

DOI: <https://doi.org/10.71317/kjard.2.10.2026.772>

Kashmir Journal of Academic Research and Development

Journal homepage: <https://rjsaonline.org/index.php/KJARD>

Bioinformatics-Driven Environmental DNA Metabarcoding and Conservation Genomics for Non-Invasive Wildlife Biodiversity Monitoring under Climate Change

Azmat Ali

Department of Zoology, Government Post Graduate Jahanzeb College, Saidu Sharif, Swat, Khyber Pakhtunkhwa, Pakistan

Email: azmatalizoologist.edu@gmail.com

ORCID iD: 0009-0007-5385-3946

Ibrar Hussain

Research Officer, Veterinary Research and Disease Investigation Centre (VR&DIC), Balogram, Swat

ARTICLE INFO**ABSTRACT****Received:**

August 25, 2026

Revised:

September 05, 2026

Accepted:

September 23, 2026

Available Online:

October 04, 2026

Keywords:

Environmental DNA, metabarcoding, conservation genomics, bioinformatics, climate change, non-invasive monitoring, wildlife biodiversity.

Corresponding Author:**Azmat Ali****Email:**azmatalizoologist.edu@gmail.com

Global biodiversity is declining rapidly as anthropogenic climate change interacts with habitat loss, disease emergence and overexploitation to reshape species distributions and population viability. Conventional survey methods, including live trapping, visual census and invasive tissue sampling, are frequently slow, costly and disturbing to the organisms under study, so they cannot keep pace with the scale at which ecosystems are now changing. Environmental DNA (eDNA) metabarcoding, supported by conservation genomics and modern bioinformatic pipelines, offers a non-invasive alternative capable of detecting entire assemblages of organisms from water, soil, air, sediment and even surface swabs. This review synthesises recent methodological and empirical literature to examine how bioinformatics-driven eDNA metabarcoding and population genomic tools are being integrated for wildlife biodiversity monitoring under a changing climate. It traces the workflow from sample collection through amplicon and shotgun sequencing to taxonomic classification, highlighting the computational tools, reference databases and quality control steps that determine the reliability of biodiversity inferences. It then considers how population genomic approaches complement eDNA surveys by estimating effective population size, inbreeding, adaptive potential and demographic history in threatened species. Case studies spanning critically endangered birds, marine reptiles, river ecosystems, airborne biomonitoring and disease surveillance in migratory waterfowl illustrate the practical reach of these tools. The review closes by discussing persistent limitations, including reference database incompleteness, taxonomic classifier disagreement, degradation of DNA under variable environmental conditions, and the computational capacity gap faced by many biodiversity-rich but resource-limited regions, before outlining priorities for standardisation and capacity building relevant to zoological research in Pakistan and other developing regions.

Introduction

Climate change is one the most widespread causes of ecological change on earth with measurable changes in species distribution, phenology and interspecific interactions in terrestrial, freshwater and marine biomes [1]. A recent meta-analytical synthesis of marine fish and invertebrates demonstrates that ocean warming and acidification produce significant deviations in calcification, survival, metabolism, reproduction and behaviour even at moderate intensities, and that these deviational effects are roughly twice as common as directional changes reported in earlier literature [2]. At the same time, population-level assessments across thousands of vertebrate species indicate a pattern of biological annihilation, with widespread range contraction and local population loss occurring even among taxa not yet formally classified as threatened [3]. This is further

complicated by interacting stressors as the environmental changes of temperature and precipitation are also changing the transmission dynamics of wildlife pathogens and the distribution of their hosts and vectors [4].

Monitoring tools that are sensitive enough to detect subtle changes in populations, scalable enough to monitor large and remote landscapes, and gentle enough to minimize additional stressors to already vulnerable populations are needed to address these overlapping pressures [5]. Traditional methods of biodiversity assessment, like electrofishing, mist netting, live trapping, invasive blood or tissue sampling, and extended visual or acoustic surveys, typically meet only one or two of these three criteria; they are often taxonomically limited, labour intensive and can have adverse impacts on the species under investigation [5]. Environmental DNA, defined as genetic material shed by organisms into water, soil, air, sediment or other environmental substrates through skin cells, mucus, faeces, urine, gametes or decomposing tissue, has emerged as a practical solution to this monitoring gap because it allows entire biological communities to be characterised from a single environmental sample without ever encountering the source organism [6]. Coupled with high-throughput sequencing, eDNA metabarcoding can simultaneously screen for dozens or hundreds of taxa spanning microbes, invertebrates, fish, amphibians, reptiles, birds and mammals within one assay, a capability that has driven its rapid adoption across biodiversity science over the past decade [7].

The reliability of any eDNA survey, however, depends heavily on the bioinformatic decisions made after sequencing, including the choice of taxonomic classifier, the completeness of the reference database, and the thresholds applied to filter sequencing artefacts from genuine biological signal, since these choices can alter detection outcomes as much as the underlying ecology does [7]. In parallel, conservation genomics has matured from a specialised academic pursuit into a practical management tool, providing wildlife agencies with genome-wide estimates of inbreeding, effective population size, adaptive capacity and demographic history that were previously unattainable for most non-model species [8]. Because both eDNA metabarcoding and conservation genomics rely on the same underlying sequencing and bioinformatic infrastructure, the two fields are increasingly converging into a single non-invasive genomic toolkit capable of monitoring species presence, population health and genetic diversity from the same sample stream [9]. This convergence is particularly consequential for critically endangered species with small, geographically restricted populations, where every additional handling event carries a disproportionate risk and where genomic information on inbreeding and adaptive potential is most urgently needed [10].

