

Bioteecnologie sostenibili

COLLANA EDAGRICOLE UNIVERSITÀ & FORMAZIONE

Agricoltura sostenibile [a cura di Michele Pisante]
Microbiologia enologica [a cura di Giovanna Suzzi e Rosanna Tofalo]
Igiene degli alimenti [a cura di Maria Schirone e Pierina Visciano]
L'acqua in agricoltura [a cura di Marcello Mastroianni]
Difesa sostenibile in agricoltura [a cura di Paola Battilani]
Fertilizzazione sostenibile [a cura di Carlo Grignani]
Agricoltura di precisione [a cura di Raffaele Casa]
Malattie delle piante ornamentali
[Angelo Garibaldi, Domenico Bertetti, Stefano Rapetti, M. Lodovica Gullino]
Biotecnologie Sostenibili
[a cura di Massimo Galbiati, Alessandra Gentile, Stefano La Malfa, Chiara Tonelli]
Oli e grassi [a cura di Giuliano Mosca]
Politica agraria e di sviluppo rurale [a cura di Angelo Frascarelli]
Gestione della qualità e conservazione dei prodotti ortofrutticoli [Giancarlo Colelli e Paolo Inglese]

DIRETTORE SCIENTIFICO

Michele Pisante

COMITATO SCIENTIFICO

Marco Acutis, Aniello Anastasio, Paolo Balsari, Paola Battilani, Marco Bindi, Raffaele Casa, Luisella Celi,
Riccardo D'Andria, Guido D'Urso, Stefania De Pascale, Rosa Draisci, Angelo Frascarelli, Dario Frisio,
Massimo Galbiati, Alessandra Gentile, Carlo Grignani, Maria Lodovica Gullino, Paolo Inglese,
Stefano La Malfa, Rosalba Lanciotti, Marcello Mastroianni, Fabio Molinari, Giuliano Mosca, Erasmo Neviani,
Michele Perniola, Maria Schirone, Fabio Stagnari, Giovanna Suzzi, Rosanna Tofalo, Chiara Tonelli,
Sandra Torriani, Giovanni Vannacci, Pierina Visciano

Biotecnologie sostenibili

Scienza e innovazione in agricoltura
per affrontare le sfide della sicurezza alimentare
e della sostenibilità ambientale

a cura di

Massimo Galbiati, Alessandra Gentile,
Stefano La Malfa, Chiara Tonelli



1ª edizione: dicembre 2017



© Copyright 2017 by «Edagricole - Edizioni Agricole di New Business Media srl»
via Eritrea 21 - 20157 Milano
Redazione: Piazza G. Galilei, 6 - 40123 Bologna

Vendite: tel. 051/6575833; fax 051/6575999 - email: libri.edagricole@newbusinessmedia.it
<http://www.edagricole.it>

5534

Proprietà letteraria riservata - printed in Italy

La riproduzione con qualsiasi processo di duplicazione delle pubblicazioni tutelate dal diritto d'autore è vietata e penalmente perseguibile (art. II della legge 22 aprile 1941, n. 633). Quest'opera è protetta ai sensi della legge sul diritto d'autore e delle Convenzioni internazionali per la protezione del diritto d'autore (Convenzione di Berna, Convenzione di Ginevra). Nessuna parte di questa pubblicazione può quindi essere riprodotta, memorizzata o trasmessa con qualsiasi mezzo e in qualsiasi forma (fotomeccanica, fotocopia, elettronica, ecc.) senza l'autorizzazione scritta dell'editore. In ogni caso di riproduzione abusiva si procederà d'ufficio a norma di legge.

Realizzazione grafica: Emmegi prepress, via F. Confalonieri, 36 – 20124 Milano
Impianti e stampa: Rotolito Lombarda S.p.A. via Sondrio, 3 – 20096 Seggiano di Pioltello (MI)
Finito di stampare nel dicembre 2017

ISBN 978-88-506-5534-2

Invito alla lettura

È con entusiasmo che salutiamo la pubblicazione del volume *Biotecnologie sostenibili* per i caratteri di Edagricole – ora nel Gruppo Tecniche Nuove – ed esprimiamo un vivo senso di gratitudine ai curatori, e agli altri autori per il loro impegno. Un ringraziamento anche al prof. Michele Pisante che ha ideato e coordina la collana *Università e Formazione*, di cui il presente volume fa parte, e al comitato scientifico, per i temi identificati.

Il volume non è nuovo nel panorama editoriale italiano, perché alla fine degli anni '80 del secolo scorso, il compianto prof. Gian Tommaso Scarascia Mugnozza, di lì a poco Presidente di questa Accademia, curò un volume dal titolo “Miglioramento genetico vegetale” per i tipi della Pàtron Editore. Quel volume, dopo aver ricordato la nascita della Genetica, si soffermava sui metodi di riproduzione delle piante, sulla variabilità necessaria per condurre con successo l'attività di miglioramento genetico, per poi affrontare i temi delle tecniche e degli obiettivi del miglioramento genetico di diverse specie di interesse agrario, e concludere con gli aspetti legislativi della produzione e commercio delle novità vegetali.

Fra le tecniche di miglioramento genetico non mancavano, accanto all'ibridazione, alla citologia e alla mutagenesi, gli aspetti del trasferimento genico con metodi molecolari. È a quest'ultimo aspetto che è interamente dedicato il volume “Biotecnologie sostenibili”, che tuttavia non si sofferma solo sui metodi e sugli obiettivi in modo specifico, ma affronta gli argomenti a più vasto raggio, in tono discorsivo pur se scientificamente rigoroso, portando ad argomentazioni e riflessioni sui pro e sui contro delle tecniche mentre ne ricorda gli aspetti generali. Così, ad esempio, viene illustrato come il sequenziamento dei genomi stia consentendo di spiegare la diversità esistente in natura e nelle piante coltivate e fornisca una massa di marcatori utili per identificare caratteri e diversità genetica finora sconosciuti. Viene spiegato anche come queste conoscenze permettano di correggere particolari sequenze geniche – come si fa, o si sta tentando di fare, per combattere alcune malattie che colpiscono gli esseri umani – e ottenere, in tempi estremamente più brevi, nuove varietà assolutamente non diverse da quelle ottenute con uno dei metodi appena ricordati e che oggi sono considerati classici o tradizionali. Non manca un capitolo sulla nascita delle piante geneticamente modificate e sulla loro parabola, sulle varietà resistenti agli erbicidi e agli insetti nocivi - con riferimenti anche agli effetti sul contenuto in micotossine nella granella in queste varietà - prima di affrontare aspetti del miglioramento genetico delle specie di interesse ortofrutticolo, ed in particolare di specie importanti per la frutticoltura italiana, come limone, pesco e melo, per i quali il miglioramento genetico mediante incroci presenta difficoltà legate allo stadio giovanile delle piante derivanti da seme e alla loro prevalente natura eterozigote che porta ad un'enorme massa di segreganti nella discendenza.

Seguono tre capitoli su temi di importanza capitale per la situazione attuale e ancor più per quella futura in cui la questione ambientale e i cambiamenti climatici imporranno nuove strategie per ottimizzare l'efficienza d'uso di nutrienti minerali, richiederanno varietà in grado di adattarsi agli stress ambientali e ad agenti patogeni nuovi e finora sconosciuti o di poca importanza per l'agricoltura italiana. Sono tre capitoli che illustrano l'avanzamento delle conoscenze e le prospettive di raggiungere una più piena conoscenza dei meccanismi con cui il materiale vegetale reagisce al variare delle condizioni ambientali. Tutte conoscenze indispensabili per costituire, in modo rapido e mirato, nuove varietà più adatte alle situazioni che oggi appaiono cambiare velocemente.

Seguono tre temi abbastanza nuovi per i non addetti ai lavori, quali il miglioramento genetico per la costituzione di varietà di piante agrarie che producano: i) alimenti bio-fortificati, ossia alimenti ricchi di vitamine

Invito alla lettura

e minerali e poveri di anti-nutrienti, in grado quindi di combattere le malattie croniche; ii) alimenti funzionali, ossia alimenti ricchi di molecole in grado di svolgere un'azione preventiva sulla salute; iii) principi farmaceutici, anticorpi, enzimi terapeutici e reagenti ad uso commerciale; iv) derrate per la produzione di energia, ossia piante in grado di produrre energie rinnovabili in grado di valorizzare, attraverso la fotosintesi clorofilliana, quella fonte energetica rinnovabile rappresentata dal sole, in vista di una produzione agraria che contribuisca più efficientemente ad un miglioramento delle qualità della vita e dell'ambiente.

Il volume si chiude con tre capitoli illuminanti su argomenti indispensabili per l'impostazione strategica delle politiche agrarie: gli aspetti economici, compresa la diffusione nel mondo e l'impatto delle varietà GM sulle aziende di piccole e grandi dimensioni; i riferimenti normativi per la valutazione del rapporto rischi/benefici dell'introduzione di innovazioni, compreso il quadro normativo italiano e le Direttive europee; i principi di precauzione e di equivalenza; la necessità di ripensare l'approccio alle biotecnologie vegetali; i risultati di ricerche sul livello di conoscenze e fonti di informazione del pubblico; sulla percezione delle biotecnologie, su come questa percezione sia stata costruita e si sia evoluta negli ultimi venti anni e sui fattori che incidono sulle scelte di acquistare o non acquistare alimenti GM.

In conclusione, un volume scritto, a più mani, in modo piano e dalla lettura piacevole, oltre che scientificamente corretta, che incuriosisce e invoglia a proseguire nella lettura.

Un volume che i laureandi e laureati dei corsi di laurea afferenti al sistema agricolo, nelle sue accezioni agroforestale, agro-alimentare e agro-industriale, dovrebbero conoscere ed utilizzare per organizzare discussioni e confronti fra loro su un argomento così pregnante.

Un volume che i giornalisti dovrebbero leggere per arricchire i loro scritti di argomentazioni, sollevare interrogativi, porre questioni, chiedere conto del loro operato agli scienziati delle più diverse discipline, ai politici che emanano norme, agli opinionisti che indirizzano le masse.

Un volume che il pubblico dovrebbe leggere per conoscere e riflettere e prendere decisioni informate. I mezzi di comunicazione di massa sommergono tutti noi di informazioni, ma queste sono tante e in direzioni diverse, e spesso risulta difficile orientarsi. Il volume raccoglie e presenta in modo ordinato una serie di punti che consentono a ognuno di noi di formarsi un'opinione ragionata.

Non è un manuale, ma una narrazione pacata e semplice su un insieme di cose, situazioni e norme utili per affrontare un argomento complesso come quello delle biotecnologie sostenibili.

Modificare le piante. Perché modificarle e non lasciarle come si sono evolute in natura? Nel porsi questa domanda non andrebbe dimenticato che non appena l'uomo ha iniziato a coltivare le piante, fossero esse cereali o leguminose, le ha modificate, deliberatamente anche se empiricamente; le ha portate ad esempio a non disseminare i loro semi, per poterli raccogliere più facilmente, e a liberare i semi dagli involucri di protezione che li avvolgono, sicuramente poco appetibili. Nel diffondersi nelle terre e nei continenti, l'uomo ha portato con sé le piante che coltivava, esponendole a nuove situazioni in cui manifestavano tutta la loro variabilità e l'uomo sceglieva quelle più adatte, quelle più gradite per produzione, sapidità o perché non attaccate da batteri, funghi e insetti, o perché sopportavano alluvioni e siccità. La scelta non era facile e spesso avveniva solo dopo fallimenti che comportavano desolazione e morte.

