SHORT COMMUNICATION Bhartiya Krishi Anusandhan Patrika

कृषि जैव-विविधता के संरक्षण के लिए जीनोमिक क्षरण सूचकांकों की जांच हेतु कार्यप्रवाह

अनुपमा रॉय[1,2], लशिका मीना[1], नीलेश जोशी[3], मीर आसिफ इकबाल[1], दिनेश कुमार[1], सारिका जायसवाल[1]

सारांश

खाद्यान उत्पादन हेतु बोई गयी फसलों की प्रजातियों, किस्मों और पशुधन की नस्लों के गायब होने के साथ, बिना काटे गई प्रजातियों की एक विस्तृत श्रृंखला भी गायब हो रही है, औद्योगिक करण कृषि भूमि का विस्तारण विन्यास, अतिचारण के संबंध में नुकसान का पैमाना व्यापक है और पिछली शताब्दियों में प्रजातियों का अत्यधिक शोषण हुआ है। आनुवंशिक विविधता में कमी, आंतरिक प्रजनन में वृद्धि और जनसंख्या में हानिकारक उत्परिवर्तन के संचय द्वारा विशेषता 'जीनोमिक क्षरण' प्रक्रियाएँ। फिर भी, आधुनिक समय की आबादी के जीनोमिक क्षरण अनुमानों में घटती आबादी के आकार और कृषि जैव विविधता की संरक्षण स्थिति के साथ सामंजस्य का अभाव है। ऐतिहासिक नमूनों और आधुनिक नमूनों से संपूर्ण जीनोम पुनः अनुक्रमित डेटा को संसाधित करने के लिए अत्याधुनिक जैव सूचना विज्ञान उपकरणों और तकनीकों का उपयोग करके विभिन्न अवधि के नमूनों का उपयोग करके जीनोमिक क्षरण के पैटर्न की तुलना करने के लिए एक नई डिजाइन की गई अत्यधिक लचीली, स्केलेबल और एकमात्र कार्यप्रवाह। कई जीनोमिक क्षरण सूचकांकों के अनुमान तैयार करने के लिए जिनकी तुलना की जा सकती है। तदनुसार, विभिन्न किस्मों और नस्लों के विलुप्त होने के जोखिमों की पहचान की जा सकती है और कृषि जैव विविधता संरक्षण रणनीति योजनाएं तैयार की जा सकती हैं।

शब्दकुंजी: कृषि जैव विविधता, आनुवंशिक विविधता, अंतः प्रजनन, जीनोमिक क्षरण, पुनः अनुक्रमित डेटा, अत्याधुनिक।

Pipeline to Investigate Genomic Erosion Indices for the Conservation of Agrobiodiversity

Anupama Roy[1,2], Lashika Meena[1], Nilesh Joshi[3], Mir Asif Iquebal[1], Dinesh Kumar[1], Sarika Jaiswal[1] **10.18805/BKAP811**

ABSTRACT

With the disappearance of harvested species, varieties and breeds, a wide range of unharvested species is also disappearing, the scale of loss is extensive with respect to habitat loss, habitat configuration, overgrazing, and over exploitation of the species over the past centuries has led to 'genomic erosion' processes characterized by reduced genetic diversity, increased inbreeding and accumulation of harmful mutations in the population. Still, genomic erosion estimates of modern-day populations lack concordance with the declining population sizes and conservation status of agrobiodiversity. A newly designed highly flexible, scalable, and only pipeline to compare the patterns of genomic erosion using samples from different period using the state-of-the-art bioinformatics tools and techniques to process whole genome re-sequenced data from historical samples and modern samples in order to produce the estimates of several genomic erosion indices that can be compared. Accordingly, extinction risks for different varieties and breeds can be identified and agricultural biodiversity conservation strategy plans can be framed.

Key words: Agrobiodiversity, Genetic diversity, Inbreeding, Genomic erosion, Re-sequenced data, State-of-the-art.

