

Oslo
Feb. 8-10, 2023

Norwegian Biodiversity & Genomics Conference 2023

Keynotes, Abstracts, & Meeting Info

Quick links:

[February 9th, 2023 - Morning session \(0900-1200\)](#)

[February 9th, 2023 - Afternoon session \(1300-1600\)](#)

[February 9th, 2023 - European Reference Genome Atlas \(ERGA\) Open Meeting \(1600-1700\)](#)

[February 10th, 2023 - Morning session \(0900-1200\)](#)

[February 10th, 2023 - Afternoon session \(1300-1600\)](#)

This meeting is supported by UiO:Life Science and:



EBP-Nor



InvertOmics



February 9th, 2023 - Morning session (0900-1200)

Session chair: Benedicte Garmann-Johnsen

KEYNOTE: Dr. Mark Blaxter, Sanger Institute

New biology from chromosomally complete genomes at the Tree of Life



Mark Blaxter is Head of the Tree of Life Programme at the Wellcome Sanger Institute. ToL is generating and analysing reference genome sequences from thousands of species across the tree of life, with initial focus on the eukaryotic biota of Britain and Ireland (Darwin Tree of Life Project), and the genomes of symbiotic organisms (Aquatic Symbiosis Genomics project). His research portfolio focuses on using reference genomes to understand pattern and process in genome evolution, with an abiding interest in nematodes and other meiofauna. He was elected a Fellow of the Royal Society of Edinburgh in 2014.

Ana Maria Peris Tamayo, Nord University

Genomic architecture of parallel phenotypic divergence in the marine snail *Littorina saxatilis* along shore levels and latitudinal locations

Understanding the ecological and physical factors driving intraspecific divergence and the origin of species is the “holy grail” of evolutionary biology. Parallel ecological divergence can occur in heterogeneous environments, showing different patterns at the genotype level under parallel phenotypic evolution. In this study, we focused on the genomic basis of temperature adaptation in the marine snail *Littorina saxatilis* in response to low-shore and high-shore from different latitudinal locations to understand the spatial distribution of adaptive genetic variation. We tested whether the same set of loci underlie shore level temperature adaptation in different geographical locations. We used pooled whole-genome resequencing for determining the allele frequency variation among locations. We found that while some genomic regions showed location-specific patterns, others are likely to contribute to adaptive divergence across locations. We also found overlap between these regions and candidate chromosomal inversions. Chromosomal inversions can increase differentiation across large stretches of the genome and control divergent phenotypes, which can be associated with local adaptation in the presence of gene flow, favouring rapid parallel adaptation to heterogeneous environments. Chromosomal inversions can play an important role in adaptation and speciation as has been seen in other polymorphic species.

Alberto Valero-Gracia, Natural History Museum Oslo

HiFi de novo sequencing and genome architecture dynamics in nemerteans

Nemertea, a predominately marine phylum composed of over 1,300 species, is a Spiralia/Lophotrochozoa clade that can be subdivided into three major groups: Palaeonemertea, Hoplonemertea, and Pilidiophora. They are ambush hunters that use an eversible proboscis and toxins to capture prey. Moreover, nemerteans develop either directly or indirectly by primary larvae. This makes this clade a promising system for studies of toxin and life-cycle evolution. Despite its potential, nemerteans have so far been severely neglected from genomic studies with only two genomes of pilidiophoran species being published. To fill this gap, here we present six new high-quality genomes containing representatives of all three major groups. These genomes, generated de novo by means of PacBio HiFi sequencing and assembled with Hifiasm, were compared based on a whole-genome alignment together with the two publicly available Nemertea genomes. Our study is the first to support the phylogenetic relationships of the phylum with whole genome data and enables an understanding of the genome evolution dynamics within the clade. In this work we specifically investigate the evolution of selected gene clusters to address whether there is homeotic gene contiguity and whether there are major structural rearrangements of the genome architecture (i.e., gene flow modifications, transposable element behaviour, and introgression changes). In a broader perspective, our genome-based study aims to understand how conserved and lineage-specific genetic features shape evolution.

Kjetill Jakobsen, University of Oslo

The Earth BioGenome Project Norway - A national 'dugnad'

Biodiversity is increasingly threatened by human-related activities, directly and indirectly, which ultimately imperils both the global food supply and other biological resources that we depend upon. Therefore, understanding and preserving species and ecosystems are global imperatives for human survival and prosperity. Each organism and species on the planet has a unique "blueprint" encoded by its genome, i.e. its total DNA, which contains all the information that the species needs to survive, interact with its environment, and reproduce. Knowing and understanding the complete DNA sequence of all species of life on Earth will not only provide fundamental new insights into biology, biodiversity, and our ability to conserve life, but also deliver a comprehensive view of nature's toolkit for future biotechnology to provide humanity with food, medical treatment, drugs, vaccines, biofuels, and biomaterials. The Earth BioGenome Project (EBP) is a global non-profit initiative that aims to sequence and catalogue the genomes of all of Earth's 1.5 million currently described eukaryotic species over a period of ten years. EBP-Nor is the Norwegian initiative of EBP, and a partnership among the major universities in Norway (UiO, NMBU, UiB, NTNU, Uni Nord, and UiT), the research institute SINTEF, and the

non-academic institutions REVOcean, The Life Science Cluster, and ArcticZymes Technologies. Together we will sequence and catalogue all eukaryotic species occurring in Norway, an estimated 45,000 species.

