

# Caratterizzazione di OGM mediante l'applicazione di tecniche di Next Generation Sequencing a campioni del controllo ufficiale (IZS LT08/18 RC)

Katia Spinella, Davide La Rocca, Cristiana Fusco, Pamela Bonini, Daniela Verginelli e Ugo Marchesi

1. Istituto Zooprofilattico Sperimentale del Lazio e della Toscana "M. Aleandri" Centro di riferimento nazionale per la ricerca di organismi geneticamente modificati (CROGM)

## INTRODUZIONE

La normativa vigente dell'Unione Europea (UE) riguardo alimenti e mangimi geneticamente modificati (GM) (Reg. CE 1829/2003 e 1830/2003)<sup>1,2</sup>, ha stabilito i requisiti per autorizzazione, etichettatura e tracciabilità di organismi geneticamente modificati (OGM) e prodotti derivati, nel mercato alimentare europeo. Sul territorio nazionale **le principali criticità sono legate** al riscontro di positività nei prodotti di importazione destinati all'alimentazione umana (es: riso cinese), spesso attribuibili alla presenza di eventi non autorizzati nell'UE, poco caratterizzati dal punto di vista molecolare e non sottoposti alle procedure di valutazione del rischio imposte dalla normativa. Per garantirne la risoluzione ed il contenimento, si rende quindi necessario lo sviluppo di protocolli di caratterizzazione molecolare che si avvalgano dell'ausilio delle tecnologie di sequenziamento di nuova generazione (*Next Generation Sequencing*), a prescindere dalla conoscenza a priori degli elementi genetici esogeni eventualmente presenti in questi eventi GM ignoti.

**Il progetto mira a** sviluppare strategie analitiche su piattaforme NGS di seconda e terza generazione (Illumina e ONT) per la caratterizzazione molecolare di OGM autorizzati e non autorizzati. In Figura 1 è mostrato il workflow proposto. A tale scopo sono state esplorate tre strategie: la prima riguarda un protocollo di arricchimento a partire dal promotore 35S (P35S) per la caratterizzazione delle regioni nucleotidiche fiancheggianti, utilizzando un approccio di amplicon sequencing con *genome walking* (Fig.2A e 2B). La seconda riguarda un protocollo, sempre di "amplicon sequencing", ma (Fig.3) con rilevazione di 36 target. Il terzo approccio riguarda una strategia di whole genome sequencing (WGS) su piattaforma di terza generazione MinION Oxford Nanopore Technologies (ONT) (Fig.4).

Presenza di OGM autorizzato o non autorizzato?

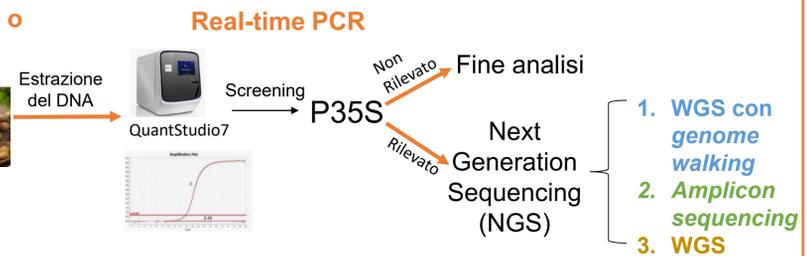
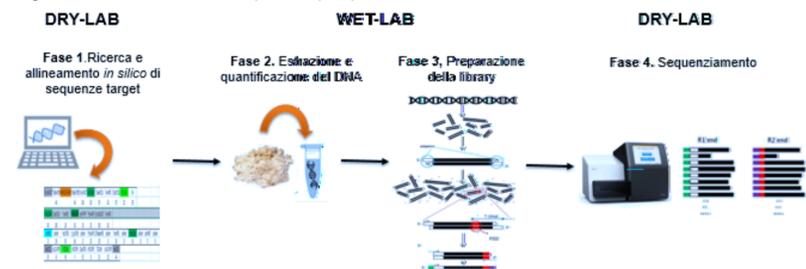


Fig. 1 Workflow dell'approccio in studio

## RISULTATI

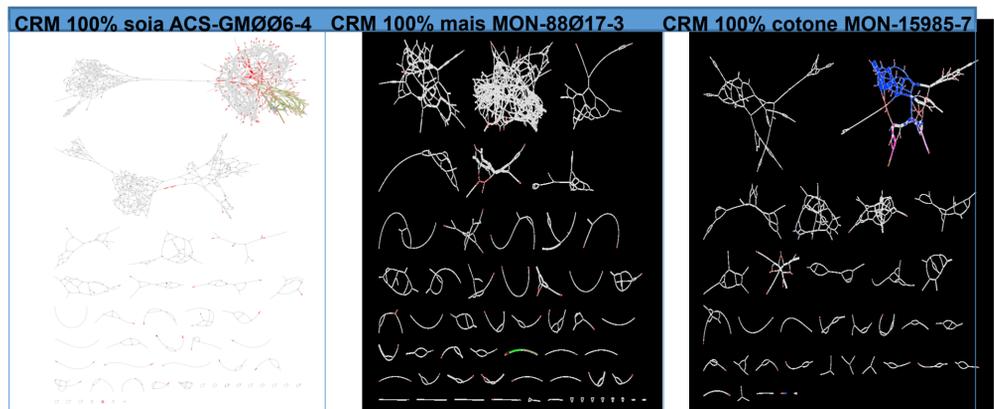
### Strategia di amplicon sequencing con Genome Walking