I caratteri da modificare nelle piante si sono evoluti nel tempo, ma l'esigenza resta immutata ancora oggi per arrivare a produrre di più e in maniera sostenibile, per avere derrate più rispondenti ad una società che non sta ferma, ma si evolve in modo rapido quanto mai in precedenza. Una società che vuole chiamarsi e considerarsi evoluta, ma che dimostra, nelle sue scelte, di non avere una formazione scientifica, di non essere costituita da persone che si documentano, analizzano, si formano una opinione su fatti scientificamente provati. Certo la scienza non è priva di incertezze ed errori, ma mette continuamente in discussione i risultati precedenti e, con il procedere delle conoscenze, corregge se stessa, costruisce sistemi di controllo. Rappresenta quindi una forza propulsiva dello sviluppo della società, della sua modernizzazione ed in definitiva della sua democrazia.

Emilia Chiancone
Presidente
dell'Accademia Nazionale delle Scienze

Gli Autori

Selene Baschieri

ENEA Centro Ricerche Casaccia, Divisione Biotecnologie e Agroindustria, Roma

Giorgia Batelli

Consiglio Nazionale delle Ricerche (CNR), Istituto di Bioscienze e Biorisorse, Portici (NA)

Teodoro Cardì

CREA Centro di ricerca Orticoltura e Florovivaismo, Pontecagnano Faiano (SA)

Luigi Cattivelli

CREA Centro di ricerca Genomica e Bioinformatica, Fiorenzuola d'Arda (PC)

Felice Cervone

Dipartimento di Biologia e Biotecnologie "Charles Darwin", Università degli Studi di Roma "La Sapienza"

Monica Colombo

Dipartimento Genomica e Biologia delle Piante da Frutto, Centro Ricerca e Innovazione, Fondazione Edmund Mach, San Michele all'Adige (TN)

Lucio Conti

Dipartimento di Bioscienze, Università degli Studi di Milano

Roberto Defez

Consiglio Nazionale delle Ricerche (CNR), Laboratorio di Biotecnologie Microbiche, Istituto di Bioscienze e Biorisorse, Portici (NA)

Marcello Donini

ENEA Centro Ricerche Casaccia, Divisione Biotecnologie e Agroindustria, Roma

Gabriele Di Gaspero

Istituto di Genomica Applicata, Udine

Dario Frisio

Dipartimento di Economia, Management e Metodi Quantitativi (DEMM), Università degli Studi di Milano

Vincenzo Lionetti

Dipartimento di Biologia e Biotecnologie "Charles Darwin", Università degli Studi di Roma "La Sapienza"

Chiara Lico

ENEA Centro Ricerche Casaccia, Divisione Biotecnologie e Agroindustria, Roma

Enzo Loner

Dipartimento di Sociologia e Ricerca Sociale, Università degli Studi di Trento

Gli Autori

Carla Marusic

ENEA Centro Ricerche Casaccia, Divisione Biotecnologie e Agroindustria, Roma

Bruno Mezzetti

Dipartimento di Scienze Agrarie, Alimentari ed Ambientali, Università Politecnica delle Marche

Michele Morgante

Dipartimento di Scienze Agrarie ed Ambientali, Università degli Studi di Udine

Silvia Morgutti

Dipartimento di Scienze Agrarie e Ambientali - Produzione, Territorio, Agroenergia, Università degli Studi di Milano

Fabio Francesco Nocito

Dipartimento di Scienze Agrarie e Ambientali - Produzione, Territorio, Agroenergia, Università degli Studi di Milano

Paolo Pesaresi

Dipartimento di Scienze Agrarie e Ambientali - Produzione, Territorio, Agroenergia, Università degli Studi di Milano

Katia Petroni

Dipartimento di Bioscienze, Università degli Studi di Milano

Giuseppe Leonardo Rotino

CREA Centro di ricerca Orticoltura e Florovivaismo, Montanaso Lombardo (LO)

Silvia Sabbadini

Dipartimento di Scienze Agrarie, Alimentari ed Ambientali, Università Politecnica delle Marche

Gian Attilio Sacchi

Dipartimento di Scienze Agrarie e Ambientali - Produzione, Territorio, Agroenergia, Università degli Studi di Milano

Stefano Tartarini

Dipartimento di Scienze Agrarie, Alma Mater Studiorum-Università di Bologna

Giampiero Valè

CREA Centro di ricerca Cerealicoltura e colture industriali, Vercelli

Vera Ventura

Dipartimento di Economia, Management e Metodi Quantitativi (DEMM), Università degli Studi di Milano

Indice generale

Invito alla lettura	Pag.	V
Gli Autori	"	VII
1. Next Generation Breeding: le conoscenze genomiche rivoluzionano il miglioramento genetico (Luigi Cattivelli, Giampiero Valè)	"	1
1.1 Introduzione	"	1
1.2 Il sequenziamento dei genomi spiega la diversità delle specie coltivate	"	4
1.3 Un'illimitata disponibilità di marcatori molecolari	"	6
1.4 Associazioni carattere-marcatore in collezioni di germoplasma	"	6
1.5 Predizione del <i>breeding value</i> e selezione genomica	"	9
1.6 <i>Next generation breeding</i>	"	10
Bibliografia	"	11
2. La parabola degli OGM e la prima generazione di piante transgeniche (Roberto Defez)	"	15
2.1 Definizione di OGM	"	15
2.2 La cassetta degli attrezzi del giardiniere genetico	"	15
2.3 Le prime piante OGM vedono la luce	"	19
2.3.1 Le varietà HR	"	20
2.3.2 Le varietà IR (varietà Bt)	"	20
2.4 L'impatto ambientale delle varietà HR e IR	"	21
2.4.1 Sviluppo d'infestanti resistenti a erbicidi	"	21
2.4.2 Sviluppo d'insetti resistenti al Bt	"	22
2.4.3 Effetti delle colture Bt su insetti non-target	"	22
2.4.4 Effetti delle colture Bt sull'accumulo di micotossine nella granella	"	23
2.5 Uno sguardo alla situazione italiana	"	24
2.6 Considerazioni conclusive	"	25
Bibliografia	"	25
3. Dalla trans-genesi alla cis-genesi e al genome editing (Michele Morgante e Gabriele Di Gaspero)	"	27
3.1 Introduzione	"	27
3.2 Miglioramento genetico 4.0, ossia basato sulla conoscenza	"	28
3.3 La rigenerazione <i>in vitro</i>	"	28
3.4 Cisgenesesi	"	29
3.5 <i>Genome editing</i>	"	30
Bibliografia	"	33

4. Breeding di specie ortofrutticole

(Teodoro Cardi, Alessandra Gentile, Stefano La Malfa, Bruno Mezzetti, Giuseppe Leonardo Rotino, Silvia Sabbadini, Stefano Tartarini)

	Pag.	37
4.1	Introduzione	37
4.2	I principali obiettivi del miglioramento genetico per le specie ortofrutticole	37
4.3	Metodi di miglioramento genetico e ruolo delle biotecnologie	38
4.4	Alcuni casi studio	40
4.4.1	Il miglioramento genetico del limone per resistenza a patogeni fungini	40
4.4.2	La resistenza a ticchiolatura (<i>Venturia inaequalis</i>) del melo	43
4.4.3	Tecniche di ingegneria genetica applicate a pesco per la resistenza a PPV	48
4.4.4	Induzione di resistenza a Dorifora (<i>Leptinotarsa decemlineata</i> Say.) in melanzana	50
4.4.5	Ingegnerizzazione del carattere partenocarpia in melanzana e pomodoro	51
4.5	Conclusioni e prospettive	55
	Bibliografia	56

5. Efficienza d'uso dei nutrienti minerali nelle piante (Fabio Francesco Nocito, Silvia Morgutti, Gian Attilio Sacchi)

	"	63
5.1	Introduzione	63
5.2	L'efficienza d'uso dei nutrienti	63
5.3	L'efficienza di assorbimento dei nutrienti: il ruolo dell'architettura radicale e del trasporto	66
5.4	L'efficienza di utilizzazione dei nutrienti tra fotosintesi e senescenza	68
5.5	Strategie per il miglioramento dell'efficienza d'uso dei nutrienti	69
5.5.1	Il miglioramento genetico della NUE: alcune considerazioni	70
5.5.2	Indici diagnostici dello stato nutrizionale di una coltura	71
5.6	Le piante sentinella dello stato nutrizionale	71
5.6.1	Il concetto di fusione genica in bioindicazione	72
5.7	Conclusioni e prospettive	74
	Bibliografia	75

6. Adattamento e resistenza delle piante agli stress ambientali

(Lucio Conti, Giorgia Batelli)

	"	79
6.1	Introduzione	79
6.2	Definizione degli stress ambientali, loro impatto in agricoltura	79
6.3	Strategie di tolleranza a stress	81
6.4	Tipi di stress ambientale, loro percezione e segnale	81
6.5	Le proteine di segnalazione dello stress: mediatori dell'interazione pianta-ambiente	82
6.5.1	Percezione del segnale di stress: possibili sensori e il calcio come messaggero del segnale	83
6.5.2	Mediatori della omeostasi ionica in risposta a stress salino	83
6.6	Dopo i primi segnali: il ruolo chiave dell'ABA	84
6.7	Regolazione genica e risposta a stress abiotici	85
6.8	Il ciclo dell'acqua	86
6.8.1	Approcci biotecnologici per limitare il consumo d'acqua	86
6.9	Controllo della crescita e implicazioni nella resistenza agli stress	87
6.10	Applicazioni biotecnologiche per migliorare la tolleranza agli stress in agricoltura	88
6.11	Conclusioni	89
	Bibliografia	90

7. La resistenza ai patogeni: approcci biotecnologici sostenibili per il miglioramento della resistenza alle malattie (Vincenzo Lionetti, Felice Cervone)

	"	93
7.1	Introduzione	93
7.2	Le interazioni pianta-patogeno	93
7.3	Protezione delle piante e miglioramento della resistenza	96
7.3.1	Biostimolanti ed elicitori	96
7.3.2	Transgenesi	98

7.3.3	Cisgenesi	Pag.	99
7.3.4	<i>Editing</i> del genoma	"	100
Bibliografia	"	103
8.	Biofortificazione e alimenti funzionali (Katia Petroni)	"	107
8.1	Introduzione	"	107
8.2	Strategie di ingegneria metabolica	"	107
8.3	Biofortificazione per combattere la malnutrizione	"	110
8.3.1	Vitamine	"	110
8.3.2	Minerali e antinutrienti	"	118
8.4	Alimenti funzionali per combattere le malattie croniche	"	119
8.4.1	I polifenoli	"	120
8.4.2	Gli Omega-3	"	122
Bibliografia	"	124
9.	<i>Molecular farming</i>. Piante come 'biofabbriche' per la produzione di biofarmaceutici (Marcello Donini, Carla Marusic, Chiara Lico, Selene Baschieri)	"	127
9.1	Introduzione	"	127
9.2	Vantaggi e svantaggi della produzione di biofarmaci in pianta	"	128
9.3	Strategie utilizzate per la produzione di biofarmaci in pianta	"	130
9.3.1	Trasformazione stabile	"	130
9.3.2	Trasformazione transiente o epicromosomale	"	131
9.4	Produzione di anticorpi in pianta	"	134
9.4.1	Anticorpi monoclonali	"	135
9.4.2	Frammenti anticorpali ricombinanti	"	135
9.5	Produzione di molecole a scopo vaccinale	"	136
9.6	Produzione di enzimi terapeutici e reagenti ad uso commerciale	"	139
9.7	Il futuro del <i>Molecular Farming</i>	"	141
Bibliografia	"	141
10.	Bioteecnologie per la produzione sostenibile di bioenergia (Paolo Pesaresi, Monica Colombo)	"	145
10.1	Introduzione	"	145
10.2	Il problema energetico nel mondo	"	145
10.2.1	La politica energetica dell'Unione Europea	"	146
10.2.2	Le energie rinnovabili nel mondo	"	147
10.3	La biomassa: una fonte di energia rinnovabile	"	148
10.4	Ricavare energia dalla biomassa	"	149
10.4.1	Processi termochimici	"	149
10.4.2	Processi chimici/biochimici	"	150
10.4.3	La filiera dei combustibili solidi	"	150
10.4.4	La filiera dei combustibili liquidi	"	151
10.4.4.1	Il bioetanolo	"	151
10.4.4.2	Biodiesel	"	152
10.4.4.3	Il bio-olio	"	153
10.4.5	La filiera dei combustibili gassosi	"	153
10.4.5.1	Il biogas	"	153
10.4.5.2	Syngas	"	153
10.5	I biocarburanti	"	154
10.5.1	Biocarburanti di prima generazione	"	154
10.5.1.1	Produzione di bioetanolo da zucchero ed amido	"	154
10.5.1.2	Produzione di biodiesel da colture olearie	"	154
10.5.1.3	La sostenibilità dei biocarburanti di prima generazione	"	154
10.5.2	Biocarburanti di seconda generazione	"	155
10.5.2.1	Colture energetiche	"	155