कृषि जैव विविधता पर्यावरण, आनुवंशिक संसाधनों और प्रबंधन प्रणालियों और सहस्राब्दियों से सांस्कृतिक रूप से विविध लोगों द्वारा उपयोग की जाने वाली प्रथाओं के बीच बातचीत का परिणाम है, जिसमें किसान, चरवाहे और मछुआरे शामिल हैं, और इसलिए भूमि और जल संसाधनों का उपयोग विभिन्न तरीकों से उत्पादन के लिए किया जाता है। भोजन, औषधियों और आजीविका की सुरक्षा के लिए, मानव की जरूरतों को पूरा करने के लिए कृषि विविधता आवश्यक है। पृथ्वी पर जीवन की विविधता, इसके विकास और पर्यावरणीय परिवर्तनों के प्रति इसकी प्रतिक्रिया का अध्ययन विकासवादी और पारिस्थितिक

[1]Division of Agricultural Bioinformatics, ICAR-Indian Agricultural Statistics Research Institute, New Delhi-110 012, India.
[2]The Graduate School, ICAR-IARI, New Delhi-110 012, India.
[3]Pulse Chickpea Laboratory, Division of Genetics, ICAR-IARI, New Delhi-110 012, India.

Corresponding Author: Sarika Jaiswal, Division of Agricultural Bioinformatics, ICAR-IARI, New Delhi-110 012, India.
Email: sarika@icar.gov.in, aijaiswal@gmail.com

How to cite this article: Roy, A., Meena, L., Joshi, N., Iquebal, M.A., Kumar, D. and Jaiswal, S. (2025). Pipeline to Investigate Genomic Erosion Indices for the Conservation of Agrobiodiversity. *Bhartiya Krishi Anusandhan Patrika*. 1-5. doi: 10.18805/BKAP811.

Submitted: 26-10-2024 **Accepted:** 14-11-2025 **Online:** 09-12-2025

जनसंख्या आनुवंशिक अध्ययन द्वारा किया जा सकता है। कृषि जैव विविधता के लिए डेटा एकत्र करना और पारिस्थितिकी तंत्र की कार्यक्षमता को बनाए रखने में इसकी भूमिका को समझना एक बड़ी चुनौती है। अनुक्रमण प्रौद्योगिकियों में क्रांति के साथ, आणविक डेटा संग्रह इतना बढ़ गया है कि जीनोमिक्स, जहां अरबों न्यूक्लियोटाइड की एक साथ जांच की जाती है, अब जैविक टूलबॉक्स का एक महत्वपूर्ण हिस्सा है (Christiansen *et al.*, 2021)। भारतीय पालतू पशुओं में किए गए जीनोम–व्यापी अध्ययनों ने यह दर्शाया है कि स्थानीय नस्लों में आनुवंशिक विविधता लगातार परिवर्तित हो रही है, जो जीनोमिक क्षरण की बढ़ती चिंताओं की ओर संकेत करती है (Saravanan *et al.*, 2022)। पूरे जीनोम पुनःअनुक्रमण तथा SNP विश्लेषण जैसे अत्याधुनिक अनुक्रमण उपकरण अब भारतीय कृषि–प्रणालियों में आनुवंशिक विविधता की निगरानी और संरक्षण रणनीतियों को सुदृढ़ करने में सहायक बन रहे हैं (Mahalakshmi, 2022)। जनसंख्या के आकार में धीरे–धीरे कमी के कारण आनुवंशिक बहाव और अंतःप्रजनन के परिणामस्वरूप विलुप्त होने के जोखिमों में वृद्धि हुई है, जिसके परिणामस्वरूप आनुवंशिक विविधता का नुकसान हुआ है और फिटनेस में कमी आई है, इस प्रक्रिया को यहां आनुवंशिक क्षरण के रूप में जाना जाता है (Bijlsma and Loeschcke, 2012)। जीनोमिक क्षरण केवल लुप्तप्राय वनस्पतियों और जीवों तक ही सीमित नहीं है, कृषि विज्ञान के भाग्य को आकार देने में जीनोमिक क्षरण की भूमिका की पहचान करने के उद्देश्य से अध्ययन चलन में आया है।