Existing reviews have generally treated eDNA metabarcoding methodology, conservation genomics and climate change biology as separate literatures, each with its own dedicated syntheses, so that a researcher wishing to understand how bioinformatic pipeline choices propagate through to population-level genomic inference under a changing climate must currently consult several disconnected bodies of work rather than a single integrated account [8]. The present review addresses this gap by deliberately tracing a single continuous pipeline, from raw sequence read to conservation-relevant genomic inference, while foregrounding the climate change context in which that pipeline is increasingly being deployed [7]. It gives explicit attention throughout to the practical accessibility of these tools for zoology researchers working in biodiversity-rich but resource-constrained settings such as South Asia, where laboratories may lack the sequencing infrastructure of larger institutions [8]. The review begins with the molecular and ecological principles of eDNA metabarcoding, proceeds through the bioinformatic pipelines that convert raw sequence reads into taxonomic assignments, examines the population genomic toolkit used for non-invasive conservation assessment, surveys emerging technologies such as long-read nanopore sequencing, shotgun metagenomics, airborne eDNA and autonomous sampling platforms, and concludes with a discussion of persistent limitations and priorities for future research and capacity building [11].

The scope of this review is deliberately centred on wildlife biodiversity monitoring, and it therefore does not attempt a comprehensive treatment of agricultural or plant eDNA applications, forensic uses of environmental genetic material, or detailed mathematical methods used in population genetic models because these topics are outside the scope of this review.

Principles of Environmental DNA (eDNA) and Metabarcoding

Environmental DNA exists in two broad physical forms within a sample, intracellular DNA still enclosed within shed cells and extracellular DNA that has been released following cell lysis, and the two forms differ in their persistence, transport and degradation behaviour within air, water, soil and sediment matrices [6]. Once released, eDNA is subject to physical, chemical and biological degradation processes that are strongly influenced by temperature, ultraviolet exposure, pH, microbial activity and substrate type, so that detection windows can range from hours in warm turbulent water to centuries or millennia within cold, anoxic sediments [5]. This variation in eDNA persistence is important for climate change research because warmer water and changes in water flow may speed up eDNA degradation in many freshwater and marine environments. This could reduce the ability to detect species at a time when biodiversity changes are happening rapidly [7]. Controlled mesocosm experiments support this concern, showing that higher water temperatures increase the removal of small eDNA particles from flowing streams. This occurs through faster microbial breakdown and changes in how eDNA is physically trapped in the environment. As a result, species may become harder to detect as stream temperatures increase with climate change [12].

Because that removal effect varies with particle size and flow regime, seasonal and geographic differences in water temperature can introduce a climate-linked bias into eDNA-based abundance and occupancy estimates that survey designers need to account for explicitly [12]. Metabarcoding, the amplicon-based branch of eDNA analysis, relies on universal or semi-universal polymerase chain reaction (PCR) primers that target short, variable regions of conserved genes, most commonly mitochondrial 12S and 16S ribosomal RNA genes for vertebrates and the cytochrome c oxidase subunit I gene for invertebrates, flanked by highly conserved primer-binding sites [11]. Following amplification, samples are pooled and sequenced on high-throughput platforms, generating millions of short reads that must then be denoised, clustered or resolved into amplicon sequence variants before taxonomic identity can be assigned [13]. Marine biomonitoring surveys applying this workflow across representative tropical reef habitats have shown that a single water sample can recover sequences spanning the full tree of life, from bacteria and protists through to fish, corals, molluscs and marine mammals, in a manner that traditional single-taxon survey methods cannot replicate within comparable time and cost budgets [14].

The non-invasive character of eDNA sampling is especially valuable for rare, cryptic, nocturnal or dangerous species, and for habitats where physical access is difficult or ecologically sensitive, since a water, soil or sediment sample can be collected without disturbing the animals or their surroundings [6]. Recognition of this advantage has led to a rapid diversification of sampling substrates beyond open water, including soil, snow tracks, faecal deposits, nesting substrate, and even DNA adhering to vegetation surfaces browsed or touched by animals [9]. Vertebrate DNA recovered from simple cotton swabs wiped across the leaves of tropical forest vegetation has, for example, allowed researchers to detect the presence of primates, ungulates and other mammals along transects without any need for camera traps or direct observation [15]. Sampling efficiency in physically and chemically challenging tropical habitats requires its own methodological calibration, since a study conducted in a Malaysian blackwater peat swamp found that the volume of acidic, tannin-rich water filtered had a strong positive correlation with recovered DNA concentration, with larger filtered volumes needed to achieve reliable detection under these extreme conditions compared with clearer temperate waters [16].

Similar reasoning has motivated the collection of eDNA from beach sand at sea turtle nesting sites, where accumulated DNA shed by nesting females and hatchlings enables both species detection and, when population genomic methods are applied downstream, the genetic characterisation of the reproductive population without ever restraining an animal [17]. Taken together, these developments illustrate that the substrate from which eDNA is recovered has become almost as flexible as the questions researchers wish to ask, extending the reach of non-invasive monitoring far beyond its original aquatic origins [7].

Table 1. Environmental DNA sampling substrates and representative applications discussed in this review.

Substrate	Representative application	Reference
Open water (marine and freshwater)	Whole-community biomonitoring across the tree of life	[14]
Flowing river and stream water	Temperature-dependent eDNA removal and persistence studies	[12]
Beach and nesting sand	Sea turtle species detection and nesting-cohort population genomics	[17]
Soil (forest floor, display sites, nests)	Individual identification of the critically endangered kakapo	[9]
Vegetation surface (leaf swabs)	Detection of primates and other mammals without direct observation	[15]
Air (passive and active filtration)	Whole-biome and fungal community characterisation	[31,32]
Tropical peat swamp water	Calibration of filtration volume against DNA yield in acidic waters	[16]
Faeces / scat	Rapid nanopore-based species identification of elusive mammals	[24]

Sampling design decisions, including the number of replicate samples collected, the volume of water or mass of soil processed, and the season and time of day at which collection occurs, all influence detection probability and must therefore be matched to

the ecology of the target species and the climatic conditions of the study site [5]. Standardised sampling manuals developed for widely used marker systems have therefore become an important complement to the underlying molecular methods, since they translate general principles of eDNA ecology into concrete field procedures that non-specialist survey teams can follow consistently [6].