Indice generale

10.5.2.2 Residui agricoli e forestali	Pag. 157
10.5.2.3 Rifiuti urbani ed industriali	" 157
10.5.2.4 Sfide nella commercializzazione	" 158
10.5.2.5 Il miglioramento genetico delle colture energetiche: breeding classico e biotecnologie	" 159
10.5.3 Biocarburanti di terza generazione	" 165
10.5.3.1 Vantaggi delle microalghe	" 166
10.5.3.2 Il processo produttivo	" 168
10.5.3.3 Il miglioramento delle microalghe	" 169
10.5.4 Biocarburanti di quarta generazione	" 175
10.6 Conclusioni	" 175
Bibliografia	" 175
11. Aspetti economici delle biotecnologie (Dario Frisio, Vera Ventura)	" 183
11.1 Introduzione	" 183
11.2 Ricerca e sviluppo – piante GM	" 185
11.2.1 <i>New Breeding Techniques</i>	" 187
11.3 Fase sperimentale	" 187
11.4 Approvazioni e forme di mercato	" 191
11.5 Diffusione	" 194
11.5.1 Status delle coltivazioni GM nel mondo	" 194
11.6 Impatto economico	" 196
11.6.1 Impatto a livello aziendale	" 198
11.6.2 Impatto sulle piccole aziende	" 198
11.6.3 Effetti sull'ambiente	" 199
11.6.4 Coesistenza	" 200
11.7 Considerazioni finali. Innovazione e sostenibilità	" 202
Bibliografia	" 203
12. Valutazione del rischio e normativa sugli Organismi Geneticamente Modificati (Bruno Mezzetti)	" 205
12.1 Introduzione	" 205
12.2 Quadro normativo di riferimento: internazionale, europeo ed italiano	" 206
12.2.1 Il protocollo di Cartagena: il più importante riferimento normativo internazionale	" 206
12.2.2 Le normative europee: dalla Direttiva 90/220 alla Direttiva 2001/18	" 208
12.2.2.1 Indicazioni in materia di geni marcatori e di etichettatura dei prodotti	" 209
12.2.3 Il regolamento 1829/2003 relativo all'importazione di alimenti e mangimi OGM	" 211
12.2.3 Il regolamento 1830/2003 relativo alla tracciabilità ed etichettatura di alimenti e mangimi OGM	" 212
12.3 La normativa italiana	" 213
12.4 L'ultima direttiva comunitaria e i decreti di recepimento: la scelta agli stati membri di vietare la coltivazione degli OGM	" 215
12.5 Principio di sostanziale equivalenza e il principio di precauzione	" 216
12.6 Valutazione rischi e benefici piante GM	" 217
12.6.1 Rischio per la salute del consumatore	" 218
12.6.2 Rischi per l'ambiente	" 219
12.7 Ripensare l'approccio alle biotecnologie vegetali	" 220
Bibliografia	" 222
13 La percezione pubblica delle biotecnologie in Italia (Enzo Loner)	" 223
13.1 Introduzione	" 223
13.2 I risultati delle ricerche precedenti	" 223
13.3 I dati sulle biotecnologie in Italia	" 225
13.4 Cosa pensano gli italiani delle biotecnologie	" 226
13.5 Il livello di conoscenza e le fonti di informazione	" 227

13.6	L'utilizzo delle biotecnologie per modificare le piante	Pag.	228
13.7	L'utilizzo delle biotecnologie sugli animali	"	230
13.8	I cibi geneticamente modificati	"	232
13.9	Quali fattori incidono sulla scelta di (non) acquistare cibo GM	"	234
13.10	Conclusioni	"	238
	Bibliografia	"	240
	Glossario	"	243

1 Next Generation Breeding: le conoscenze genomiche rivoluzionano il miglioramento genetico

Luigi Cattivelli, Giampiero Valè

1.1 Introduzione

Nel corso degli ultimi cinquant'anni, i metodi tradizionali di miglioramento genetico sono stati integrati con nuove tecnologie capaci di migliorarne l'efficienza o anche di generare nuovi caratteri non presenti nel germoplasma coltivato (Fig. 1.1).

A partire dagli anni '60 la mutagenesi è stata estesivamente utilizzata per generare nuova diversità genetica da inserire direttamente nei programmi di *breeding*. Successivamente, l'utilizzo delle colture *in vitro* ha consentito di indurre nuova diversità (variazione somaclonale) e ha rappresentato la premessa metodologica per la trasformazione genetica che dagli anni '80 ha reso possibile il trasferimento di geni anche tra organismi distanti tra loro.

Dai primi anni '90 la disponibilità di marcatori molecolari ha consentito di determinare associazioni tra specifici marcatori ed i caratteri più rilevanti e di utilizzare questi marcatori per seguire i caratteri associati nel corso della selezione (*Marker Assisted Selection*, MAS). La selezione assistita ha permesso di ottenere risultati eccellenti quando applicata a

caratteri semplici o comunque controllati da pochi loci con grande effetto fenotipico, tipicamente i loci che controllano la resistenza alle malattie. Nel caso di caratteri complessi controllati da molti loci con piccoli effetti quantitativi (QTL) e largamente influenzati dall'ambiente, come ad esempio la produzione o la tolleranza alla siccità, la selezione assistita non ha raggiunto i risultati desiderati e l'applicazione della selezione genomica (*Genomic Selection*, GS) al miglioramento genetico vegetale potrebbe consentire di superare questi limiti e di sviluppare una selezione basata sul genotipo. Le avanzate conoscenze sul genoma delle piante e sulla funzione dei geni hanno aperto la strada a diverse tecniche capaci di modificare in modo predeterminato la sequenza di un gene target (*genome editing*).

Lo sviluppo delle conoscenze relative ai genomi delle piante è strettamente connesso allo sviluppo delle tecnologie per il sequenziamento del DNA. A partire dai primi anni 2000, il progresso tecnologico in questo settore ha compiuto enormi progressi con la comparsa di sequenziatori ad alta processività (*Next Generation Sequencing* o NGS) (Scheda 1.1)

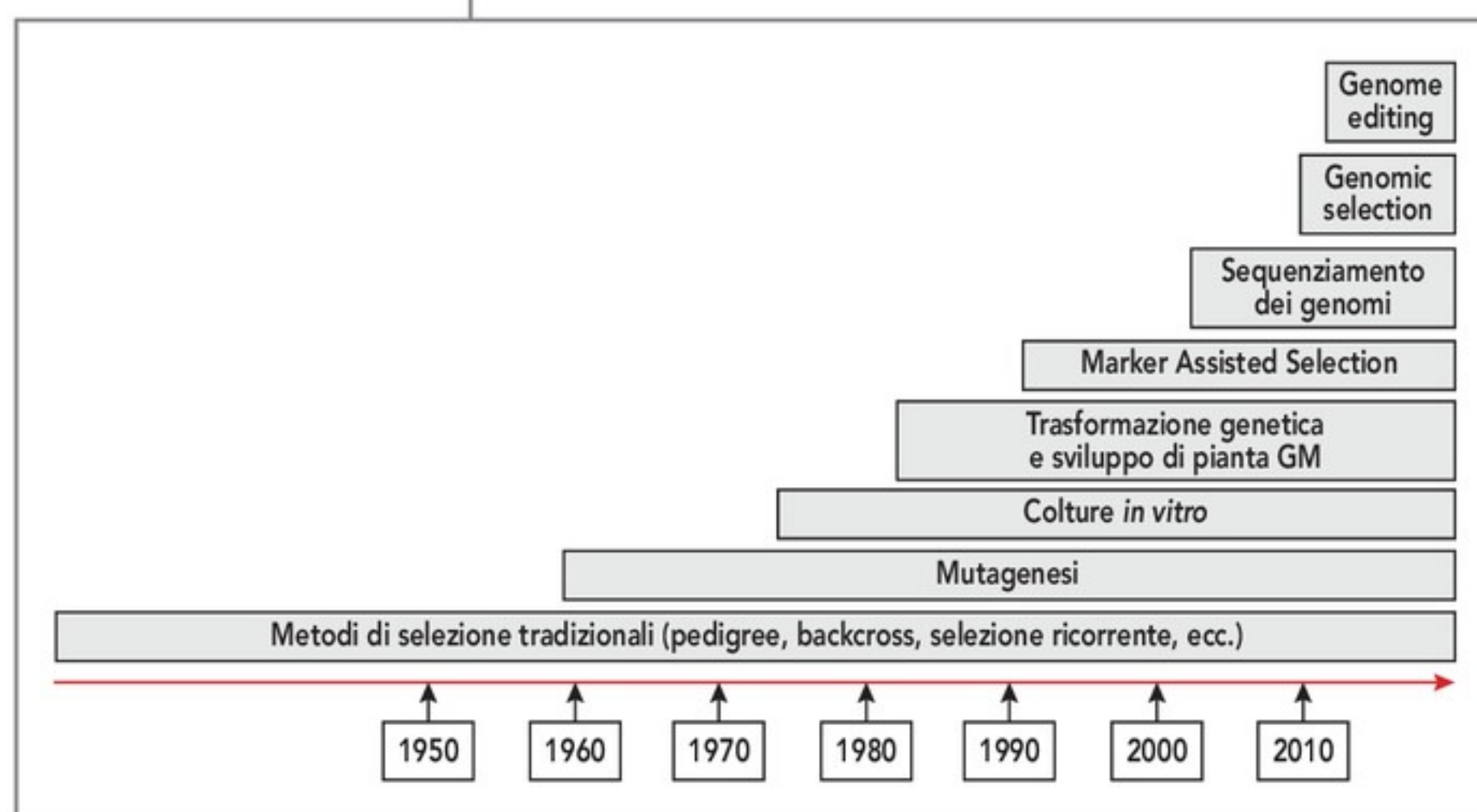


Figura 1.1 - Evoluzione dei metodi e delle biotecnologie per il miglioramento genetico delle piante.

Strumentazioni NGS che producono brevi sequenze (da 100 fino a 1.000 bp)

In questi strumenti il DNA è amplificato in modo clonale, generando migliaia di copie identiche di un dato frammento di DNA in un'area definita. La procedura consente di creare diversi milioni di centri di reazione individuali, ognuno con i suoi templati clonali di DNA, in modo che il segnale generato dalla successiva reazione di sequenziamento sia di entità tale da essere distinguibile dal rumore di *background*. Un sequenziatore NGS può raccogliere le informazioni da diversi milioni di centri di reazione simultaneamente e quindi sequenziare diversi milioni di molecole di DNA in parallelo. Due diverse strategie sono state messe a punto per la generazione di popolazioni clonali di DNA: i) sistemi di *emulsion PCR* in cui il DNA template legato a sequenze adattatori presenti su *beads* è incluso in micelle acquose dove sono presenti dNTP, primer e DNA polimerasi e l'amplificazione avviene nelle micelle in modo che ogni bead avrà migliaia di copie della stessa sequenza di DNA (sequenziatori 454 e Ion Torrent); ii) amplificazione in fase solida, in cui il DNA frammentato è legato a una sequenza adattatore a cui si appaia un primer entrambi immobilizzati su un supporto solido (sequenziatori Illumina).