जीनोमिक क्षरण सूचकांकों में परिवर्तन की दर का पता लगाने में दो सीमाओं का सामना करना पड़ रहा है। सबसे पहले, जीनोमिक क्षरण सूचकांकों और संरक्षण स्थिति के बीच कमजोर सहसंबंध, जो फसलों, पालतू जानवरों, मछलियों और पेड़ों के बीच खतरे की श्रेणियां स्थापित करने के लिए एक मानदंड के रूप में आनुवंशिक मापदंडों को शामिल करने की जटिलताओं को जन्म देता है। इस मुद्दे का एक प्रामाणिक और शक्तिशाली समाधान ऐतिहासिक नमूनों से जीनोमिक डेटा प्राप्त करना है जो हाल की जनसांख्यिकीय गिरावट से पहले का है। इस डेटा का उपयोग पूर्व-गिरावट आधार रेखाएं स्थापित करने और हाल ही में जैव विविधता में गिरावट के परिणामस्वरूप जीनोमिक क्षरण सूचकांकों में परिवर्तन की दर की प्रत्यक्ष मात्रा निर्धारित करने में सक्षम बनाने के लिए किया जा सकता है (Díez-Del-Molino *et al.*, 2018) ।

इसके अतिरिक्त, जीनोमिक अध्ययन और संरक्षण में अनुप्रयोगों के बीच अंतर को पाटने में जैव सूचना विज्ञान ज्ञान की कमी (Shafer *et al.*, 2015)। संरक्षण जीवविज्ञानियों के पास अक्सर बड़े पैमाने पर जीनोमिक डेटासेट और प्रोग्रामिंग कौशल को संसाधित करने में विशेषज्ञता का अभाव होता है। जीनोमिक क्षरण सूचकांक केवल तभी सार्थक और तुलनीय उपाय हो सकते हैं, यदि परिणाम प्रतिलिपि प्रस्तुत करने योग्य हों, क्योंकि अन्य वैज्ञानिक विषयों में इसकी कमी है। इस प्रकार, उपयोगकर्ता के अनुकूल सॉफ्टवेयर और प्रतिलिपि प्रस्तुत करने योग्य परिणाम उत्पन्न करने वाली संरचित पाइपलाइनों तक पहुंच में वृद्धि, संरक्षण के लिए जीनोमिक डेटा के अनुप्रयोग में महत्वपूर्ण भूमिका निभा सकती है (van Oosterhout, 2020)।

संरक्षण जीनोमिक्स के लिए जीनोमिक डेटा के उपयोगकर्ता–अनुकूल अनुप्रयोग पर केंद्रित एक अत्यधिक लचीली और मॉड्यूलर कार्यप्रवाह का उद्देश्य इन सीमाओं को संबोधित करना है। यह संपूर्ण–जीनोम पुनः अनुक्रमण डेटा का विष्लेशण करता है, जो जीनोम क्षरण के पैटर्न की जांच करने के लक्ष्य के साथ अस्थायी रूप से नमूना किए गए जीनोम के विष्लेशण को सक्षम बनाता है। यह कार्यप्रवाह जीनोमिक डेटा को संसाधित करने के लिए अत्याधुनिक तकनीकों का उपयोग करती है जो प्राचीन/ऐतिहासिक और आधुनिक दोनों नमूनों को एक साथ संभाल सकती है। यह भूगर्भिक समय पैमाने की दो समयावधियों की तुलना करने के लिए जीनोमिक क्षरण सूचकांक, जीनोम–व्यापक विविधता, इनब्रीडिंग और उत्परिवर्तनीय भार उत्पन्न करता है (Kutschera *et al.*, 2022)।