Bioinformatic Pipelines: From Raw Sequence Reads to Taxonomic Assignment

Converting raw sequencing output into a reliable list of detected species is a multistep computational process in which each stage introduces its own sources of error and uncertainty [11]. Quality filtering and adapter trimming are typically performed first to remove low-quality bases and residual primer or adapter sequence, since uncorrected sequencing artefacts can otherwise be misinterpreted as genuine biological variation at later stages [13]. Negative controls, consisting of field and laboratory blanks carried through the entire workflow without added template, and positive controls of known composition are essential accompaniments to every sequencing run, since they allow contamination and cross-sample carryover to be identified and, where necessary, subtracted from the dataset before any ecological interpretation is attempted [5]. Index hopping, in which sequencing reads are misassigned between multiplexed samples sharing a sequencing lane, can further introduce low-level false positive detections and is best mitigated through unique dual-indexing strategies and conservative post-hoc filtering of rare, cross-sample reads [18].

Denoising algorithms then resolve the filtered reads into amplicon sequence variants that represent true biological sequences rather than sequencing errors, a step that has substantially improved the taxonomic resolution achievable from short marker gene fragments compared with older similarity-based clustering approaches [13]. Chimeric sequences, artificial hybrids formed when a partially extended amplicon primes on a different template during PCR, are identified and removed at this stage, since their inclusion would otherwise be misread as novel or divergent taxa that do not actually exist in the sampled community [13]. Community analysis platforms built around modular plugin architectures allow these denoising, filtering and classification steps to be assembled into reproducible, version-controlled workflows that can be shared between laboratories and reanalysed as reference databases improve [18]. Taxonomic classification, the step in which each amplicon sequence variant is matched against a reference database to assign a putative species, genus or family identity, is where much of the remaining uncertainty in eDNA studies resides [11].

A systematic comparison of nine taxonomic classifiers applied to curated marine vertebrate reference databases found that alignment-based tools using stringent percentage identity cutoffs generally achieved higher accuracy and fewer false positives than naive Bayes classifiers for short mitochondrial markers, while naive Bayes approaches performed comparatively better for the longer cytochrome c oxidase subunit I marker [11]. The same evaluation demonstrated that when reference databases were experimentally depleted to simulate real-world incompleteness, several classifiers began to report bacterial sequences as false-positive marine vertebrate detections, while a separate comparison found that at least one classifier mislabelled a least-concern eel species as its near-threatened congener, illustrating how both database gaps and classifier choice can directly distort conservation-relevant conclusions [11]. Non-Bayesian taxonomy classifiers designed for ribosomal and internal transcribed spacer sequences represent one attempt to combine the speed of k-mer based methods with more principled statistical confidence estimation for ambiguous or database-poor taxa [19]. Underlying all of these classifiers is the quality and taxonomic breadth of the public sequence repositories against which query sequences are compared, and the continued growth and curation of these central nucleotide archives remains a foundational, if often underappreciated, requirement for accurate species identification from environmental samples [20].

Because reference databases remain systematically better populated for well-studied northern hemisphere taxa than for many tropical, freshwater or regionally endemic species, classification accuracy in biodiversity-rich developing regions is disproportionately constrained by this uneven global sequencing effort [11]. Beyond amplicon-based metabarcoding, shotgun metagenomic sequencing of eDNA analyses the entire pool of genetic material present in a sample without amplification bias, recovering nuclear and mitochondrial sequence simultaneously and enabling downstream analyses such as population genetics and pathogen variant calling that are unavailable to marker-based approaches [21]. This approach traces its conceptual lineage to early whole-community shotgun sequencing of complex natural samples such as ocean water, which first demonstrated that untargeted sequencing could recover genomic information from thousands of co-occurring taxa within a single environmental sample [22]. A related low-cost alternative, genome and metagenome skimming, sequences high-copy genomic regions such as mitochondrial, chloroplast and ribosomal DNA at shallow coverage, and has recently been proposed as a bridge between the cost-effectiveness of amplicon metabarcoding and the taxonomic breadth of full shotgun metagenomics for eDNA applications currently limited mainly to ancient sedimentary samples [23].

Nanopore adaptive sampling, which selectively enriches for pre-specified mitochondrial reference sequences during sequencing itself rather than through prior PCR amplification, has similarly been used to obtain rapid, field-deployable species

identifications directly from faecal DNA of elusive mammals, correctly resolving species identity in the great majority of samples tested against a national mitochondrial reference database [24]. That faecal identification approach further identified a case of species misidentification in an earlier study based on physical appearance alone, illustrating how bioinformatic genetic confirmation can correct errors that purely field-based non-invasive monitoring would otherwise perpetuate [24]. Amplicon metabarcoding remains considerably cheaper per sample and better suited to high-throughput screening of large numbers of sites, whereas shotgun approaches recover a richer genomic signal from each sample at correspondingly higher sequencing cost, a distinction that should guide study design according to the specific monitoring question being addressed [11]. Table 1 summarises representative bioinformatic tools referenced in this review, together with their principal function within the eDNA and conservation genomics workflow [11].

Table 2. Representative bioinformatic tools and resources used in eDNA metabarcoding and conservation genomic workflows.

Tool / resource	Primary function	Reference
Field and laboratory negative/positive controls	Detection and subtraction of contamination and cross-sample carryover	[5]
Dual-indexing / unique tag combinations	Mitigation of index hopping between multiplexed samples	[18]
DADA2	Denoising of amplicon reads into amplicon sequence variants and removal of chimeric sequences	[13]
QIIME 2	Modular, reproducible metabarcoding workflow platform	[18]
SINTAX	Non-Bayesian taxonomy classification for marker genes	[19]
GenBank	Central nucleotide sequence reference repository	[20]
Taxonomic classifiers (alignment-based, naive Bayes, k-mer based)	Assignment of sequence variants to reference taxa	[11]
Nanopore long-read sequencing platforms	Field-deployable long-read and adaptive sequencing	[9,24,30]
Shotgun / genome-skimming pipelines	Untargeted or shallow whole-sample sequence recovery	[21,22,23]

Conservation Genomics for Non-Invasive Population Assessment

Conservation genomics extends the logic of eDNA metabarcoding from species detection to population-level genetic characterisation, applying genome-wide sequencing to estimate parameters such as effective population size, inbreeding coefficients, genetic load and adaptive capacity in wild populations [8]. These parameters matter because reduced genetic diversity and elevated inbreeding are strongly associated with lowered fitness and diminished capacity to adapt to environmental change, making population genomic monitoring a direct complement to demographic and distributional surveillance [25]. Genomic approaches allow such estimates to be obtained from minimally invasive samples, including feathers, shed hair, faeces, buccal swabs and, increasingly, from the same environmental substrates used for species-level metabarcoding, removing the need for repeated capture of the same rare individuals [8]. The application of this integrated approach to one of the world's rarest birds illustrates its practical value, since whole genome sequencing of the surviving island population and historical mainland specimens of the critically endangered kakapo (*Strigops habroptilus*), of which only around two hundred individuals remain, was undertaken to characterise the genomic health of the species [10].