Sequenziatori 454 e Ion Torrent. In queste strumentazioni, ognuno dei quattro nucleotidi viene aggiunto singolarmente alla reazione di sequenza con un processo iterativo. Conseguentemente questi sistemi non richiedono che i dNTP siano bloccati, dal momento che la assenza del dNTP successivo nella reazione di sequenza previene l'allungamento. Nel primo sequenziatore NGS ad essere sviluppato, denominato 454, il template legato a beads è distribuito in pico-piastre insieme a beads che contengono un cocktail enzimatico. Quando un dNTP è incorporato, si verifica una cascata enzimatica che risulta in un segnale bioluminescente che è rilevato da una camera CCD. La strumentazione *Ion Torrent*, invece, rileva ioni H⁺ che sono rilasciati quando un dNTP è incorporato. Il cambiamento di pH risultante (che è proporzionale al numero di nucleotidi rilevati inseriti nella sequenza) è rilevato da un sensore costituito da un ossido-metallico-semiconduttore (CMOS) e da un transistor sensibile a ioni (ISFET). Le diverse versioni delle strumentazioni 454 (oggi non più in produzione) consentono di ottenere da 35 a 700 Mb di sequenze con sequenze singole lunghe da 450 a 1.000 bp. Le varie strumentazioni Ion Torrent disponibili consentono di ottenere da 50 Mb fino a 10 Gb di sequenze con sequenze singole lunghe da 200 a 400 bp.

Sequenziatori Illumina. Nei sequenziatori Illumina, al DNA template è aggiunto un adattatore al quale successivamente si appaia un primer ed in questa regione di DNA a doppio filamento avviene il legame della DNA polimerasi. Durante ogni ciclo viene aggiunta una miscela di reazione contenente i 4 dNTP marcati con fluorofori in cui il gruppo 3'OH del ribosio è bloccato; questo comporta che ogni successivo allungamento è impedito. Dopo l'incorporazione di un singolo dNTP, che ha quindi allungato il filamento complementare, i dNTP non legati vengono rimossi e il sistema rileva quale dNTP è stato incorporato. Il fluoroforo ed il blocco al ribosio vengono quindi rimossi ed un nuovo ciclo può cominciare. Nella maggior parte delle piattaforme Illumina ogni dNTP è coniugato a uno specifico fluoroforo, per cui la identificazione del dNTP aggiunto è consentita grazie a quattro canali laser che consentono di rilevare ogni fluoroforo aggiunto (solo le piattaforme NextSeq e Mini-Seq usano un sistema a due fluorofori). I sequenziatori Illumina, dominanti in questo settore, comprendono sia piccole strumentazioni che consentono di ottenere da 1,6 a 7,5 Gb di sequenze con sequenze singole lunghe fino a 150 bp, sia grandi strumenti capaci di generare 800-900 Gb di sequenza con sequenze singole lunghe fino a 150 bp. I sequenziatori Illumina consentono inoltre di sequenziare i frammenti di DNA a partire da entrambe le estremità in

SEGUE SCHEDA 1.1

modo da generare due corte sequenze appaiate ad una distanza prefissata, una conoscenza molto utile in fase di assemblaggio.

Strumentazioni NGS che producono lunghe sequenze (da 8 fino a 200 Kb)

Le lunghe sequenze prodotte da queste strumentazioni possono risolvere regioni genomiche complesse o ripetute mediante sequenze singole e continue, permettendo di eliminare le ambiguità nella loro posizione e dimensioni. Le sequenze lunghe sono anche utili nelle analisi trascrittomiche in quanto possono includere interi mRNA, consentendo la precisa identificazione di esoni e prodotti di splicing alternativi.

Strumentazioni per il sequenziamento in real-time di singole molecole. Queste strumentazioni, diversamente da quelle che producono corte sequenze, non si basano su una popolazione, clonale di frammenti di DNA per generare un segnale rilevabile, come anche non necessitano cicli chimici per ogni dNTP aggiunto. La tecnologia di Pacific Biosciences (PacBio) è quella attualmente più utilizzata. Lo strumento usa una cella a flusso contenente molte migliaia di pozzetti a picolitri individuali con fondo trasparente denominati ZWM (*Zero-Mode Waveguides*). I frammenti del DNA template sono ligati ad adattatori con struttura hairpin ad ogni estremità, risultante in una molecola di DNA circolare a singolo filamento alle estremità e doppio filamento nelle altre zone. La DNA polimerasi è fissata alla base del ZWM e il DNA scorre nel pozzetto, permettendo il processamento di una singola molecola. L'incorporazione di ogni dNTP marcato da parte della polimerasi è visualizzata in real time tramite un laser e letta da una camera CCD. La strumentazione PacBio consente di ottenere da 500 Mb fino a 7 Gb di sequenze per reazione con lunghezze delle singole sequenze di 8-20 Kb. Un'alternativa a PacBio è rappresentata dalla tecnologia di Oxford Nanopore, in questo caso il DNA è inizialmente frammentato in dimensioni di 8-10 Kb e successivamente due diversi adattatori, una sequenza leader ed una *hairpin*, sono ligati alle estremità dei frammenti di DNA a doppio filamento. La sequenza leader interagisce con un poro proteico e determina il passaggio di un singolo filamento di DNA attraverso il poro, questo passaggio causa una variazione di voltaggio dipendente dalla composizione nucleotidica della sequenza che attraversa il poro. Queste variazioni di voltaggio sono registrate e utilizzate per identificare la composizione nucleotidica della sequenza. Quando la sequenza arriva alla struttura hairpin, il DNA continua ad essere traslocato attraverso il poro così che anche il filamento complementare viene sequenziato. Questo consente di avere sequenze forward e reverse che sono usate per creare una sequenza consenso. La strumentazione MinION di Oxford Nanopore consente di ottenere fino 1,5 Gb con sequenze singole lunghe fino a 200 Kb, tuttavia si caratterizza per un elevato *tasso di errore*.

Strumentazioni per il sequenziamento basato su approcci sintetici. Queste tecnologie permettono di suddividere frammenti di DNA di grosse dimensioni in micropozzetti o in un'emulsione in modo che poche grandi molecole di DNA siano presenti in ogni partizione. In ciascuna partizione i templati sono frammentati e i frammenti ottenuti sono identificati con un *barcoding*. Successivamente i frammenti sono sequenziati con sequenziatori Illumina e le sequenze dotate del medesimo *barcode* (e quindi derivanti dal medesimo frammento di grosse dimensioni) sono ri-assemblate tra loro. Il sistema di *barcoding* proposto da Illumina (*synthetic long-read sequencing*) prevede la selezione di frammenti di DNA di circa 10 Kb che vengono poi distribuiti in una micropiastra in modo tale che ogni pozzetto contenga circa 3.000 molecole. Ogni frammento è quindi digerito in dimensioni di circa 350 bp e sottoposto a *barcoding* con un *barcode* diverso per ogni pozzetto. Il DNA con il *barcode* è

SEGUE SCHEDA 1.1

quindi raggruppato in pool e sequenziato. La tecnologia proposta da 10X Genomics si basa invece su un sistema microfluidico per emulsionare e ripartire il DNA. Lunghi frammenti di DNA (fino a 100 Kb) sono separati in micelle chiamate GEM in cui sono presenti delle *beads* contenenti i *primers* con il *barcode* specifico, l'enzima ed i dNTP in modo da consentire l'amplificazione di corti frammenti di DNA (circa 350 bp) dotati del *barcode* che identifica il GEM. L'emulsione viene quindi rotta ed i piccoli frammenti amplificati vengono sequenziati. Successivamente la ricostruzione della sequenza di un determinato frammento è effettuata sia partendo dalle sequenze che condividono lo stesso *barcode* sia considerando che nella preparazione iniziale sono presenti diversi lunghi frammenti derivanti dalla medesima regione genomica. Entrambi i sistemi forniscono sequenze singole di circa 100 Kb, con una capacità di sequenziamento dipendente dalla piattaforma di sequenziamento utilizzata.

che hanno reso estremamente semplice ed economico sequenziare i genomi delle specie coltivate. Il sequenziamento a basso costo permette di descrivere compiutamente la diversità genetica attraverso estesi lavori di ri-sequenziamento dell'intero genoma o di parti di esso come ad esempio la parte espressa (esoma). Il sequenziamento NGS ha inoltre generato un illimitato numero di marcatori molecolari a basso costo, prevalentemente SNP (*Single Nucleotide Polymorphism*), aprendo la strada all'uso esteso dei marcatori molecolari per l'analisi genetica di vaste collezioni di germoplasma, l'isolamento di geni e la selezione genomica (Barabaschi *et al.*, 2016).

Le conoscenze genomiche degli ultimi due decenni hanno cambiato il miglioramento genetico aggiungendo alla tradizionale selezione fenotipica un importante componente molecolare, sia in termini di marcatori per selezione assistita o per selezione genomica, sia in termini di conoscenze sulla funzione dei geni che influenzano caratteri agronomici fondamentali per creare nuove piante transgeniche, cisgeniche o con sequenze modificate mediante *genome editing*. Queste tendenze verranno ulteriormente accentuate nel prossimo futuro quando la sempre maggiore diffusione di tecnologie NGS, di marcatori molecolari ad alta processività e basso costo e di conoscenze sulla funzione dei geni trasformerà il *breeding* attuale in quello che è stato definito come *Next Generation Breeding* (Barabaschi *et al.*, 2016). Nel seguito di questo capitolo si darà conto dei recenti progressi ottenuti tramite l'impiego delle tecniche di sequenziamento e dei marcatori molecolari, mentre la presentazione degli aspetti legati allo sviluppo di piante modificate mediante trasformazione o *genome editing* sarà approfondita nei capitoli successivi (cfr. Capp. 2 e 3).

1.2 Il sequenziamento dei genomi spiega la diversità delle specie coltivate

Il genoma di riferimento di molte specie coltivate è stato completato (Wendel *et al.*, 2016) ed altri progetti di sequenziamento sono in corso (<http://www.ncbi.nlm.nih.gov/genomes/leuks.cgi>; <http://plants.ensembl.org/index.html>; <http://phytozome.jgi.doe.gov/pz/portal.html>).

Queste conoscenze hanno consentito di riscrivere le relazioni filogenetiche tra le specie e di comprendere le dinamiche evolutive che hanno portato alle specie oggi esistenti (Barabaschi *et al.*, 2011). Studi basati sul confronto tra i genomi hanno identificato riarrangiamenti, duplicazioni, ed eventi di poliploidizzazione che hanno segnato la storia evolutiva delle piante coltivate sino a consentire la ricostruzione *in silico* dei proto-cromosomi ancestrali (Salse, 2012).

Il sequenziamento ha definito compiutamente le dimensioni dei genomi, il numero, l'organizzazione e la ridondanza delle sequenze espresse, ed ha consentito di gettare luce sul ruolo degli elementi ripetuti (principalmente trasposoni e retrotrasposoni) che costituiscono larga parte dei genomi delle piante coltivate. I genomi delle specie coltivate differiscono largamente in fatto di dimensioni e due principali fattori sono responsabili di queste differenze: la quantità di elementi ripetuti e il livello di ploidia. In generale è stata osservata una diretta relazione tra la dimensione del genoma e il contenuto in elementi ripetuti, una considerazione supportata dal confronto sia tra specie lontane sia tra specie strettamente correlate dal punto di vista filogenetico. All'interno delle monocotiledoni, ad esempio, l'aumento di dimensioni tra il genoma del riso (389

Mb) e quello del mais (2.300 Mb) o dell'orzo (5.000 Mb) è ascrivibile in misura significativa anche al diverso contenuto in elementi ripetuti (35% in riso, più dell'80% in mais e orzo) (Wendel *et al.*, 2016). Mentre la differenza tra i genomi dell'orzo e quelli del frumento duro (12.000 Mb) o del frumento tenero (17.000 Mb) è spiegata da due successivi eventi di poliploidizzazione, il primo dei quali ha dato origine ai frumenti tetraploidi circa mezzo milione di anni fa, mentre il secondo ha originato i frumenti esaploidi circa 10.000 anni fa (Middleton *et al.*, 2014). Malgrado la dimensione dei genomi sia estremamente variabile il numero di geni nella maggior parte delle specie coltivate è compreso tra 30.000 e 60.000 con poche eccezioni, ad esempio il frumento tenero con circa 124.000 geni accumulatisi a seguito di recenti eventi di poliploidizzazione (Wendel *et al.*, 2016).