कार्यप्रवाह रूपरेखा

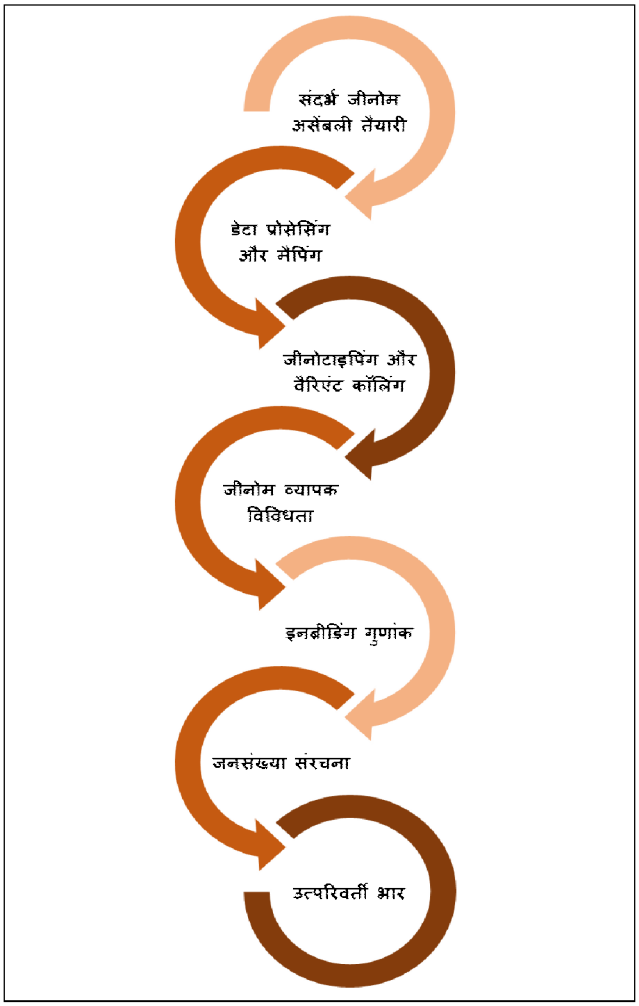
विभिन्न जीनोमिक क्षरण सूचकांकों का अनुमान लगाने के लिए कार्यप्रवाह को कई डाउनस्ट्रीम विश्लेषण निष्पादित करने के लिए संदर्भ जीनोम असेंबली और संपूर्ण जीनोम पुनरू अनुक्रमण डेटा की आवश्यकता होती है। पाइपलाइन को आकृति 1 में दर्शाया गया है।

संदर्भ आधारित जीनोम असेंबली

संदर्भ जीनोम असेंबली को BWA (Li and Durbin, 2009) और samtools (Li *et al.*, 2009) के साथ अनुक्रमित किया गया है। कार्यप्रवाह RepeatModeler (Flynn *et al.*, 2020) का उपयोग करके जीनोम के दोहराव वाले क्षेत्रों की नए सिरे से पहचान भी करती है RepeatMasker (Smit *et al.*, 2017) का उपयोग करके उन्हें मास्क करती है। गलत मानचित्रण के कारण होने वाले पूर्वाग्रहों से बचने के लिए, दोहराए जाने वाले क्षेत्रों को सभी डाउनस्ट्रीम चरणों से बाहर रखा जाएगा।

डेटा प्रसंस्करण और मैपिंग

fastp (Chen *et al.*, 2018) का उपयोग करके, कच्ची FASTQ फाइलों को ट्रिम किया जाता है। ऐतिहासिक नमूनों के लिए, fastp एक साथ रीड को ट्रिम करता है और ओवरलैपिंग



आकृति 1: विभिन्न जीनोमिक के आकलन के लिए कार्यप्रवाह की रूपरेखा।

पेयर–एंड रीड को मर्ज करता है। मर्ज किए गए रीड्स को BWA aln (Li and Durbin, 2009) का उपयोग करके संदर्भ जीनोम में मैप किया जाता है। आधुनिक नमूनों के लिए, FASTQ फाइलों को fastp के साथ ट्रिम किया जाता है और फिर BWA mem (Li, 2013) के साथ संदर्भ जीनोम असेंबली में मैप किया जाता है। Picard MarkDuplicates (https://broadinstitute.github.io/picard/) का उपयोग करके, प्रत्येक नमूने के बीच पी.सी.आर. डुप्लिकेट को चिह्नित किया जाता है।

