okozza. A termesztett búza poliploid növény, amely hatalmas méretű kromoszómákkal rendelkezik. Az allohexaploid genom összességében mintegy 16 gigabázisnyi (Gb) DNS-t tartalmaz (Bennett és Smith 1976), szemben a rizs 450 Mb-nyi genom méretével (Sasaki és Sederoff 2003). A különbségért elsősorban repetitív, nem kódoló, szekvenciák felelősek, a két faj génkészlete ugyanis nagymértékben hasonló (pl. (Sorrells és mtsai. 2003). A rizs genom becslések szerint valamivel kevesebb, mint 40000 fehérje kódoló gént tartalmaz (Bennetzen és mtsai. 2004). A búza gének számára

vonatkozó becslések elég tág intervallumba szórnak, de minden bizonnyal jóval meghaladják a rizs gének számát.

A rizs, és az árpa mint a gabonafélék kutatásának modell növényei, sok hasznos információval szolgálhatnak a búza kutatása és akár nemesítése számára is (Appels és mtsai. 2003, Feuillet és Keller 2002, Ware és mtsai. 2002). Ezeknek az információknak az alkalmazása és a búza specifikus tulajdonságainak kutatása megköveteli azonban, hogy a búza esetében is rendelkezésre álljanak az alapvető genomikai megközelítési lehetőségek. Mivel a búza genom mérete nem teszi lehetővé a teljes genom szekvencia meghatározását az elkövetkezendő néhány éven belül, ezért előtérbe került a búza cDNS szekvenciák (EST-k) minél nagyobb léptékű összegyűjtése az adatbankokban. Ehhez változatos, a búza különböző fejlődési stádiumait és eltérő környezeti hatásokra adott válaszait reprezentáló cDNS bankok előállítására van szükség. Laboratóriumunk a búza petesejt és zigóta cDNS bank elkészítésével ehhez a törekvéshez járult/járul hozzá. Az általunk meghatározott 620 EST szekvencia közül 154 az, ami az adatbázisokban fellelhető, nagyságrendileg több százezer, búza EST szekvenciával nem mutatott azonosságot (95% alatti a DNS szintű azonosság), azaz feltehetően eddig ismeretlen búza gént vagy allélt reprezentál. Ezen kívül számos esetben olyan ismert gének szekvenciáját is sikerült bővítenünk, amelyeket már reprezentált az általunk meghatározott szekvenciával átfedő adatbázis szekvencia. Azokban az esetekben pedig, ahol az általunk meghatározott szekvencia már ismert volt, új információként kerülhetett az adatbázisokba a gén kifejeződésének helyeként a petesejt és/vagy a zigóta.

Az cDNS szekvenálásból származó szekvencia információ egyik lehetséges felhasználási lehetősége a „DNS chip” vagy „DNS mikroarray” készítése és átfogó génkifejeződési változások nyomon követése. A búza esetében, mint korábban említésre került, teljes genomot reprezentáló DNS chip még nem készülhetett el, hiszen a búza genom szekvenálása még nem fejeződött be. A legrepresentatívabb búza DNS chip-et az Affymetrix cég (www.affymetrix.com) állítja elő és árulja. Ez a DNS „array” 55052 transzkriptumot reprezentál a termesztett búza mind a három genomját képviselve. A „DNS chip” rövid (25 bp), génspecifikus oligonukleotid szekvenciákat hordoz, amely biztosítja a magas fokú specificitást, de speciális leolvasó készüléket igényel. Ez

számunkra jelenleg nem hozzáférhető. Az Affymetrix búza DNS „chip” hátránya még, hogy drága. Egyetlen „chip” ára 500 Euro körül mozog.

A kutatók, hogy a költségeket csökkentsék és szélesebb körben elérhetővé tegyék a technikát, saját DNS „chip”-ek fejlesztésébe kezdtek. Az Egyesült Királyságban 35 különböző búza cDNS bankból összesen mintegy 10000 cDNS szekvenciát izoláltak és vittek fel DNS „chip”-re (Wilson és mtsai. 2004). Ebben az esetben, a teljes cDNS szekvenciák használata miatt, a specifitás foka alacsonyabb, mint oligonukleotidok alkalmazásakor. Ez a DNS chip kutatási célokra költség áron használható, de a kísérleteket az angliai laboratóriumban lehet csak elvégezni/elvégeztetni.

Japánban, egy búza genomikai konzorcium, 22000 búza gént reprezentáló oligonukleotid „chip”-et hozott létre, 60-70 bp hosszúságú génspecifikus oligonukleotidokat szintetizáltatva (Kawaura és mtsai. 2006). Ez a megoldás a két fentebb említett DNS „chip” kompromisszumának tekinthető: a specifitás mértéke a kettő közé tehető, de a leolvasás olcsóbb készülékekkel is megoldható. A japán kutatók sikerrel használták ezt az oligonukleotid chip-et só stressz hatására aktiválódó/csendesedő búza gének azonosítására (Kawaura és mtsai. 2006).

