Un team internazionale di ricercatori sta lavorando alla mappatura del genoma del girasole e uno dei contributi fondamentali per arrivare al traguardo è arrivato dall'Università di Pisa grazie a Lucia Natali, ricercatrice del Dipartimento di Scienze Agrarie, Alimentari e Agro-ambientali che ha coordinato il gruppo di genetica molecolare. Lo studio, i cui risultati sono stati recentemente pubblicati sulla rivista internazionale "BMC Genomics", è stato condotto in collaborazione con i ricercatori dell'Università della British Columbia di Vancouver, della Università del Colorado e dell'Istituto di Genomica Applicata di Udine e ha permesso di produrre un database di sequenze ripetute (SUNREP) che sarà utilizzato per decifrare la sequenza del genoma di girasole, una volta che questa sarà completata dal consorzio internazionale guidato dal professore Loren Rieseberg della Università canadese della British Columbia.

"Il girasole ha un genoma molto grande, stimato in 3,3 miliardi di nucleotidi poco più grande di quello dell'uomo – ha spiegato Lucia Natali – e le nostre analisi, hanno dimostrato che oltre l'80% del genoma di questa specie (oltre 2,5 miliardi di nucleotidi) è costituito da decine di migliaia di copie di sequenze che hanno una origine comune ai retrovirus e che si sono accumulate nel corso dell'evoluzione".

Il sequenziamento dei genomi delle piante coltivate è importante perché consente di decifrare i meccanismi che ne regolano lo sviluppo e la produttività. Per il girasole, i tratti genetici più utili da decifrare sono il contenuto in olio del seme, i meccanismi che regolano la fioritura e la dormienza dei semi, la produzione di biomassa e la resistenza a malattie e a carenza idrica.

In particolare il gruppo pisano, che ha condotto lo studio nell'ambito di un progetto di ricerca di interesse nazionale finanziato dal MIUR e coordinato dalla stessa Lucia Natali, ha utilizzato delle tecniche innovative di sequenziamento di DNA e di RNA ad elevato parallelismo (Next Generation Sequencing) che permettono di ottenere rapidamente un altissimo numero di sequenze a costi relativamente bassi e offrono un mezzo estremamente potente per affrontare l'analisi dei genomi e del loro funzionamento.