Fig. 2A Flusso di lavoro per la preparazione delle libraries



Ricerca e allineamento *in silico* di sequenze nucleotidiche target note appartenenti a comuni elementi di screening di eventi GM (P35S, NptII, T-nos, CTP-CP4-EPSPS, CP4EPSPS, FMV, PAT e T35S) effettuata attraverso consultazione della banca dati dell'NCBI. L'approccio di *genome walking* è stato eseguito a partire dal target selezionato, il P35S, in direzione 3'-5'. Per l'esperimento pilota sono stati selezionati in base ai risultati degli allineamenti *in silico*, tre materiali di riferimento certificati (CRMs) AOCS 100% relativi a cotone MON-15985-7, a mais MON-88017-3 e a soia ACS-GM006-4. Il DNA, estratto con Promega Maxwell RSC, quantificato fluorimetricamente, e preparato con protocollo ILLUMINA, è stato sequenziato su MiSeq ILLUMINA.

Fig. 2B Output dell'assembly utilizzando il programma Bandage e il plug-in del programma Blast. I segmenti colorati rappresentano i target evento specifici allineati contro gli assembly.



Per l'analisi delle reads è stata realizzata una pipeline con il programma *nextflow*. Le letture grezze sono state filtrate dai nucleotidi a bassa qualità di sequenziamento con il programma TRIMMOMATIC e sono state mantenute solo quelle contenenti la sequenza del primer P35S. Infine è stato eseguito un *assembly* per ogni campione e i risultati sono stati visualizzati tramite il programma Bandage. I risultati dell'esperimento pilota sui 3 CRMs confermano la presenza degli elementi di screening attesi contenuti nella cassetta transgenica specifica per l'evento GM di riferimento.

## CONCLUSIONI

Il sequenziamento contemporaneo in NGS di migliaia di frammenti di DNA provenienti dal campione in esame può rivelare la presenza di sequenze esogene e consentire la caratterizzazione delle porzioni genomiche circostanti.

- L'approccio di amplicon sequencing con *genome walking* su piattaforma ILLUMINA ha permesso, a partire da un elemento noto, di caratterizzare regioni genomiche sconosciute e la ricerca di OGM non autorizzati
- La strategia di WGS su piattaforma ONT ha permesso di sequenziare l'intero genoma di soia e cotone e di sviluppare una pipeline con la prospettiva di ricostruire nuovi eventi GM senza avere conoscenze a priori degli elementi genetici contenuti

### Strategia di amplicon sequencing; piattaforma multiscreening

Il saggio multiscreening di caratterizzazione in NGS è stato realizzato definendo un pannello di 36 target comprendenti: 10 geni endogeni, 10

elementi di screening, 12 costrutti, 2 organismi donatori e 2 target di identificazione di specie (Tab.1). Successivamente è stato messo a punto un protocollo di "targeted sequencing". Per l'esperimento pilota sono stati selezionati 5 CRMs al 100%: MON-15985-7 (AOCS 0804-D2), MON-89034-3 (AOCS 0906-E2), ACS-GM006-4 (AOCS 0707-C8), KM-000H71-4 (ERM 1053), BPS-25271-9 (ERM 889). È stata effettuata la preparazione delle *libraries* e si è proceduto al sequenziamento su piattaforma ILLUMINA. In generale i risultati osservati sono in linea con quelli attesi. In Figura 2 è mostrato l'esempio del CRM 0.1% cotone MON-15985-7: le barre blu mostrano gli elementi attesi, le arancioni quelli inattesi dovuti probabilmente ad un *carry-over*, tuttavia il numero di *cluster* di queste sequenze è trascurabile.