Laboratóriumunkban praktikus okokból mi is ezt az utóbbi megközelítést használtuk egy búza DNS „chip” előállítására. Célunk az volt, hogy egy viszonylag olcsó, így akár a nemesítők által is használható, speciális búza DNS „mikroarray”-t hozzunk létre, amely alkalmas a búza szemfejlődésével kapcsolatos génkifejeződési változások vizsgálatára. Ehhez mintegy 1500-2000 olyan búza cDNS szekvenciát választottunk ki az adatbázisokból, amelyek kizárólag vagy nagyobb gyakorisággal a búza szemfejlődéshez kapcsolódó cDNS bankokban fordulnak elő. Ezek a cDNS-ek reprezentálják a megtermékenyítéstől a szem feltöltődéséig szerepet játszó géneket, ezen belül a legfontosabb anyagcsere utak és szabályozási mechanizmusok génjeit.

Természetesen az elkészült DNS „chip” közel sem tartalmazza a szemfejlődés rendkívül komplex folyamatában szerepet játszó valamennyi gént, de bízunk abban, hogy hazai és külföldi együttműködések révén a DNS „chip” továbbfejleszthető lesz. Ennek első jele, hogy a SZBK „Sejtosztódási ciklus és stressz adaptáció” kutató csoportjában folyó kutatások eredményeként mintegy 1500 oligonukleotid készült el, amelyek a búza

környezeti stresszválaszában szerepet játszó géneket reprezentálnak. Ezek felkerültek az immár közös DNS „chip”-re, amely így közel 3000 búza gént tartalmaz.

Szekvencia adatok kezelése és feldolgozása

A modern „High throughput” technikák alkalmazásával az elérhető információ mennyisége olyan mértékben megnőtt, hogy nagy kapacitású, összetett számítástechnikai rendszerek segítségével nélkül lehetetlen elemezni az adatokat. Ennek érzékeltetésére egy példa: Az átlagos ember kb. 2-3000 szót használ a mindennapi beszédben és kb. 2-30000-t ismer. Ez nagyságrendileg megegyezik az ember vagy rizs genomjában található gének számával. Ha ehhez hozzátesszük, hogy az egyedet alkotó több száz szövet, szerv transzkripciós profilja mind különbözik egymástól és ez a profil folyamatosan változik az az egyed fejlődés során vagy különböző külső ingerek hatására (szárazság stressz, gyógyszerkezelés), akkor könnyen beláthatjuk, hogy a feltett kérdések megválaszolásához már nem elég a papír-ceruza módszer.

Mostanában egyre inkább követelmény, hogy a gének funkcióját, ne csak egy specializált kísérletben vizsgáljuk, hanem általános következtetéseket is levonjunk, beleillesztve az újonnan megszerzett tudást az eddigi ismeretekbe. Ehhez elengedhetetlen, hogy a nagyszámú speciális adatbázis (3D szerkezet, transzkripciós profil, kölcsönható gének, fehérje motívum, szignál transzdukciós utak stb.) között a kapcsolatok megléte. Létrehoztam egy olyan általános célú adatbázis-kezelő programot, mely számos a molekuláris biológiában gyakran használt fájl formátumot ismer. A program lehetővé teszi, hogy az adatbázis nyelvek használatában járatlan felhasználó is egyszerűen, grafikus felületen, képes legyen SQL lekérdezéseket tervezni. A program úgyszintén támogatja új adattáblák létrehozását, mind lekérdezések eredményéből, mind új adatok importjából. Speciális funkció, hogy támogatja a KEGG hierarchikus és „Gene Ontology” adatbázisok kapcsolását is. Ezen kívül olyan specifikus funkciók vannak beépítve, melyek megkönnyítik a még nem szekvenált organizmusok EST-éiből készített génlisták elemzését, szekvencia adatikból oligonukleotidok tervezését, elemzését. Lehetőség van új EST-k azonosítására, valamint klónkönyvtárak között a közös EST-k megkeresésére is.

Publikációs lista

Dorjgotov D, Szucs A, Otvos K, Szakonyi D, Kelemen Z, Lendvai A, Ponya Z, Barnabas B, Brown S, Dudits D, Feher A.

Specific features of RHO GTPase-dependent signaling in plants.

Cell Biol Int. 2003;27(3),191-2.

impakt faktor: 1.09

Otvos K, Pasternak TP, Miskolczi P, Domoki M, Dorjgotov D, Szucs A, Bottka S, Dudits D, Feher A. Nitric oxide is required for, and promotes auxin-mediated activation of, cell division and embryogenic cell formation but does not influence cell cycle progression in alfalfa cell cultures.

Plant J. 2005 Sep;43(6),849-60.

impakt faktor: 6.37

Szucs A, Dorjgotov D, Otvos K, Fodor C, Domoki M, Gyorgyey J, Kalo P, Kiss GB, Dudits D, Feher A.

Characterization of three Rop GTPase genes of alfalfa (Medicago sativa L.).

Biochim Biophys Acta. 2006 Jan-Feb;1759(1-2),108-15.

impakt faktor: 2.04

Szűcs A, Ignáth I, Jäger K, Barnabás B, Fehér A (2007) Fertilization induces gene expression changes in wheat egg cells within 8 hours. Plant Cell Reports (in preparation)