जीनोटाइपिंग और वैरिएंट कॉलिंग

प्रत्येक नमूने में bcftools (Bonfield *et al.*, 2021) का उपयोग करके वेरिएंट को कॉल किया जाता है। यदि वांछित हो, तो एलीलिक असंतुलन और इनडेल्स फिल्टर लागू किया जाता है। मर्ज की गई वीसीएफ फाइल फिल्टरिंग के बाद उत्पन्न होती है जिसे ऐतिहासिक नमूनों की एक वी.सी.एफ. फाइल और आधुनिक नमूनों की एक वीसीएफ फाइल में विभाजित किया जाता है।

जीनोम व्यापक विविधता

जनसंख्या आनुवंशिकी में सबसे महत्वपूर्ण मानकों में से एक है $\theta = 4N_e\mu$, जहाँ Ne प्रभावी जनसंख्या आकार है और μ प्रति जीन प्रति पीढ़ी उत्परिवर्तन की दर है। यह नमूना लिए गए समुदाय के विकासात्मक इतिहास में एक झलक प्रदान करता है (Lazaridis *et al.*, 2014)। उसतेव एक सॉफ्टवेयर टूल (Haubold *et al.*, 2010) है जो जनसंख्या के डीएनए अनुक्रमों में देखी गई हेटेरोजीगोशिटी के आधार पर θ का अनुमान लगाता है। यह अनंत साइटों के मॉडल का उपयोग करता है, जो मानता है कि प्रत्येक उत्परिवर्तन डीएनए में एक अद्वितीय साइट पर होता है।

इनब्रीडिंग गुणांक

चरण 3 से उत्पन्न फिल्टर किए गए जीनोटाइप की वी. सी. एफ फाइल के साथ, कार्यप्रवाह 'plink - homozyg' के स्लाइडिंग विंडो दृष्टिकोण का उपयोग करके runs of homozygosity (ROH) की पहचान करता है। ऐसे आरओएच में आवंटित प्रत्येक नमूने के जीनोम के अंश को इनब्रीडिंग गुणांक (FROH) के रूप में परिभाषित किया जा सकता है, जिसका अनुमान और कल्पना कार्यप्रवाह रिपोर्ट में प्रदान किए गए प्लॉट में की जाती है।

जनसंख्या संरचना

चरण 3 से फिल्टर किए गए जीनोटाइप की वीसीएफ फाइल का उपयोग करके, सभी नमूनों के लिए और अकेले ऐतिहासिक और आधुनिक नमूनों के लिए पीसीए प्लिंक का उपयोग करके कार्यप्रवाह द्वारा उत्पन्न किया जाता है इस विष्लेशण का उपयोग आनुवंशिक जनसंख्या संरचना की जांच करने के साथ–साथ नमूनों के बीच आनुवंशिक आउटलेर्स या पूर्वाग्रहों की पहचान करने के लिए किया जाता है।

उत्परिवर्ती भार

उत्परिवर्तनीय भार किसी जनसंख्या में संचित हानिकारक उत्परिवर्तनों के परिणामस्वरूप होने वाला कुल आनुवंशिक बोझ है। कार्यप्रवाह उत्परिवर्तनीय भार का अनुमान लगाने के लिए दो पूरक तरीके प्रदान करती है।

पहली विधि में, GTF2 प्रारूप में जीनोम एनोटशन की आवश्यकता होती है, जिसके आधार पर SnpEff (Cingolani *et al.*, 2012) का उपयोग करके चरण 3 से अलग–अलग वेरिएंट के लिए एनोटेशन किया जाता है। एसएनपीईएफएफ

प्रति नमूने सभी पहचाने गए वेरिएंट को एनोटेट करने और प्रोटीन या अमीनो एसिड पर उनके कार्यात्मक प्रभावों की एक सरल भविष्यवाणी का अनुमान लगाने की अनुमति देता है जिसकी तुलना नमूनों या नमूनों के समूहों (यानी, ऐतिहासिक बनाम आधुनिक) के बीच की जा सकती है।