La disponibilità di tecnologie di sequenziamento a basso costo permette di sequenziare per intero i genomi di ampie collezioni di germoplasma (cultivar, landrace e accessioni selvatiche) consentendo di comprendere le basi genetiche degli eventi di domesticazione che hanno condotto alla riduzione della diversità dei loci sottoposti a selezione (*bottleneck of selection*) (Huang *et al.*, 2012), di descrivere compiutamente tutti gli alleli di uno stesso locus, di identificare inserzioni/delezioni e mutazioni (Xu *et al.*, 2012) e finanche di spiegare alcuni dei meccanismi alla base dell'eterosi* (Schnable e Springer, 2013). Un esempio eclatante è il lavoro di ri-sequenziamento di circa 3.000 accessioni di riso che rappresentano larga parte della diversità genetica e funzionale disponibile per questa specie (3.000 *Rice Genomes Project*) (Li e Zeigler Wang, 2014). Disporre di collezioni di genomi interamente sequenziati offre la possibilità di affrontare lo studio della biodiversità da un punto di vista diverso rispetto a quello tradizionale basato su osservazioni fenotipiche. Nel caso di caratteri per i quali sono noti i geni che li determinano è infatti possibile ricercare varianti alleliche di tali geni *in silico* e valutare a posteriori l'impatto delle varianti più significative sul fenotipo, tracciare l'evoluzione di alleli ed aplotipi e individuare marcatori specifici per i singoli alleli da usare nella selezione assistita. Per esempio, diversi alleli di geni di resistenza con caratteristiche funzionali differenziate e capaci pertanto di conferire resistenze verso razze diverse di un determinato patogeno, sono stati identificati in diverse specie quali patata (Wang *et al.*, 2008), grano (Bhullar *et al.*, 2010), riso (Wang *et al.*, 2014) ed orzo (Biselli *et al.*, 2013), aumentando la disponibilità di fonti di resistenze per il miglioramento genetico. Il ri-sequenziamento può essere applicato a collezioni di piante derivanti

da mutagenesi (*TILLING-by-Sequencing*). In un primo approccio, pool di DNA ottenuti da decine/centinaia di piante M2 (seconda generazione dei semi sottoposti a trattamento mutageno) sono utilizzati per amplificare sequenze di geni per cui si ricercano mutazioni (sequenze target). I prodotti di amplificazione sono quindi sequenziati con piattaforme *Next Generation Sequencing* e analisi bioinformatiche dei dati di sequenza consentono la identificazione degli SNP presenti nei mutanti rispetto alla sequenza wild type (Tsai *et al.*, 2011). Il sistema consente di identificare piante che portano mutazioni a carico di specifici geni e utilizzarle sia per l'analisi funzionale sia nei programmi di *breeding* (Kim e Tai, 2014). Più recentemente, collezioni di alcune migliaia di piante M2 sono state sottoposte al sequenziamento dell'intera frazione espressa del genoma, selezionata tramite procedure di *exome capture* in modo da costituire un catalogo di tutte le mutazioni all'interno dei geni presenti nella collezione. I dati sono consultabili online, in modo tale che risulta possibile selezionare piante con mutazioni in specifici geni. Le piante portanti le mutazioni desiderate sono poi richieste ai curatori della risorsa genetica in modo che ogni ricercatore può disporre di specifici mutanti senza dover effettuare ogni volta la procedura di selezione (Henry *et al.*, 2014; Krasileva *et al.*, 2017). Una selezione di mutanti basati sul sequenziamento consente anche di selezionare mutazioni nelle diverse copie di un gene multiallelico (ad esempio nel caso di alleli omeologhi localizzati sui diversi genomi delle specie poliploidi) e accumularle in un unico individuo tramite incroci sino ad esprimere il fenotipo. Seguendo questa strada è stato possibile accumulare mutanti per le tre copie omeologhe del gene *Waxyin* frumento esaploide con produzione di amido costituito in maggior parte da amilopectina e quindi dotato di particolari proprietà fisico-chimiche (Slade *et al.*, 2005).

Accanto alla descrizione della diversità allelica, il ri-sequenziamento dei genomi consente anche l'identificazione di geni presenti solo in alcune accessioni e non in altre consentendo di descrivere l'intero *pan-genoma* di una determinata specie. Con il termine *pan-genoma* si definisce l'intero set di geni propri di una specie, esso è costituito da geni comuni a tutte le accessioni della specie (*core genome*) e da geni presenti solo in alcune accessioni (*dispensable genome*) (Morgante *et al.*, 2007). Il confronto tra genomi di accessioni diverse di una stessa specie ha spesso evidenziato, oltre ad innumerevoli SNP e piccole inserzioni/delezioni (*indel*), la presenza/assenza di larghi tratti di DNA anche contenenti interi geni espressi e variazioni nel numero di copie di un determinato gene. Questi due tipi di polimorfismi,

1. Next Generation Breeding: le conoscenze genomiche rivoluzionano il miglioramento genetico

noti come PAV (*Presence Absence Variation*) e CNV (*Copy Number Variation*), rappresentano un ulteriore componente della diversità genetica capace di spiegare importanti differenze fenotipiche (Saxena *et al.*, 2014). La rilevanza di queste variazioni strutturali non è ancora ben definita, tuttavia diversi esempi suggeriscono che esse interessano una rilevante frazione del genoma. Ad esempio, il ri-sequenziamento di sole quattro varietà di uve da tavola ha evidenziato che ben l'8% del genoma di vite è coinvolto in variazioni strutturali con oltre 700 regioni contenenti oltre 2.000 geni soggette a CNV (Cardone *et al.*, 2016). Similmente, il ri-sequenziamento di sette accessioni di soia ha evidenziato come diversi geni coinvolti nell'adattamento all'ambiente (geni di resistenza a malattie, geni che controllano l'epoca di fioritura) o nella sintesi di metaboliti secondari costituiscono varianti strutturali del genoma presenti solo in alcune delle linee sequenziate (Li *et al.*, 2014). Infine, un recente lavoro realizzato ri-sequenziando dieci diverse accessioni di *Brassica oleracea* (colza) ha dimostrato come circa il 20% dei geni sia soggetto a PAV e tra essi diversi geni codificanti per importanti caratteri come la resistenza a malattie, l'epoca di fioritura e il metabolismo dei glucosinolati (Golicz *et al.*, 2016).

1.3 Un'illimitata disponibilità di marcatori molecolari

I marcatori molecolari sono noti da più di trenta anni, tuttavia la recente diffusione dei sequenziatori NGS ha determinato un cambio di passo. Prima della diffusione delle nuove tecnologie di sequenza una tipica mappa molecolare era rappresentata da alcune centinaia, raramente un migliaio, di marcatori; con l'utilizzo delle tecnologie NGS, diverse decine di migliaia di marcatori possono essere evidenziati con una sequenza parziale di pochi genomi (o trascrittom) e inclusi in qualsiasi mappa molecolare anche in specie con pochissime informazioni di sequenza disponibili a priori. In questo contesto gli SNP rappresentano i marcatori d'elezione in virtù della loro facile automazione attraverso piattaforme strumentali capaci di operare sia con singoli marcatori sia con gruppi di SNP (da poche unità a centinaia di migliaia).

Da un punto di vista tecnologico gli SNP sono evidenziati mediante l'uso di sonde oligonucleotidiche sequenza-specifiche. Queste sonde possono essere aggiunte alla miscela di reazione per il saggio di singoli polimorfismi oppure fissate su supporti solidi di pochi cm² (SNP array) capaci di portare sino a

centinaia di migliaia di sonde. L'uso di array di SNP consente di automatizzare l'analisi di un numero enorme di marcatori su vaste collezioni di germoplasma. Array di SNP sono disponibili per molte specie coltivate quali vite, mais, pomodoro, pesco, soia, orzo, riso, frumento e melo (Ganal *et al.*, 2012; Voss-Fels e Snowdon, 2016).

Lo sviluppo di array di SNP richiede un significativo investimento di risorse per la ricerca e la validazione dei singoli marcatori e per la produzione dell'array, inoltre il particolare set di SNP incluso in un array potrebbe non cogliere appieno la diversità genetica presente in una data collezione di germoplasma. In questi casi le tecnologie NGS offrono la possibilità di procedere ad un sequenziamento parziale con basse coperture tale da consentire l'identificazione dei polimorfismi tramite il confronto delle sequenze dei genomi dei genotipi di specifico interesse. Queste tecniche sono basate sul sequenziamento di una frazione del genoma ottenuta tramite digestione enzimatica e successiva amplificazione con adattatori selettivi. Il fatto che questi metodi non richiedono una conoscenza a priori dei polimorfismi li rende utili per l'analisi genetica anche in specie dove nessuna sequenza di riferimento è disponibile. Tra le tecnologie di questo tipo le più diffuse sono *Restriction-site Associated DNA sequencing* (RAD) (Miller *et al.*, 2007) e *Genotyping by Sequencing* (GBS) (Elshire *et al.*, 2011).

1.4 Associazioni carattere-marcatore in collezioni di germoplasma

Le basi genetiche di molti caratteri sono state identificate tramite procedure convenzionali come analisi di linkage in popolazioni segreganti (Doppio Aploidi – DHs, linee inbred ricombinanti – RILs, linee di introgressione – ILs) oppure utilizzando linee quasi isogeniche (NILs) sviluppate mediante ripetuti cicli di re incrocio (Scheda 1.2).

Le basi genetiche identificate con questo tipo di materiali sono spesso specifiche per i genotipi studiati e per pochi altri geneticamente simili a quelli analizzati, ed il limitato numero di eventi di ricombinazione che sono utilizzati per generare le popolazioni segreganti (uno solo nel caso DH o RIL) limitano la possibilità di identificare marcatori strettamente associati ai geni di interesse. Nell'ultima decade, la disponibilità di piattaforme di genotipizzazione ad alta risoluzione ha aperto la strada a studi di *Genome-Wide Association Scan* (GWAS) (Tondelli *et al.*, 2013; Biscarini *et al.*, 2016) (Scheda 1.3).

SCHEDA 1.2

POPOLAZIONI SPERIMENTALI

Doppio-aploidi (DH): piante ottenute a seguito di diploidizzazione spontanea o indotta tramite trattamento con colchicina a partire da gameti della generazione F₂, tipicamente il polline delle antere ottenute da ibridi F₁. Ciascuna pianta DH risulta omozigote a tutti i loci, mentre una popolazione DH rappresenta nell'insieme la segregazione dei caratteri tipica di una F₂.

Linee inbred ricombinanti (RIL): rappresenta una popolazione segregante costituita da linee omozigoti nelle quali gli eventi di ricombinazione tra i cromosomi ereditati da due linee omozigoti sono conservati. Le RILs sono generate tramite incrocio di due parentali seguito dalla generazione delle piante F₂ e quindi da diverse generazioni di autofecondazione per consentire il raggiungimento dell'omozigosi.

Linee di introgressione (IL): rappresenta una popolazione di linee derivate dall'incrocio tra una singola pianta (in genere una varietà o linea di elite) e, solitamente, una variante selvatica della stessa specie. L'incrocio iniziale è seguito da reincroci (in genere 2 o 3) con la forma coltivata e da generazioni di autofecondazione per consentire il raggiungimento dell'omozigosi. Una popolazione di ILs contiene le linee che hanno incorporato frammenti di DNA che sommati tra loro rappresentano l'intero genoma della specie selvatica donatrice.