That sequencing effort found that the present-day population carries fewer harmful mutations than population genetic theory would predict, apparently because long-term inbreeding facilitated the natural purging of deleterious variants rather than their accumulation [10]. That same genomic dataset also highlighted persistently high rates of egg hatching failure in the species and called for further genomic investigation into its underlying causes, illustrating how whole genome sequencing can both resolve

long-standing uncertainties and identify precise questions for future non-invasive monitoring to address [10]. Building directly on this genomic foundation, soil eDNA metabarcoding combined with long-read nanopore sequencing has since been used to detect individual kakapo from soil samples collected at display sites and nests, matching recovered haplotypes against the known population reference panel to identify specific birds without any physical capture [9]. That same study additionally detected trace reads of the invasive Polynesian rat (*Rattus exulans*) from the accumulated soil eDNA signal on a predator-free island sanctuary, demonstrating that a single non-invasive sampling protocol can simultaneously serve species monitoring, individual identification and biosecurity surveillance objectives [9].

A comparable logic has been applied to marine reptiles, where environmental DNA recovered from nesting beach sand has been used to confirm the presence of nesting sea turtle species, including the green turtle (*Chelonia mydas*) and loggerhead turtle (*Caretta caretta*) [17]. The same beach sand samples have also been used to reconstruct population genomic signatures of the nesting cohort, offering a route to monitor genetic diversity at rookeries that are otherwise accessible only during a narrow nesting season [17]. Population genomic monitoring is particularly well suited to detecting genetic erosion that precedes visible demographic decline, since the loss of rare alleles and heterozygosity can occur well before a population's census size falls to a level that would trigger conventional conservation alarm [8]. Beyond estimating genetic diversity itself, epigenetic clocks that infer an individual's chronological or biological age from DNA methylation patterns offer a further non-lethal genomic biomarker, and a recently published design workflow validated against an extensive Arctic polar bear dataset demonstrates that reliable wildlife epigenetic clocks can be constructed even from limited sample sizes typical of small conservation budgets [26].

Age structure and biological ageing rate estimates of this kind complement effective population size and inbreeding metrics by revealing whether a population's demographic profile is skewed toward older, potentially less reproductively active individuals, a pattern that could otherwise go undetected under climate-driven stress [26]. Under accelerating climate change, this early warning capacity is increasingly valuable, because populations that retain greater standing genetic variation are generally considered better placed to respond adaptively to novel selective pressures such as shifting thermal regimes, altered precipitation and emerging disease [25].

Table 3. Selected non-invasive genomic monitoring case studies discussed in this review.

Species / system	Genomic approach	Reference
Kakapo (<i>Strigops habroptilus</i>)	Whole genome sequencing for pedigree and inbreeding assessment	[10]
Kakapo (<i>Strigops habroptilus</i>)	Soil eDNA and long-read sequencing for individual identification	[9]
Sea turtles (<i>Chelonia mydas</i> , <i>Caretta caretta</i>)	Beach sand eDNA for species detection and population genomics	[17]
Leopard cat (<i>Prionailurus bengalensis</i>), Pakistan	Faecal DNA and next-generation sequencing for non-invasive diet analysis	[38]
Sympatric carnivores (leopard, fox, jackal, civets), Pakistan	Molecular scatology to correct morphological scat misidentification	[37]
Wild and domestic animal gut microbiota	Shotgun metagenomics for antimicrobial resistance surveillance	[27]
Migratory waterfowl	Mechanistic modelling of climate effects on avian influenza dynamics	[28]
Various wildlife (Arctic polar bear dataset)	Epigenetic clocks for non-lethal age and ageing-rate estimation	[26]
Amphibians and reptiles (global)	Systematic audit of genomic resource coldspots	[39]

The same sequencing infrastructure that supports these applications is also being used to characterise the wider community of organisms sharing an ecosystem with a focal threatened species, including the assessment of antimicrobial resistance genes circulating among wild and domestic animal populations that may share pathogens or landscapes with vulnerable wildlife [27].

Climate Change, Disease Emergence and the Case for Integrated Monitoring

Climate change rarely acts on wildlife populations in isolation, and its effects on infectious disease dynamics illustrate why biodiversity monitoring increasingly needs to integrate genomic, ecological and epidemiological data streams [4]. Mechanistic modelling of highly pathogenic avian influenza in migratory waterfowl demonstrates that warming-driven shifts in spring migration timing can extend the duration of outbreaks and increase mortality by lengthening the period birds spend at shared breeding grounds, while simultaneously altering the environmental persistence of the virus itself [28]. That same modelling work found that the relative importance of transmission related viral traits versus environmental decay depended on whether a viral strain relies primarily on direct bird-to-bird transmission or on environmentally persistent virus, indicating that climate impacts on disease dynamics cannot be predicted from a single generic pathway [28]. Long-term capture data from resident, non-migratory understory birds in the Brazilian Amazon provide a further example of how shifting climatic baselines are reshaping wildlife population trajectories, with a one-degree Celsius increase in dry-season temperature associated with a substantial reduction in survival across the resident bird community even within structurally undisturbed, pristine rainforest [29].