दूसरी विधि में जीनोमिक इवोल्यूशनरी रेट प्रोफाइलिंग (जी.ई.आर.पी.) स्कोर (Cooper *et al.*, 2005; Davydov *et al.*, 2010) शामिल है जो विकासवादी बाधित क्षेत्रों में व्युत्पन्न एलील्स की संख्या से प्रत्येक नमूने के सापेक्ष उत्परिवर्तन भार का अनुमान लगाता है। आउटग्रुप प्रजातियों के एक पैनल पर आधारित जीनोम का। आउटग्रुप के रूप में कार्य करने के लिए विभिन्न प्रजातियों से कम से कम 30 जीनोम असेंबलियों की आवश्यकता होती है और NEWICK प्रारूप में शामिल प्रजातियों की एक दिनांकित फाइलोजेनी होती है। उच्च जी.ई.आर.पी. स्कोर जीनोम में अत्यधिक संरक्षित क्षेत्र को संदर्भित करता है जो हानिकारक होने की संभावना है जबकि कम जीईआरपी स्कोर उन साइटों को इंगित करता है जो उत्परिवर्तन के प्रति तटस्थ हैं।

यह अब तक की पहली और एकमात्र कार्यप्रवाह है जिसे विशेष रूप से एक ही प्रजाति के ऐतिहासिक और आधुनिक अनुक्रमण डेटा को संसाधित करने के लिए डिजाइन किया गया है ताकि उन्हें तुलनीय बनाया जा सके, क्योंकि इसमें ऐतिहासिक और आधुनिक डेटा के प्रसंस्करण के लिए अलग–अलग अनुकूलित ट्रैक हैं जो विभिन्न समय अवधि से संबंधित नमूनों के बीच बेहतर तुलना को सक्षम करते हैं।

कार्यप्रवाह जीनोमिक विविधता सूचकांकों के अनुमानों को प्रदर्शित करती है, जैसे कि जीनोम विविधता और इनब्रीडिंग, म्यूटेशनल लोड, समय के माध्यम से जीनोमिक क्षरण की मात्रा निर्धारित करती है। इन जीनोमिक क्षरण सूचकांकों को विभिन्न समयावधियों के बीच तुलना करने पर प्रजातियों के विलुप्त होने के जोखिम का आकलन करने में प्रासंगिक होने के लिए प्रलेखित किया गया है।

निष्कर्ष

भोजन, दवाइयों और आजीविका की सुरक्षा के लिए मानव की जरूरतों को पूरा करने के लिए, कृषि जैव विविधता संरक्षण उच्च उपज, सुधार के साथ–साथ फसलों, पालतू जानवरों, मछली और वृक्षारोपण पेड़ों की संरक्षण स्थिति पर ध्यान केंद्रित करने का महत्वपूर्ण और आवश्यक कारण बन गया है। किस्म/नस्ल और रोग प्रतिरोधी। यह कार्यप्रवाह शोधकर्ताओं को प्रत्येक जीव की प्रजातियों की गिनती के बजाय जनसंख्या के आकार, उत्परिवर्तन भार, अंत:प्रजनन गुणांक और विषमयुग्मजीता पर नजर रखने की सहमति देती है। यह विविध खाद्य पदार्थों और खाद्य प्रणालियों से संभावित पोषण संबंधी लाभों की पहचान करने में सहायता करता है, और अपने सभी आयामों में खाद्य सुरक्षा सुनिश्चित करते हुए वैश्विक स्तर पर मानव मांग–संचालित दबावों को कम करने की आवष्यकता के साथ–साथ सतत कृषि को सुनिश्चित करता है जो भूमि, पानी और अन्य का कुशल उपयोग करता है।

हितों का टकराव घोषणा

लेखक घोषणा करते हैं कि उनके कोई ज्ञात प्रतिस्पर्धी वित्तीय हित या व्यक्तिगत संबंध नहीं हैं, जो इस पेपर में बताए गए कार्य को प्रभावित करते प्रतीत हो सकते हैं।

संदर्भ

Bijlsma, R. and Loeschcke, V. (2012). Genetic erosion impedes adaptive responses to stressful environments. *Evolutionary Applications*. **5(2)**: 117-129. https://doi.org/10.1111/j.1752-4571.2011.00214.x

Bonfield, J.K., Marshall, J., Danecek, P., Li, H., Ohan, V., Whitwham, A., Keane, T. and Davies, R.M. (2021). HTSlib: C library for reading/writing high-throughput sequencing data. *GigaScience*. **10(2):** giab007. https://doi.org/10.1093/gigascience/giab007

Christiansen, H., Heindler, F.M., Hellemans, B. *et al.* (2021). Facilitating population genomics of non-model organisms through optimized experimental design for reduced representation sequencing. *BMC Genomics*. **22:** 625. https://doi.org/10.1186/s12864-021-07917-3.