Linee quasi isogeniche (NIL): sono coppie di linee che differiscono tra di loro unicamente per un singolo tratto genomico implicato nella determinazione di uno specifico carattere. Sono generate dall'incrocio iniziale tra un parentale ricorrente e la linea che possiede il carattere da trasferire (parentale donatore). L'incrocio iniziale è seguito da reincroci con il parentale ricorrente affiancati dalla selezione per il carattere che si intende trasferire (*forward selection*) e talora dalla selezione per il genoma del ricorrente (*background selection*). Dopo alcune generazioni di reincrocio le linee ottenute saranno geneticamente quasi identiche al ricorrente (quasi isogeniche) con la eccezione del carattere trasferito dal donatore.

Multi-parent Advanced Generation Inter-Cross (MAGIC): rappresenta un tipo di popolazione multiparentale sviluppata dall'incrocio di linee fondatrici diverse, il cui numero varia da quattro a otto, a cui seguono diversi cicli di inter-incrocio e di autofecondazioni per consentire il raggiungimento dell'omozigosi. Tali popolazioni sono generate per incrementare la precisione e la risoluzione di mappaggio di QTL, ottenibile grazie al maggior numero di alleli ed eventi di ricombinazione rispetto alle popolazioni di mappaggio bi-parentali.

Nested association mapping (NAM): rappresenta una popolazione multiparentale in cui una varietà o linea di riferimento è incrociata con diverse altre varietà o linee e le F₁ ottenute sono autofecondate per diverse generazioni fino all'ottenimento di una popolazione di RILs omozigoti.

L'approccio GWAS è basato sul *linkage disequilibrium* (LD) tra marcatori e geni coinvolti nell'espressione di specifici caratteri valutato su grandi collezioni di germoplasma. GWAS sfrutta gli eventi di ricombinazione accumulati durante l'evoluzione e, quando l'analisi è effettuata con un'elevata copertura di marcatori uniformemente distribuiti sul genoma, permette di incrementare la risoluzione genetica dell'associazione tra marcatore e carattere, talvol-

ta sino a livelli tali da consentire l'identificazione diretta della variazione di sequenza che ha causato un determinato fenotipo (Ogura e Busch, 2015).

La risoluzione di un'analisi GWAS dipende da diversi fattori:

1. il grado di LD, e quindi il grado di risoluzione nel determinare l'associazione tra marcatore e carattere, varia tra le diverse specie (le specie autoga-

ASSOCIAZIONE TRA CARATTERI FENOTIPICI E MARCATORE MOLECOLARE

SCHEDA 1.3

Analisi di linkage: consente di definire il livello di concatenazione tra geni, tra geni e marcatori e tra marcatori. Il test statistico del valore di LOD (*Logarithm Of Odds*) è utilizzato per analizzare i rapporti di concatenazione (linkage) in popolazioni tramite il confronto della probabilità di riscontrare i valori osservati se i due loci sono effettivamente sullo stesso cromosoma (con una determinata frequenza di ricombinazione) rispetto alla probabilità di ottenere gli stessi risultati se i due loci non sono in linkage, quindi esclusivamente per caso.

Studi di associazione genome-wide (GWAS): questi studi utilizzano collezioni di linee, varietà, non imparentate che sono genotipizzate e fenotipizzate per caratteri di interesse. Associazioni statisticamente significative (in *linkage disequilibrium*) tra polimorfismi del DNA (marcatori) e variazione dei caratteri identificano regioni genomiche nelle quali sono localizzati geni che governano i caratteri in studio.

Linkage Disequilibrium (LD): indica la presenza di associazione non casuale, statisticamente significativa, tra geni, marcatori, geni e marcatori, che costituiscono di solito un particolare aplotipo ancestrale, diffuso nella popolazione in cui è rilevato perché trasmesso lungo la discendenza da un comune progenitore. LD è comunemente valutato calcolando i valori di correlazione (r^2) tra i marcatori in intervalli che includono un numero variabile di marcatori in modo da identificare un r^2 medio tra tutti i marcatori ottenuto come funzione delle distanze fisiche tra i marcatori. Il decadimento del LD determina la risoluzione attesa nelle analisi GWAS, minore è la distanza a cui LD decade maggiore è la risoluzione e viceversa, a patto di avere sufficienti marcatori per coprire uniformemente il genoma con una distanza media analoga al valore di LD.

Struttura di popolazione: presenza di sottogruppi (denominati sub-popolazioni) nella popolazione in analisi causata, nella maggioranza dei casi, dall'aver intrapreso diversi percorsi evolutivi dal periodo di addomesticazione o dall'essersi originati da eventi di addomesticazione distinti. Un limite dell'analisi GWAS è dovuto a situazioni in cui l'organizzazione genetica della popolazione (struttura) è strettamente collegata alla distribuzione della sua variabilità fenotipica. In questi casi si possono avere associazioni spurie, o nel caso di correzioni applicate all'analisi GWAS per la struttura della popolazione, condurre alla eliminazione di associazioni positive collegate alla struttura della popolazione, creando false associazioni negative.

me hanno generalmente valori superiori di LD rispetto alle specie allogame), ma anche tra diverse popolazioni di una stessa specie e tra diverse regioni dello stesso genoma;

2. la struttura della popolazione che può condurre a false associazioni;
3. l'effetto di alleli rari (anche in popolazioni di rilevanti dimensioni) che potrebbe non essere rilevato dall'analisi GWAS se la frequenza di tali alleli è inferiore a determinate soglie (generalmente < 5%).

La capacità di rilevare associazioni significative tra marcatore e carattere dipende inoltre dalla qualità dei dati fenotipici, dalla dimensione della popola-

zione, dalla base genetica ed ereditabilità del carattere (Korte e Farlow, 2013).

Nuovi schemi d'incrocio e nuovi tipi di popolazioni sperimentali sono state proposte per superare alcune delle limitazioni descritte (es. struttura della popolazione) e per incrementare la potenza statistica e la risoluzione dell'associazione tra marcatore e carattere e quindi facilitare la identificazione dei geni responsabili della variazione fenotipica (Gupta et al., 2013). Queste includono le popolazioni *Multi-parent Advanced Generation Intercross* (MAGIC) e le popolazioni *Nested Association Mapping* (NAM). Le prime sono create tramite l'inter-incrocio di diverse linee parentali (in genere otto o più) e successive autofecondazioni delle progenie per generare RILs.

Rispetto alle tipiche popolazioni bi-parentali, l'utilizzo di più accessioni parentali negli incroci iniziali consente di catturare una maggior quota di diversità allelica, mentre i cicli multipli d'inter-incrocio rendono possibili maggiori ricombinazioni, migliorando di conseguenza la precisione nella localizzazione di QTLs. Nelle popolazioni NAM, diverse linee fondatrici sono incrociate con la medesima linea di riferimento per sviluppare progenie *half-sib* geneticamente correlate. Il vantaggio di questa popolazione è rappresentato dalla possibilità di incorporare un elevato numero di alleli del pool genico di una determinata specie.

La disponibilità di grandi set di marcatori molecolari con cui genotipizzare a basso costo estese collezioni di germoplasma ben caratterizzate dal punto di vista fenotipico e la possibilità di ricercare le associazioni tra marcatori e i loci responsabili dei caratteri descritti sta generando una grande mole di informazioni genetiche che se opportunamente analizzate consentono di predire le performance delle piante sulla base dei soli dati genotipici aprendo la strada alla selezione genomica.

1.5 Predizione del *breeding value* e selezione genomica

Ci sono due principali strade per supportare il miglioramento genetico con l'uso di marcatori molecolari: utilizzare marcatori localizzati in vicinanza o entro i loci causali degli effetti fenotipici (MAS) o sfruttare tutti i marcatori disponibili come predittori del valore di *breeding* (GS). La MAS è utilizzata per la selezione di un numero relativamente piccolo di geni che hanno effetti rilevanti sul fenotipo (Xu e Crouch, 2008). Tuttavia, il successo delle nuove varietà non dipende solo da pochi loci con grande effetto fenotipico, infatti, il ruolo di particolari combinazioni di loci aventi ognuno piccoli effetti sul fenotipo è spesso superiore a quello dei pochi geni selezionabili tramite MAS. GS consente di valutare l'effetto complessivo di questi loci nel processo di selezione. GS calcola il valore genetico degli individui, il quale, a sua volta, stima l'attitudine degli individui a generare progenie con un fenotipo superiore. Marcatori distribuiti sull'intero genoma sono utilizzati per calcolare indici genomici (*Genomic Estimated Breeding Value*, GEBV) che stimano l'effetto dei singoli marcatori e delle diverse combinazioni aplo-tipiche sui caratteri d'interesse. La definizione dei valori di GEBV è realizzata attraverso la comparazione di informazioni genotipiche e fenotipiche in una collezione di germoplasma (*training population*) che rap-

presenta il pool genico utilizzato nei programmi di miglioramento genetico. Successivamente i valori di GEBV sono utilizzati per selezionare popolazioni di piante (*breeding population*) genotipizzate con i marcatori molecolari ma non sottoposte alla valutazione fenotipica (Desta e Ortiz, 2014). Rispetto alla tradizionale selezione fenotipica, GS permette di stimare in anticipo il *breeding value* degli individui senza attendere la progenie della generazione successiva riducendo la necessità di estese valutazioni in campo e, a parità d'intensità di selezione, permette di realizzare lo stesso o un maggiore progresso genetico per unità di tempo ad un costo inferiore (Jonas e de Koning, 2013). Una volta che i dati fenotipici della popolazione del programma di breeding sono disponibili, questi possono essere usati per reiterare il processo sviluppando uno schema di selezione ricorrente (Fig. 1.2) che permette una maggiore correlazione tra il vero valore di *breeding* e il GEBV (definita come *prediction accuracy*) ed un progresso genetico più elevato di quello ottenibile con altri schemi di selezione (Bassi *et al.*, 2016).

Diversi modelli statistici basati prevalentemente su regressione lineare corretta, *Best Linear Unbiased Prediction* (BLUP), e metodi di regressione Bayesiani (Desta e Ortiz, 2014) sono stati proposti per determinare il GEBV. L'accuratezza di predizione dei diversi metodi è dibattuta; mentre in alcuni casi tutti i modelli forniscono simili accuratezze nelle stime (Gouy *et al.*, 2013), altri studi evidenziano come diverse caratteristiche delle popolazioni (struttura del LD, presenza di epistasi e grado di relazione fra *training* e *breeding population*) come anche la base genetica e l'ereditabilità dei caratteri possono influenzare la performance dei metodi di predizione (Desta e Ortiz, 2014). Il livello di LD rappresenta un parametro di rilevante importanza nel disegnare l'approccio di GS. LD ha dimensioni ridotte nelle piante allogame (es. LD in mais 0,1-1,5 kb (Tenailon *et al.*, 2001), mentre si estende considerevolmente in specie autogame come riso (da 75 kb nel *background* della sub-specie *indica* a 500 kb nella sub-specie *japonica* temperati) o l'orzo dove il decadimento del LD avviene dopo circa 5-10 cM, che rappresentano approssimativamente 20-40 Mb (Pasam *et al.*, 2012). In specie con modesti valori di LD la dimensione della *training population* può essere minore ma servono molti più marcatori per coprire in modo efficace l'intero genoma. Anche le dimensioni delle popolazioni di *training* e di *breeding* rappresentano parametri critici. In generale, un maggiore rapporto *training vs. breeding population* è richiesto per predizioni accurate del GEBV nel caso di bassa ereditabilità dei caratteri e della presenza di un numero elevato di QTLs che controllano i