Bioinformatics-enabled eDNA and metagenomic approaches are well positioned to support this kind of integrated surveillance, since a single shotgun metagenomic sequencing run of river water can simultaneously screen for pathogens, pollutants and the wider vertebrate, invertebrate and microbial community sharing that habitat [21]. Applying this logic to a representative Irish river system, from a mountain tributary through to its tidal outlet, long-read shotgun metagenomic sequencing recovered genomic information spanning viruses, bacteria, fungi and animals within the same set of water samples, illustrating the feasibility of single-assay biodiversity and health surveillance across an entire catchment [21]. Because such combined surveillance depends on recovering low-abundance, fragmented genetic material from complex mixtures, the same core bioinformatic principles of reference database completeness and classifier accuracy discussed earlier directly determine whether an emerging pathogen or a climate-sensitive indicator species will be correctly detected [11].

Emerging Technologies: Long-Read Sequencing, Shotgun Metagenomics, Airborne eDNA and Autonomous Sampling

Long-read nanopore sequencing has substantially expanded what non-invasive genomic monitoring can achieve in the field, since portable sequencing devices can be deployed at remote sampling sites without reliance on centralised sequencing facilities [30]. The same long-read technology underpins the individual-level kakapo identification described earlier, since sufficiently long reads are required to construct robust haplotypes from low-concentration soil eDNA that would otherwise be indistinguishable from sequencing noise using short-read platforms alone [9]. Air has emerged as an unexpectedly rich and logistically convenient substrate for this expanded eDNA toolkit, with airborne shotgun metagenomic sequencing shown to recover genomic information sufficient for whole biome characterisation, population genetic inference and detection of genomic variation from filtered air samples alone [31]. Complementing this airborne approach, passive air filtration surveys targeting fungal communities have revealed spatially and seasonally predictable patterns of airborne DNA deposition that mirror known phenological and biogeographic patterns, supporting the ecological validity of air as a genuinely informative sampling medium rather than merely a novelty substrate [32].

Robotic and autonomous platforms are simultaneously being integrated into eDNA sampling protocols to reach habitats that are difficult, dangerous or ecologically sensitive for human collectors to access directly, including forest canopies and cliff faces [33]. A drone-based sampling system designed to collect eDNA directly from tree branches without ground-level access has demonstrated that autonomous platforms can extend biodiversity monitoring into vertically structured habitats that are otherwise chronically undersampled [33]. Camera trap imagery, long a mainstay of non-invasive wildlife monitoring, has itself been transformed by deep learning classification systems capable of automatically identifying, counting and describing the behaviour of dozens of species across millions of images with accuracy comparable to trained human analysts, freeing substantial volunteer and researcher labour for other conservation tasks [34]. Beyond passive detection, eDNA-guided decision making is also being incorporated directly into active management interventions, since a targeted invasive fish removal programme that used quantitative PCR detection of loach eDNA to generate spatial heat maps of likely occurrence achieved substantially higher trapping efficiency than a fixed, non-informed trapping design [35].

That programme's experience further showed that eDNA and conventional trapping methods agreed on species presence in roughly four out of five cases, and that traps failed to detect the target species at a substantial proportion of sites where eDNA had already returned a positive result, reinforcing the value of combining molecular and physical detection methods within a single management workflow [35]. As these technologies mature, questions of data governance and international data sharing become increasingly pressing, since eDNA samples collected in one jurisdiction frequently contain genetic information relevant to species, pathogens or populations of interest to researchers and regulators elsewhere [36].

Relevance to Wildlife Monitoring in Pakistan and Comparable Regions

Pakistan and neighbouring montane regions of South and Central Asia contain terrain that is logistically difficult to survey, a challenge illustrated directly by field work in the Himalayan foothills of Pakistan, where morphological identification of carnivore scats collected from rugged, elevated terrain was found to misclassify a substantial proportion of samples until molecular scatology was used to confirm species identity from the same faecal DNA [37]. This kind of field evidence, that genetic methods can correct identification errors which conventional morphological survey alone would have missed, is precisely the advantage that the low-disturbance, single-sample nature of eDNA metabarcoding offers more broadly for mountain, riverine and rangeland ecosystems across the region [5]. Non-invasive molecular approaches are not a purely hypothetical prospect for Pakistan, since faecal DNA from the leopard cat (*Prionailurus bengalensis*) collected in two ecologically distinct areas of the country was used more than a decade ago to characterise the species' diet through next-generation sequencing of a short mitochondrial marker, demonstrating that the core genomic principles reviewed here have an established track record in Pakistani wildlife research [38]. Soil, snow and stream substrates already validated for eDNA recovery in other high-altitude and temperate systems provide a template that could extend this existing local expertise to a wider range of priority species with comparatively modest additional laboratory investment, provided that sequencing and computational infrastructure are available locally or through regional collaboration [9].

Taxonomic classification accuracy has been shown to depend directly on the regional representation of reference sequence databases even within well-studied marine vertebrate systems, indicating that classification of samples from any less-represented region or taxonomic group is likely to be similarly constrained until reference coverage improves [11]. This concern is not merely theoretical, since a systematic global audit of genomic resources for amphibians and reptiles has documented pronounced coldspots in genomic sampling that align closely with regions of high species richness and predicted global change exposure, concluding that equitable international collaboration and local capacity building are needed to close this gap rather than further concentration of sequencing effort in already well-resourced countries [39]. A parallel initiative producing the first reference genome assemblies for seven endemic bird families of the Western Ghats of India illustrates one practical route to redressing this imbalance, generating high-quality, locally led genomic resources for a South Asian biodiversity hotspot that had previously lacked dedicated reference assemblies [40]. Following this example, locally generated reference genomes and barcode libraries for the montane mammals, birds and reptiles of Pakistan's northern mountain ranges would provide a comparable foundation for eDNA and conservation genomic work in the region, building on the existing molecular groundwork already established for at least one Pakistani carnivore species, rather than requiring researchers to rely on more distantly related reference sequences of uncertain applicability [38].