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics (Oxford, England)*. **34(17):** i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

Cingolani, P., Platts, A., Wang, l., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X. and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly*. **6(2):** 80-92. https://doi.org/10.4161/fly.19695.

Cooper, G.M., Stone, E.A., Asimenos, G., NISC Comparative Sequencing Program, Green, E.D., Batzoglou, S. and Sidow, A. (2005). Distribution and intensity of constraint in mammalian genomic sequence. *Genome Research*. **15(7):** 901-913. https://doi.org/10.1101/gr.3577405.

Davydov, E.V., Goode, D.L., Sirota, M., Cooper, G.M., Sidow, A. and Batzoglou, S. (2010). Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Computational Biology*. **6(12):** e1001025. https://doi.org/10.1371/journal.pcbi.1001025

Díez-Del-Molino, D., Sánchez-Barreiro, F., Barnes, I., Gilbert, M. and Dalén, L. (2018). Quantifying temporal genomic erosion in endangered species. *Trends in Ecology and Evolution*. **33(3):** 176-185. https://doi.org/10.1016/j.tree.2017.12.002.

Flynn, J.M., Hubley, R., Goubert, C., Rosen, J., Clark, A.G., Feschotte, C. and Smit, A.F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences*. **117(17):** 9451-9457.

Haubold, B., Pfaffelhuber, P. and Lynch, M. (2010). mlRho - a program for estimating the population mutation and recombination rates from shotgun-sequenced diploid genomes. *Molecular Ecology.* **19 Suppl 1(Suppl 1):** 277-284. https://doi.org/10.1111/j.1365-294X.2009.04482.x.

Kutschera, V.E., Kierczak, M., van der Valk, T. *et al*. (2022). GenErode: A bioinformatics pipeline to investigate genome erosion in endangered and extinct species. *BMC Bioinformatics*. **23:** 228. https://doi.org/10.1186/s12859-022-04757-0.

Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H. *et al.* (2014). Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*. **513(7518):** 409-413. https://doi.org/10.1038/nature13673.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.

Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England).* **25(14):** 1754-1760. https://doi.org/10.1093/bioinformatics/btp324.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England).* **25(16):** 2078-2079. https://doi.org/10.1093/bioinformatics/btp352

Mahalakshmi, L. (2022). Identification of genome-wide single nucleotide polymorphisms in indigenous cattle breeds of Tamil Nadu. *Indian Journal of Animal Research.* doi: 10.18805/IJAR.B-4934.

Saravanan, R., Murali, N., Thiruvenkadan, A.K. and Das, D. N. (2022). Comparative genome sequence analysis of bovine lymphocyte antigen *BoLA-DRB3.2* alleles in Deoni and Ongole (*Bos indicus*) cattle breeds of India. *Indian Journal of Animal Research.* **56(7):** 783-789. doi: 10.18805/IJAR.B-4357.

Shafer, A.B., Wolf, J.B., Alves, P.C., Bergström, L., Bruford, M.W., Brännström, I., Colling, G. *et al.* (2015). Genomics and the challenging translation into conservation practice. *Trends in Ecology and Evolution*. **30(2):** 78-87. https://doi.org/10.1016/j.tree.2014.11.009

Smit, A.F.A., Hubley, R. and Green, P. (2017). 1996-2010. RepeatMasker Open-3.0.

van Oosterhout, C. (2020). Conservation genetics: 50 Years and counting. *Conservation Letters*. e12789. https://doi.org/10.1111/conl.12789.