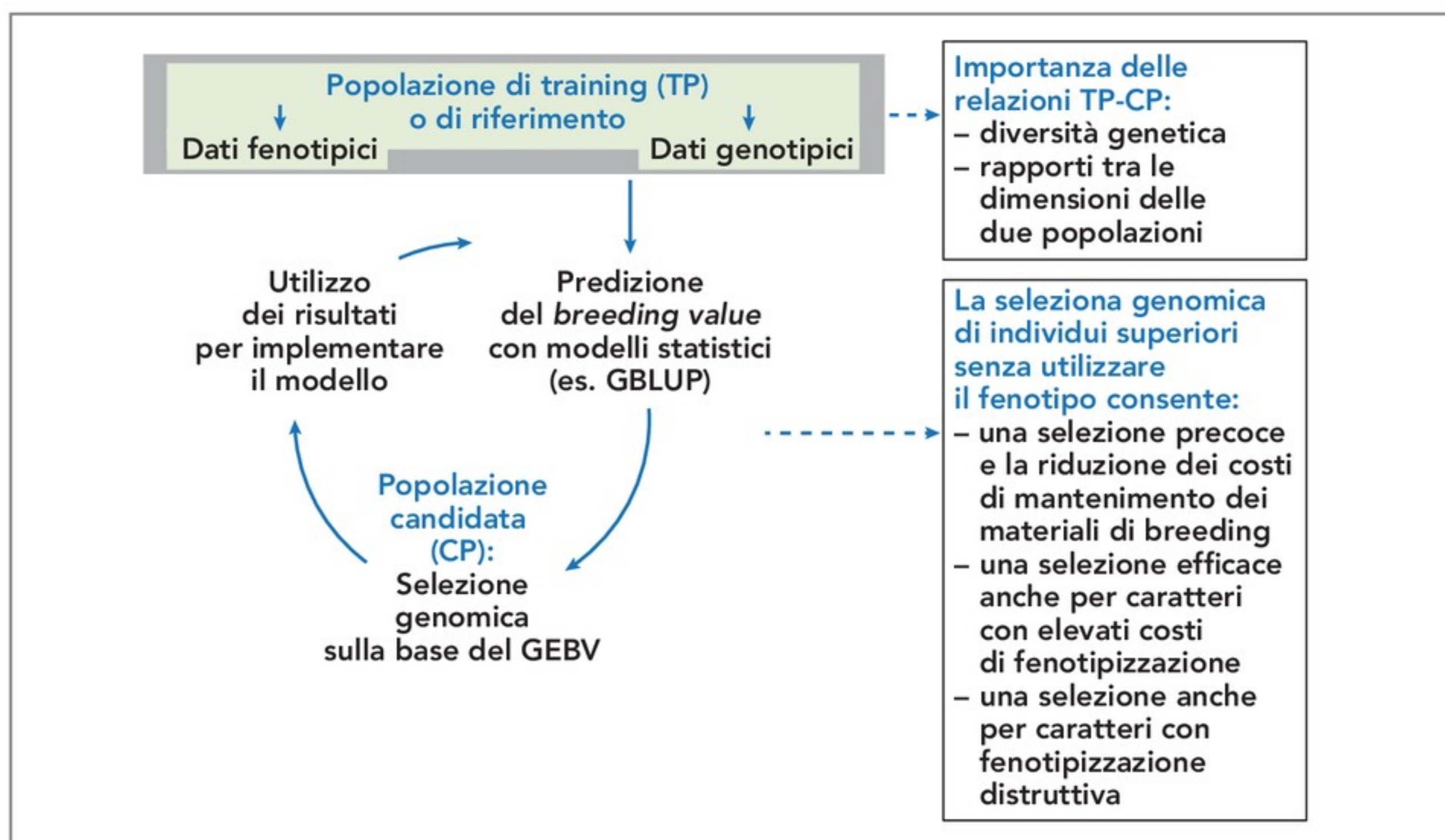


Figura 1.2 - Schema di selezione genomica.

caratteri oggetto di GS (Fig. 1.2) (Bassi *et al.*, 2016). Marcatori co-dominanti (es. SNPs, SSRs) forniscono una stima più accurata del GEBV rispetto a quelli dominanti (es. marcatori DArT) grazie alla maggiore capacità di rilevazione del LD e l'accuratezza può essere ulteriormente migliorata considerando gli aplotipi invece dei singoli marcatori (Li *et al.*, 2007). GS è stata applicata a diversi caratteri in diverse specie che includono mais, orzo, frumento tenero e riso. In prima analisi, i lavori in cui tale approccio è stato adottato suggeriscono un'ampia applicabilità della GS anche in specie con genomi complessi come frumento tenero e mais. In diversi studi GS ha fornito informazioni affidabili a livelli superiori rispetto agli approcci di selezione più tradizionali con valori di *prediction accuracy* sino a 0,85 anche per caratteri poligenici a bassa ereditabilità (Heffner *et al.*, 2009).

1.6 Next generation breeding

Gli attuali programmi di *breeding* integrano tra loro la selezione fenotipica basata sui tradizionali metodi di *breeding* (*pedigree*, *backcross*, valutazione dell'attitudine alla combinazione e sviluppo di ibridi) con le conoscenze molecolari attraverso l'uso di marcatori molecolari e, nei paesi ove consentito, della transgenesi, tuttavia in entrambi i casi la compo-

nente molecolare è limitata a pochi geni con grande effetto fenotipico (frequentemente geni di resistenza a malattie).

Lo sviluppo tecnologico dei sistemi NGS e la possibilità di accedervi a costi estremamente contenuti ha reso il sequenziamento del DNA e le sue innumerevoli applicazioni alla portata di tutti quelli che si occupano di miglioramento genetico. In figura 1.3 sono riassunti in maniera schematica le principali opportunità offerte al miglioramento genetico dalla disponibilità di vaste conoscenze genomiche. I *breeder* interessati a specifici caratteri potranno sviluppare le idonee popolazioni sperimentali e poi applicare ad esse la tecnologia NGS più avanzata ed identificare i marcatori associati a geni che influenzano caratteri oggetto del *breeding* direttamente nelle linee avanzate di loro interesse a costi sempre più contenuti. Risorse, disponibili anche *on line*, consentiranno di accedere a informazioni relative alla variabilità allelica per geni che influenzano caratteri di interesse agronomico.

L'esteso utilizzo di marcatori molecolari sposterà sempre più la selezione dalla dimensione fenotipica a quella genotipica tramite il ricorso alla selezione genomica, inoltre le conoscenze derivanti dal sequenziamento dei genomi combinate alle tecniche di *genome editing* permetteranno di ampliare la diversità genetica utilizzabile ai fini della selezione. Se da un lato, almeno per alcuni anni, le tecniche

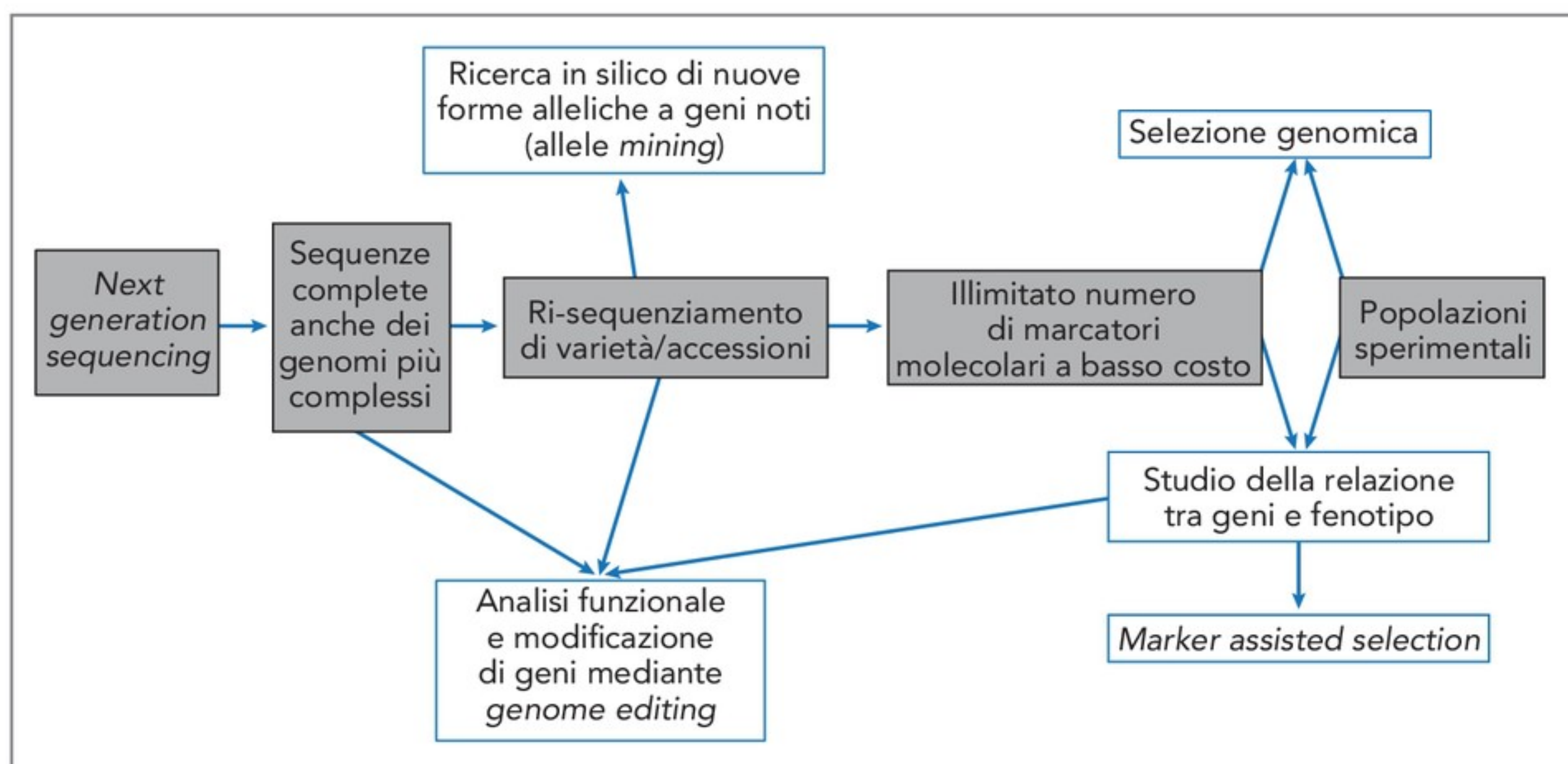


Figura 1.3 - Principali opportunità offerte dalla tecnologia NGS al miglioramento genetico.

di *genome editing* rimarranno confinate nei laboratori di ricerca, dall'altro, tutte le applicazioni basate su marcatori molecolari e ri-sequenziamento stanno diventando di uso comune nel miglioramento genetico, anche nelle piccole industrie sementiere. In questo contesto le conoscenze bioinformatiche necessarie per leggere, interpretare e sfruttare i dati genomici diventeranno un elemento essenziale. La GS ha una grande potenzialità nel miglioramento genetico vegetale, ma richiede conoscenze statistiche e bioinformatiche elevate, allo stesso modo la disponibilità dei genomi di riferimento e dei dati di ri-sequenziamento che descrivono il pan-genoma di una specie permette di identificare *in silico* nuovi alleli o nuovi marcatori associati a specifici caratteri, ma servono conoscenze bioinformatiche e biologiche rispettivamente per leggere ed interpretare informazioni contenute nel genoma.

La diversità genetica è da sempre, e rimarrà anche in futuro, la base da cui muovere per la selezione delle nuove varietà, tuttavia il miglioramento genetico sarà sempre più un'attività ad alto contenuto tecnologico e nei prossimi anni la genomica entrerà in modo preponderante nei campi di selezione varietale, trasformando il *breeding* tradizionale in *Next Generation Breeding* (Barabaschi *et al.*, 2016). Questa rivoluzione pone al miglioratore genetico pubblico e privato un'importante sfida di aggiornamento culturale e tecnologico che, nel caso del privato, rischia di mettere fuori mercato le imprese sementiere più piccole o meno propense all'innovazione. È pertanto necessario che conoscenze quali genomica, miglioramento genetico, statistica e bioinformati-

ca vengano integrate in nuove figure professionali per dare piena attuazione alle potenzialità del *Next Generation Breeding*.