Embedding bioinformatics training within zoology curricula and regional research networks, alongside investment in shared sequencing and computational facilities, would allow universities and wildlife research institutions across the region to participate directly in the generation and interpretation of eDNA and conservation genomic data rather than depending entirely on external analysis [8]. A comparable pan-continental model has already been implemented in Africa, where a coordinated genomics and bioinformatics capacity building initiative delivered dozens of workshops and trained several hundred researchers across dozens of African countries within a single year, illustrating that large-scale regional capacity building of this kind is logistically achievable when sustained institutional and international support is secured [41]. A similarly coordinated regional initiative for Pakistan and neighbouring South Asian countries, potentially linked to existing zoology departments and museum collections, could adopt a comparable workshop and fellowship model to accelerate local uptake of eDNA metabarcoding and conservation genomic methods [41].

Challenges and Limitations

Despite its rapid growth, bioinformatics-driven eDNA metabarcoding continues to face several interlinked technical challenges that constrain the confidence with which results can be interpreted [7]. Reference database incompleteness remains the most consistently cited limitation, since a taxon absent from the database cannot be correctly identified regardless of how accurate the classification algorithm applied to it may be [7]. Environmental factors that govern eDNA persistence and transport, including temperature, flow, ultraviolet exposure and substrate chemistry, remain incompletely characterised across the full range of climates and habitats in which eDNA surveys are now being deployed, introducing uncertainty into the interpretation of both positive and negative detections [7]. This uncertainty is likely to be exacerbated under climate change itself, since shifting thermal and hydrological regimes may alter degradation rates in ways that are not yet incorporated into standard eDNA sampling protocols or statistical detection models [5].

Cross-contamination during field sampling, laboratory processing and sequencing library preparation represents a further persistent risk, particularly for sensitive applications such as individual genetic identification in critically endangered species, where a single false detection could carry disproportionate management consequences [9]. Polymerase chain reaction inhibitors

naturally present in humic-rich soils, tannin-stained tropical waters and sediment-laden samples can suppress amplification and generate false negative results even when target DNA is genuinely present, a problem that typically requires additional purification steps or inhibitor-tolerant polymerases to overcome [5]. Universal or semi-universal primers rarely amplify every taxon within a community with equal efficiency, so that primer bias can systematically under-represent or entirely miss species whose marker regions diverge from the primer binding site, meaning that apparent absence in metabarcoding data does not always equate to true absence [6]. Because both PCR inhibition and primer bias produce false negatives that are difficult to distinguish from genuine non-detection, formal occupancy modelling approaches that explicitly estimate imperfect detection probability alongside site-level presence are increasingly recommended as a statistical safeguard against overinterpreting negative eDNA results [7].

Disagreement between taxonomic classifiers on ambiguous or database-poor sequences means that the choice of bioinformatic pipeline can itself become a source of variation between studies, complicating efforts to compare biodiversity trends across research groups using different analytical workflows [11]. Computational infrastructure and bioinformatics expertise remain unevenly distributed across the regions where biodiversity monitoring is most urgently needed, so that laboratories in many biodiversity-rich but resource-limited countries continue to depend on external collaborators for sequence analysis, which can slow the pace of locally led conservation decision making [8]. Ethical and legal questions surrounding the ownership, storage and international transfer of genetic material recovered from environmental samples, particularly where human or other sensitive genetic signal is incidentally captured, add a further layer of complexity that technical guidelines alone cannot resolve [36]. Genomic and molecular approaches also cannot substitute entirely for direct observation, since body condition, disease symptoms, precise age and sex ratios, reproductive status and behaviour remain observable only through physical encounter with an animal, meaning that eDNA and conservation genomics are best understood as complements to, rather than replacements for, traditional field methods [8].

Finally, translating a species or genomic detection into an actionable conservation outcome still requires integration with demographic, ecological and socio-political context, since genomic and molecular data alone cannot determine whether a detected population is stable, declining or recovering without complementary long-term monitoring [42].

Future Directions

Standardising bioinformatic workflows across laboratories, including agreed thresholds for quality filtering, classifier choice and reference database curation, would substantially improve the comparability of eDNA-based biodiversity trends reported from different regions and research groups [7]. Continued investment in reference genome and barcode sequencing for underrepresented taxa, particularly from tropical, freshwater and high-altitude regions, is likely to yield disproportionate improvements in classification accuracy relative to further refinement of the classification algorithms themselves [7]. Wider adoption of portable, field-deployable long-read sequencing platforms could shorten the interval between sample collection and actionable results, an advantage that is particularly valuable for time-sensitive applications such as invasive species detection and disease outbreak response [30]. Expanding shotgun metagenomic and airborne eDNA approaches beyond their current proof-of-concept demonstrations toward routine, cost-effective biodiversity and biosecurity surveillance represents a realistic near-term goal given the pace of recent methodological progress [31].

Integrating automated image-based species identification with genomic eDNA data streams, alongside further development of autonomous sampling platforms for structurally complex or hazardous habitats, would extend non-invasive monitoring capacity into environments that remain chronically undersampled [33]. Explicit incorporation of climate projections into eDNA sampling design, including anticipated changes in degradation rates, hydrological connectivity and species range shifts, would strengthen the ability of monitoring programmes to distinguish genuine ecological signal from methodological artefact [2]. Perhaps most importantly, long-term paired validation studies that track eDNA signal strength alongside independently derived abundance or occupancy estimates from traditional census methods at the same sites over multiple years remain scarce, and closing this gap would do more than any single technical refinement to establish the confidence with which eDNA trends can be treated as genuine proxies for population change [8]. Finally, deliberate investment in regional sequencing infrastructure, open reference database contribution and bioinformatics training, particularly across South Asia and other biodiversity-rich developing regions, would help ensure that the benefits of this rapidly advancing toolkit accrue as much to the countries hosting the biodiversity in question as to the institutions that currently lead its technical development [42]. Formal partnership programmes that pair genomic coldspot regions with well-resourced sequencing centres, of the kind already proposed for amphibian and reptile conservation genomics, offer a concrete mechanism through which this investment could be organised and sustained over the long term [39].