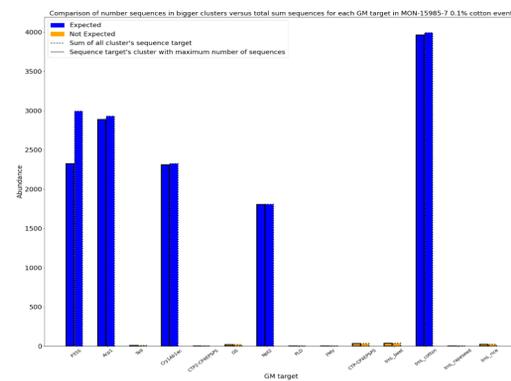


Fig. 3, CRM cotone MON-15985-7 0.1%

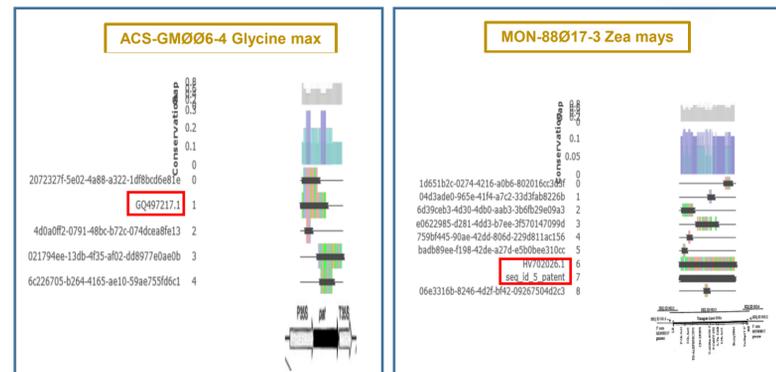
| N TARGET | TARGET                 | TIPOLOGIA                                  | ANALIZZATI |
|----------|------------------------|--|------------|
| 1        | LECTINA                | ENDOGENO Glycine max (soia)                | SI         |
| 2        | PLD                    | ENDOGENO Oryza sativa ssp Japonica (riso)  | NO         |
| 3        | HMG                    | ENDOGENO Zea mays (mais)                   | SI         |
| 4        | GS                     | ENDOGENO Beta vulgaris (barbabietola)      | SI         |
| 5        | UGPasi                 | ENDOGENO Solanum tuberosum (patata)        | SI         |
| 6        | CHIMOPAPAINA           | ENDOGENO Carica papaya (papaya)            | NO         |
| 7        | Waxy                   | ENDOGENO Triticum aestivum (frumento)      | NO         |
| 8        | SAD                    | ENDOGENO Linum usitatissimum (lino)        | NO         |
| 9        | Acp1                   | ENDOGENO Gossypium hirsutum (cotone)       | SI         |
| 10       | FATA(A)                | ENDOGENO Brassica napus (colza)            | NO         |
| 11       | LECTINA                | ORGANISMO DONATORE Pisum sativum (pisello) | NO         |
| 12       | CaMV                   | ORGANISMO DONATORE CaMV                    | NO         |
| 13       | 23579                  | IDENTIFICATORE DI SPECIE                   | SI         |
| 14       | tmL                    | IDENTIFICATORE DI SPECIE                   | SI         |
| 15       | P35S                   | SCREENING                                  | SI         |
| 16       | TNOS                   | SCREENING                                  | SI         |
| 17       | TE9                    | SCREENING                                  | SI         |
| 18       | CP4-EPSPS              | SCREENING                                  | SI         |
| 19       | CTP-CP4EPSPS           | COSTRUTTO                                  | SI         |
| 20       | CTP2-CP4EPSPS-BIS      | COSTRUTTO                                  | SI         |
| 21       | GENE NPTII             | SCREENING                                  | SI         |
| 22       | GENE PAT               | SCREENING                                  | SI         |
| 23       | GENE BAR               | SCREENING                                  | SI         |
| 24       | PFMV                   | SCREENING                                  | SI         |
| 25       | PNOS                   | SCREENING                                  | SI         |
| 26       | T35S-pCAMBIA           | COSTRUTTO                                  | NO         |
| 27       | T35S                   | SCREENING                                  | SI         |
| 28       | Cry1Ab/Cry1Ac          | COSTRUTTO                                  | SI         |
| 29       | pNOS-NptII             | COSTRUTTO                                  | SI         |
| 30       | P35S::bar              | COSTRUTTO                                  | SI         |
| 31       | P-ubi-cry1A(b)         | COSTRUTTO                                  | NO         |
| 32       | 35S-hpt                | COSTRUTTO                                  | NO         |
| 33       | P35S-hpt               | COSTRUTTO                                  | NO         |
| 34       | CptI-Tnos              | COSTRUTTO                                  | NO         |
| 35       | CryIA(c)-Tnos (TT51-1) | COSTRUTTO                                  | SI         |
| 36       | Cry1Ab/Ac-Tnos         | COSTRUTTO                                  | NO         |

Tab.1 Target selezionati

### Whole Genome Sequencing

Il WGS è stato eseguito su piattaforma di terza generazione MinION ONT, che consente di sequenziare genomi di grande dimensione e di ottenere reads lunghe (~10 Kb). A tale scopo sono stati sequenziati due CRM al 100% GM: ACS-GM006-4 soia e il mais MON-88017-3.

Fig 4: Rappresentazione degli allineamenti delle reads contro le sequenze di riferimento dei rispettivi eventi GM (a destra). A sinistra sono presenti i nomi delle sequenze utilizzate (in rosso quelle di riferimento). In basso a destra sono rappresentate le immagini degli eventi recuperati nei patent.



## BIBLIOGRAFIA

1. European Union law, EUR-Lex Access to European Union law. Regulation (EC) No 1829/2003 of the European Parliament and of the Council of 22 September 2003 on genetically modified food and feed.
2. European Union law, EUR-Lex Access to European Union law. Regulation (EC) No 1830/2003 of the European Parliament and of the Council of 22 September 2003 concerning the traceability and labelling of genetically modified organisms and the traceability of food and feed products produced from genetically modified organisms and amending Directive 2001/18/EC.