Bibliografia

- BARABASCHI D., GUERRA D., LACRIMA K., LAINO P., MICHELOTTI V., URSO S., VALÈ G., CATTIVELLI L. (2011) – Emerging knowledge from genome sequencing of crop species. *Mol. Biotech.* 50: 250-266.
- BARABASCHI D., TONDELLI A., DESIDERIO F., VOLANTE A., VACCINO P., VALÈ G. AND CATTIVELLI L. (2016) – *Next generation breeding*. *Plant Sci.* 242: 3-13.
- BASSI F.M., BENTLEY A.R., CHARMET G., ORTIZ R., CROSSA J. (2016) – Breeding schemes for the implementation of genomic selection in wheat (*Triticum* spp.). *Plant Sci.* 242: 23-36.
- BHULLAR N.K., ZHANG Z., WICKER T., KELLER B. (2010) – Wheat gene bank accessions as a source of new alleles of the powdery mildew resistance gene Pm3: a large scale allele mining project. *BMC Plant Biol.* 10: 88.
- BISCARINI F., COZZI P., CASELLA L., RICCARDI P., VATTARI A., ORASEN G., PERRINI R., TACCONI G., TONDELLI A., BISELLI C., CATTIVELLI L., SPINDEL J., MCCOUCH S., ABBRUSCATO P., VALÈ G., PIFFANELLI P., GRECO R. (2016) – Genome-wide association study for traits related to plant and grain morphology, and root architecture in temperate rice accessions. *PLoS One* 11: e0155425.
- BISELLI C., URSO S., TACCONI G., STEUERNAGEL B., SCHULTE D., GIANINETTI A., BAGNARESI P., STEIN N., CATTIVELLI L., VALÈ G. (2013) – Haplotype variability and identifica-

1. Next Generation Breeding: le conoscenze genomiche rivoluzionano il miglioramento genetico

tion of new functional alleles at the Rdg2a leaf stripe resistance gene locus. *Theor. Appl. Genet.* 126: 1575-1586.

CARDONE M.F., D'ADDABBO P., ALKAN C., BERGAMINI C., CATACCHIO C.R., ANACLERIO F., CHIATANTE G., MARRA A., GIANNUZZI G., PERNIOLA R., VENTURA R., ANTONACCI D. (2016) – Inter-varietal structural variation in grapevine genomes. *Plant J.* 88: 648-661.

DESTA Z.A., ORTIZ R. (2014) – Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci.* 19: 592-601.

ELSHIRE R.J., GLAUBITZ J.C., SUN Q., POLAND J.A., KAWAMOTO K., BUCKLER E.S., MITCHELL S.E. (2011) – A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6: e19379.

GANAL M.W., POLLEY A., GRANER E.-M., PLIESKE J., WIESEKE R., LUERSEN H., DURSTEWITZ G. (2012) – Large SNP arrays for genotyping in crop plants. *J. Bioscience* 37: 821-828.

GOLICZ A.A., BAYER P.E., BARKER G.C., EDGER P.P., KIM H., MARTINEZ P.A., CHAN C.K.K., SEVERN-ELLIS A., MCCOMBIE W.R., PARKIN I.A.P., PATERSON A.H., PIRES J.C., SHARPE A.G., TANG H., TEAKLE G.R., TOWN C.D., BATLEY J., EDWARDS D. (2016) – The pangenome of an agronomically important crop plant *Brassica oleracea*. *Nature Com.* 7:13390.

GOUY M., ROUSSELLE Y., BASTIANELLI D., LECOMTE P., BONNAL L., ROQUES D., EFILÉ J.C., ROCHER S., DAUGROIS J., TOUBI L., NABENEZA S., HERVOUET C., TELISMART H., DENIS M., THONG-CHANE A., GLASZMANN J.C., HOARAU J.Y., NIBOUCHE S., COSTET L. (2013) – Experimental assessment of the accuracy of genomic selection in sugarcane. *Theor. Appl. Genet.* 126: 2575-2586.

GUPTA P.K., KULWAL P.L., MIR R.R. (2013) – “QTL mapping: methodology and applications in cereal breeding”. In: P.K. Gupta, R.K. Varshney (Eds.), *Cereal Genomics II*, Springer, Heidelberg, 275-318.

HEFFNER E.L., SORRELLS M.E., JANNINK J.L. (2009) – Genomic selection for crop improvement. *Crop Sci.* 49: 1-12.

HENRY I.M., NAGALAKSHMI U., LIEBERMAN M.C., NGO K.J., KRASILEVA K.V., VASQUEZ-GROSS H., AKHUNOVA A., AKHUNOV E., DUBCOVSKY J., TAI T.H., COMAI L. (2014) – Efficient genome-wide detection and cataloging of EMS-induced mutations using exome capture and next-generation sequencing. *Plant Cell* 26: 1382-1397.

HUANG X., KURATA N., WEI X., WANG Z.X., WANG A., ZHAO Q., ZHAO Y., LIU K., LU H., LI W., GUO Y., LU Y., ZHOU C., FAN D., WENG Q., ZHU C., HUANG T., ZHANG L., WANG Y., FENG L., FURUUMI H., KUBO T., MIYABAYASHI T., YUAN X., XU Q., DONG G., ZHAN Q., LI C., FUJIYAMA A., TOYODA A., LU T., FENG Q., QIAN Q., LI J., HA B. (2012) – A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490: 497-501.

KRASILEVA K.V., VASQUEZ-GROSS H.A., HOWELL T., BAILEY P., PARAISO F., CLISSOLD L., SIMMONDS J., RAMIREZ-GONZALEZ R.H., WANG X., BORRILL P., FOSKER C., AYLING S., PHILLIPS A.L., UAUY C., DUBCOVSKY J. (2017) – Uncovering hidden variation in polyploid wheat. *Proc. Natl. Acad. Sci. USA* 114: E913-E921.

JONAS E., DE KONING D.J. (2013) – Does genomic selection have a future in plant breeding?. *Trends in Biotech.* 31(9): 497-504.

KIM S.I., TAI H.T. (2014) – Identification of novel rice low phytic acid mutations via TILLING by sequencing. *Mol. Breed.* 34: 1717-1729.

KORTE A., FARLOW A. (2013) – The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9: 29.

LI Y., LI Y., WU S., HAN K., WANG Z., HOU W., ZENG Y., WU R. (2007) – Estimation of multilocus linkage disequilibria in diploid populations with dominant markers. *Genetics* 176: 1811-1821.

LI J.Y., ZEIGLER WANG R.S. (2014) – The 3000 rice genomes project: new opportunities and challenges for future rice research. *GigaScience* 3: 8.

LI Y.H., ZHOU G., JIANG J.W., JIN L.G., ZHANG Z., GUO Y., ZHANG J., SUI Y., ZHENG L., ZHANG S.S., ZUO Q., SHI X.H., LI Y.F., ZHANG W.K., HU Y., KONG G., HONG H.L., TAN B., SONG J., LIU Z.X., WANG Y., RUAN H., YEUNG C.K., LIU J., WANG H., ZHANG L.J., GUAN R.X., WANG K.J., LI W.B., CHEN S.Y., CHANG R.Z., JIANG Z., JACKSON S.A., LI R., QIU L.J. (2014) – De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. *Nat. Biotechnol.* 32: 1045-1052.

MIDDLETON C.P., SENERCHIA N., STEIN N., AKHUNOV E.D., KELLER B., WICKER T., KILIAN B. (2014) – Sequencing of chloroplast genomes from wheat, barley, rye and their relatives provides a detailed insight into the evolution of the *Triticeae* tribe. *PLoS One* 9: e85761.

MILLER M.R., DUNHAM J.P., AMORES A., CRESKO W.A., JOHNSON E.A. (2007) – Rapid and cost effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res.* 17: 240-248.

MORGANTE M., DE PAOLI E., RADOVIC S. (2007) – Transposable elements and the plant pan-genomes. *Curr. Opin. Plant Biol.* 10: 149-155.

OGURA T., BUSCH W. (2015) – From phenotypes to causal sequences: using genome wide association studies to dissect the sequence basis for variation of plant development. *Curr. Opin. Plant Biol.* 23: 98-108.

PASAM R.K., SHARMA R., MALOSETTI M., VAN EEUWIJK F.A., HASENEYER G., KILIAN B., GRANER A. (2012) – Genome-wide association studies for agronomical traits in a world wide spring barley collection. *BMC Plant Biol.* 12: 16.

- SALSE J. (2012) – In silico archeogenomics unveils modern plant genome organization, regulation and evolution. *Current Opin. Plant Biol.* 15: 122-130.
- SAXENA R.K., EDWARDS D., VARSHNEY R.K. (2014) – Structural variations in plant genomes. *Brief. Funct. Genomics* 13: 296-307.
- SCHNABLE P.S., SPRINGER N.M. (2013) – Progress toward understanding heterosis in crop plants *Ann. Rev. Plant Biol.* 64: 71-88.
- SLADE A.J., FUERSTENBERG S.I., LOEFFLER D., STEINE M.N., FACCIOTTI D. (2005) – A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nat. Biotechnol.* 23: 75-81.
- TENAILLON M.I., SAWKINS M.C., LONG A.D., GAUT R.L., DOEBLEY J.F., GAUT B.S. (2001) – Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Proc. Natl. Acad. Sci. USA* 98: 9161-9166.
- TONDELLI A., XU X., MORAGUES M., SHARMA R., SCHNAITHMANN F., INGVARSDEN C., MANNINEN O., COMADRAN J., RUSSELL J., WAUGH R., SCHULMAN A., PILLEN K., RASMUSSEN S., KILIAN B., CATTIVELLI L., THOMAS W., FLAVELL A.J. (2013) – Structural and temporal variation in the genetic diversity of a European collection of barley cultivars and utility for association mapping of quantitative traits. *Plant Genome* 6: 2.
- TSAI H., HOWELL T., NITCHER R., MISSIRIAN V., WATSON B., NGO K.J., LIEBERMAN M., FASS J., UAUY C., TRAN R.K., KHAN A.A., FILKOV V., TAI T.H., DUBCOVSKY J., COMAI L. (2011) – Discovery of Rare Mutations in Populations: TILLING by Sequencing. *Plant Physiol.* 156: 1257-1268.
- VOSS-FELS K., SNOWDON R.J. (2016) – Understanding and utilizing crop genome diversity via high-resolution genotyping. *Plant Biotech. J.* 14: 1086-1094.
- WANG M., ALLEFS S., VAN DEN BERG R.G., VLEESHOUWERS V.G., VAN DER VOSSEN E.A., VOSMAN B. (2008) – Allele mining in *Solanum*: conserved homologues of Rpi-blb1 are identified in *Solanum stoloniferum*. *Theor. Appl. Genet.* 116: 933-943.
- WANG D., GUO C., HUANG J., YANG S., TIAN D., ZHANG X. (2014) – Allele-mining of rice blast resistance genes at AC134922 locus. *Biochem. Biophys. Res. Com.* 446: (2014) 1085-1090.
- WENDEL J.F., JACKSON S.A., MEYERS B.C., WING R.A. (2016) – Evolution of plant genome architecture. *Genome Biology* 17:37.
- XU Y., CROUCH J.H. (2008) – Marker-assisted selection in plant breeding: from publications to practice. *Crop Sci.* 48: 391-407.
- XU X., LIU X., GE S., JENSEN J.D., HU F., LI X., DONG Y., GUTENKUNST R.N., FANG L., HUANG L., LI J., HE W., ZHANG G., ZHENG X., ZHANG F., LI Y., YU C., KRISTIANSEN K., ZHANG X., WANG J., WRIGHT M., MCCOUCH S., NIELSEN R., WANG J., WANG W. (2012) – Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat. Biotechnol.* 30: 105-111.

Sitografia

- <http://www.ncbi.nlm.nih.gov/genomes/leuks.cgi>
<http://plants.ensembl.org/index.html>
<http://phytozome.jgi.doe.gov/pz/portal.html>