Conclusion

Bioinformatics-driven environmental DNA metabarcoding and conservation genomics together represent a substantial expansion of the wildlife monitoring toolkit available to zoologists and conservation practitioners. By allowing entire biological communities and, increasingly, the genomic health of individual populations to be characterised from water, soil, sediment, air or surface samples, these approaches directly address the scale, cost and disturbance limitations that have long constrained traditional survey methods. Their reliability nonetheless rests on a chain of interdependent bioinformatic decisions, from quality filtering through taxonomic classification to reference database curation, each of which introduces its own uncertainties that must be understood and managed rather than treated as a black box. As climate change continues to reshape species distributions, alter disease dynamics and accelerate biodiversity loss, the capacity to monitor wildlife populations non-invasively, repeatedly and at meaningful spatial scale becomes not merely convenient but essential. Realising this potential fully will require sustained investment in reference databases, computational infrastructure and regional bioinformatics capacity, so that the countries and institutions closest to the world's most threatened biodiversity are equally equipped to generate, interpret and act upon their own genomic and molecular monitoring data.

Acknowledgements

Heartfelt thanks to Prof. Dr. Muhammad Iqbal, Dr. Muhammad Aammar Tufail and Mr. Said Ali for their valuable guidance.

Conflict of interest

The authors declare no conflicts of interest.

Financial Support

This research was conducted independently without external funding.

References

- [1] Parmesan C, Yohe G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421(6918), 37-42. doi:10.1038/nature01286.
- [2] Alter K, Jacquemont J, Claudet J, Lattuca ME, Barrantes ME, Marras S, et al. (2024). Hidden impacts of ocean warming and acidification on biological responses of marine animals revealed through meta-analysis. *Nat Commun*, 15, 2885. doi:10.1038/s41467-024-47064-3.
- [3] Ceballos G, Ehrlich PR, Dirzo R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proc Natl Acad Sci U S A*, 114(30), E6089-E6096. doi:10.1073/pnas.1704949114.
- [4] Pfenning-Butterworth A, Buckley LB, Drake JM, Farner JE, Farrell MJ, Gehman ALM, et al. (2024). Interconnecting global threats: climate change, biodiversity loss, and infectious diseases. *Lancet Planet Health*, 8(4), e270-e283. doi:10.1016/S2542-5196(24)00021-4.
- [5] Bohmann K, Evans A, Gilbert MTP, Carvalho GR, Creer S, Knapp M, et al. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends Ecol Evol*, 29(6), 358-367. doi:10.1016/j.tree.2014.04.003.
- [6] Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH. (2012). Environmental DNA. *Mol Ecol*, 21(8), 1789-1793. doi:10.1111/j.1365-294X.2012.05542.x.
- [7] Deiner K, Bik HM, Machler E, Seymour M, Lacoursiere-Roussel A, Altermatt F, et al. (2017). Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Mol Ecol*, 26(21), 5872-5895. doi:10.1111/mec.14350.
- [8] Hohenlohe PA, Funk WC, Rajora OP. (2021). Population genomics for wildlife conservation and management. *Mol Ecol*, 30(1), 62-82. doi:10.1111/mec.15720.
- [9] Urban L, Miller AK, Eason D, Vercoe D, Shaffer M, Wilkinson SP, et al. (2023). Non-invasive real-time genomic monitoring of the critically endangered kakapo. *eLife*, 12, RP84553. doi:10.7554/eLife.84553.
- [10] Dussex N, van der Valk T, Morales HE, Wheat CW, Diez-Del-Molino D, von Seth J, et al. (2021). Population genomics of the critically endangered kakapo. *Cell Genomics*, 1, 100002. doi:10.1016/j.xgen.2021.100002.

- [11] Bayer PE, Bennett A, Nester G, Corrigan S, Raes EJ, Cooper M, et al. (2025). A comprehensive evaluation of taxonomic classifiers in marine vertebrate eDNA studies. *Mol Ecol Resour*, 25, e14107. doi:10.1111/1755-0998.14107.
- [12] Snyder ED, Tank JL, Pruitt AN, Peters B, Brandao-Dias PFP, Curtis EM, et al. (2025). Warming increases environmental DNA (eDNA) removal rates in flowing waters. *Environ DNA*, 7, e70094. doi:10.1002/edn3.70094.
- [13] Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods*, 13, 581-583. doi:10.1038/nmeth.3869.
- [14] Stat M, Huggett MJ, Bernasconi R, DiBattista JD, Berry TE, Newman SJ, et al. (2017). Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Sci Rep*, 7, 12240. doi:10.1038/s41598-017-12501-5.
- [15] Lynggaard C, Calvignac-Spencer S, Chapman CA, Kalbitzer U, Leendertz FH, Omeja PA, et al. (2023). Vertebrate environmental DNA from leaf swabs. *Curr Biol*, 33(16), R853-R854. doi:10.1016/j.cub.2023.06.031.
- [16] Ibrahim NAA, Azmir IA, Aqmal-Naser M, Esa Y, Ahmad A. (2025). From water to DNA: the evaluation of environmental DNA (eDNA) sampling volume in a tropical blackwater peat swamp in Dungun, Terengganu. *Malays J Fundam Appl Sci*, 21(5), 2662-2674. doi:10.1113/mjfas.v21n5.4246.
- [17] Farrell JA, Whitmore L, Mashkour N, Rollinson Ramia DR, Thomas RS, Eastman CB, et al. (2022). Detection and population genomics of sea turtle species via noninvasive environmental DNA analysis of nesting beach sand tracks and oceanic water. *Mol Ecol Resour*, 22(6), 2471-2493. doi:10.1111/1755-0998.13617.
- [18] Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*, 37, 852-857. doi:10.1038/s41587-019-0209-9.
- [19] Edgar RC. (2016). SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *bioRxiv*. doi:10.1101/074161.
- [20] Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. (2009). GenBank. *Nucleic Acids Res*, 37, D26-D31. doi:10.1093/nar/gkn723.
- [21] Nousias O, Duffy FG, Duffy IJ, McCauley M, Whilde J, Duffy DJ. (2026). Long-read nanopore shotgun metagenomic DNA sequencing for river biodiversity, wildlife, pollution, and environmental health monitoring. *NAR Genom Bioinform*, 8, lqag040. doi:10.1093/nargab/lqag040.
- [22] Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, et al. (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, 304(5667), 66-74. doi:10.1126/science.1093857.
- [23] Lu Y, Dong Y, Zhang M, Mao L. (2025). Genome and metagenome skimming: future sequencing methods for environmental DNA (eDNA) studies. *Mol Ecol Resour*, 0, e14095. doi:10.1111/1755-0998.14095.
- [24] Frank LE, Lindsey LL, Kipp EJ, Faulk C, Stone S, Roerick TM, et al. (2024). Rapid molecular species identification of mammalian scat samples using nanopore adaptive sampling. *J Mammal*, 105(4), 965-975. doi:10.1093/jmammal/gyae044.
- [25] Frankham R. (2005). Genetics and extinction. *Biol Conserv*, 126(2), 131-140. doi:10.1016/j.biocon.2005.05.002.
- [26] Newediuk L, Richardson ES, Bohart AM, Roberto-Charron A, Garroway CJ, Jones MJ. (2025). Designing epigenetic clocks for wildlife research. *Mol Ecol Resour*, 25, e14120. doi:10.1111/1755-0998.14120.
- [27] Skarzynska M, Leekitcharoenphon P, Hendriksen RS, Aarestrup FM, Wasyl D. (2020). A metagenomic glimpse into the gut of wild and domestic animals: quantification of antimicrobial resistance and more. *PLoS One*, 15(12), e0242987. doi:10.1371/journal.pone.0242987.
- [28] Teitelbaum CS, Casazza ML, Overton CT, Matchett EL, Prosser DJ. (2025). Host responses and viral traits interact to shape the impacts of climate warming on highly pathogenic avian influenza in migratory waterfowl. *PLoS Comput Biol*, 21(10), e1013451. doi:10.1371/journal.pcbi.1013451.
- [29] Wolfe JD, Luther DA, Jirinec V, Collings J, Johnson EI, Bierregaard RO, et al. (2025). Climate change aggravates bird mortality in pristine tropical forests. *Sci Adv*, 11(5), eadq8086. doi:10.1126/sciadv.adq8086.

- [30] Urban L, Holzer A, Baronas JJ, Hall MB, Braeuninger-Weimer P, Scherm MJ, et al. (2021). Freshwater monitoring by nanopore sequencing. *eLife*, 10, e61504. doi:10.7554/eLife.61504.
- [31] Nousias O, McCauley M, Stammenitz MR, Farrell JA, Koda SA, Summers V, et al. (2025). Shotgun sequencing of airborne eDNA achieves rapid assessment of whole biomes, population genetics and genomic variation. *Nat Ecol Evol*, 9(6), 1043-1060. doi:10.1038/s41559-025-02711-w.
- [32] Abrego N, Furneaux B, Hardwick B, Somervuo P, Palorinne I, Aguilar-Trigueros CA, et al. (2024). Airborne DNA reveals predictable spatial and seasonal dynamics of fungi. *Nature*, 631, 835-842. doi:10.1038/s41586-024-07658-9.
- [33] Aucone E, Kirchgeorg S, Valentini A, Pellissier L, Deiner K, Mintchev S. (2023). Drone-assisted collection of environmental DNA from tree branches for biodiversity monitoring. *Sci Robot*, 8, eadd5762. doi:10.1126/scirobotics.add5762.
- [34] Norouzzadeh MS, Nguyen A, Kosmala M, Swanson A, Palmer MS, Packer C, et al. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proc Natl Acad Sci U S A*, 115(25), E5716-E5725. doi:10.1073/pnas.1719367115.
- [35] Wiggins JJ, Tobias VD, Holcombe EF, Karpenko K, Huber ER, Goodman AC. (2024). Leveraging environmental DNA (eDNA) to optimize targeted removal of invasive fishes. *J Freshw Ecol*, 39(1), 2378841. doi:10.1080/02705060.2024.2378841.
- [36] Stammenitz MR, Hartman Scholz A, Duffy DJ. (2024). Environmental DNA without borders. *EMBO Rep*, 25, 4095-4099. doi:10.1038/s44319-024-00264-w.
- [37] Akrim F, Mahmood T, Max T, Nadeem MS, Qasim S, Andleeb S. (2018). Assessment of bias in morphological identification of carnivore scats confirmed with molecular scatology in north-eastern Himalayan region of Pakistan. *PeerJ*, 6, e5262. doi:10.7717/peerj.5262.
- [38] Shehzad W, Riaz T, Nawaz MA, Miquel C, Poillot C, Shah SA, Pompanon F, Coissac E, Taberlet P. (2012). Carnivore diet analysis based on next-generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. *Mol Ecol*, 21(8), 1951-1965. doi:10.1111/j.1365-294X.2011.05424.x.
- [39] Carneiro CM, Shields-Estrada A, Boville AE, Alves-Ferreira G, Xu T, Wong Arnott RL, et al. (2025). Toward a global science of conservation genomics: coldspots in genomic resources highlight a need for equitable collaborations and capacity building. *Mol Ecol*, 34, e17729. doi:10.1111/mec.17729.
- [40] Vinay KL, Goyal N, Warudkar A, Arvind C, Robin VV. (2026). Genome assemblies for seven families of birds from the Global South. *Mol Ecol Resour*, 26, e70162. doi:10.1111/1755-0998.70162.
- [41] Hayah I, Ezebuoro V, Kagame SP, Kuja JO, Waruhiu C, Nesengani LT, et al. (2025). Unlocking the African bioeconomy and strengthening biodiversity conservation through genomics and bioinformatics. *NPJ Biodivers*, 4, 29. doi:10.1038/s44185-025-00102-9.
- [42] Sandbrook C, Albury-Smith S, Allan JR, Bhola N, Bingham HC, Brockington D, et al. (2023). Social considerations are crucial to success in implementing the 30x30 global conservation target. *Nat Ecol Evol*, 7(6), 784-785. doi:10.1038/s41559-023-02048-2.



2026 by the authors; Journal of Kashmir Journal of Academic Research and Development. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).