Helle Tessand Baalsrud, Norwegian University of Life Sciences

The highly repetitive landscape of the snowy owl genome

Bird genomes are known to be relatively repeat-poor and stable in terms of size and gene synteny compared to other vertebrates. However, with the appearance of more avian chromosome-level assemblies we can now investigate the full spectrum of genomic repeats and evolution of genomic architecture in this lineage to test this assertion. We have generated a gold-standard genome assembly for the snowy owl (*Bubo scandiacus*) using PacBio, ONT and HiC as part of the Earth Biogenome Project Norway. We found that the relatively large snowy owl genome (1.6Gb) has the highest proportion of repeats reported for any bird genome to date (28.34% compared to an average of ~10% in other birds). The main bulk of the snowy owl genomic repeat landscape consists of transposable elements (superfamily endogenous retroviruses - ERVs), which in some chromosomes form large chunks of up to 5Mb of centromeric satellite DNA. These ERV-like centromere sequences are only found in the *Bubo* lineage and seem to be species-specific, they could therefore have implications for reproductive isolation and speciation in this lineage, as well as genomic stability. Our findings reveal a highly dynamic evolutionary history of the snowy owl genome characterized by high transposable elements activity and repeat expansions.

Bram Danneels, University of Bergen

A genomic approach to decode marine mammal toxicology

Large marine mammals fill important ecological niches in marine food webs. Their high energy intake is often accompanied with high intake of contaminants, especially for the top predators. They can thus act as good barometers for environmental pollution. Killer whales are top predators and by far the most contaminated marine mammal species, while considerably lower concentrations of pollutants are measured in filter-feeding baleen whales, and in polar bears. Many pollutants are endocrine and/or metabolic disruptors, and can disturb animal physiology, growth, reproduction, and immune response. Many (marine) organisms have developed ways to sense and deal with toxic chemicals. The collection of genes and proteins related to sensing and detoxification of chemical stressors is known as the chemical defensome. The chemical defensome has been studied in detail in teleost fish, and some key xenosensors (e.g. pxr, nr1i2) are missing in certain species (e.g. the Atlantic cod, *Gadus morhua*) with no obvious compensatory mechanisms. With many marine mammal genomes now available, their chemical defensome can be studied in more detail. Previous research showed functional loss of the pxr gene in cetaceans, while the gene is still functional in many other mammals, including polar

bears. We aim to combine comparative genomics approaches of chemical defense-related genes with toxicological studies on cell cultures of the killer whale (*Orcinus orca*), the fin whale (*Balaenoptera physalus*), and the polar bear (*Ursus maritimus*). This will provide a more complete insight into how marine mammals deal with the elevated levels of chemicals in their environment.

Lone Kvalheim, University of Oslo

Mapping the biodiversity of fish in the Oslo fjord using eDNA and Citizen science

The Oslo fjord has long been pronounced dying by locals, however documentation of the species richness and the loss thereof is much needed. The fjord is changing due to long-term pollution, fisheries, land use and temperature shifts. Traditional methods rely heavily on recorded observations which are beneficial in detecting abundance rich species prone to be caught in fishing nets or trawls, but are lacking for species adapted to hide, in low abundance, or at remote locations. eDNA methods have been presented as a viable solution to these shortcomings. This approach allows early detection of invasive species and saves economic efforts in management. Researchers present it as more cost-efficient and with less sampling efforts. When combined with citizen science, research is expected to excel as large areas can be covered in a short time. DNA information derived from the Oslo fjord is applied to metabarcoding using the MiFish primer, which identifies any fish species present in each sample. With samples of DNA and environmental parameters covering most of the Oslo fjord it is possible to map the biodiversity and locate any correlations to water quality. This will aid in creating suitable conservation measures along the fjord.

February 9th, 2023 - Afternoon session (1300-1600)

Session chair: José Cerca

KEYNOTE: Astrid Böhne, Zoological Research Museum Alexander Koenig

A tour of genomics in biodiversity research



With my team at the Leibniz Institute for the Analysis of Biodiversity change we are interested in the genomic underpinnings of biodiversity and study those to understand sex chromosome evolution in model systems of speciation (cichlid fish) and sex role reversal (pipefishes and seahorses), phenotypic novelties in vertebrates building on museum collections and traces of biodiversity decline. I am further involved in the ERGA project, especially in the first steps of the sample flow including developing standards for sample and sample metadata collection, biobanking and voucherizing. Topics to cover: sex chromosome evolution in cichlids, use of genomes in museum research and biodiversity monitoring, the importance of samples in the ERGA project

Kristina Stenlökk, Norwegian University of Life Sciences

Atlantic salmon pan-genome reveals hidden genomic variation impacting environmental adaptation

Structural variations (SVs) are widely recognized as a major source of genomic variation impacting adaptive evolution and species diversification. However, their functional importance is poorly understood largely because they have been challenging to detect with short-read sequencing. Here we present the first pan-genome for Atlantic salmon, comprising 11 long-read-based assemblies from all four phylogeographical lineages. We detected 1,061,452 SVs capturing 367 Mbp of sequence, wherein 13,038 SVs overlapped the coding sequence (CDS) of 2,725 unique genes, implicating that they directly affect gene functions. We genotyped 304,407 SVs using graph genome analysis and short-read data for 366 salmon sampled from contrasting riverine environments across the natural range. Genotype-environment association analyses revealed GO enrichment of neurological processes and highlighted 13 KEGG pathways including the GnRH signalling pathway known to be one of the main regulators of reproductive function in vertebrates. SVs overlapping CDS of genes associated with

environmental adaptation highlight the importance of several polymorphic immunoglobulin regions and an 18 kbp deletion on chromosome 28 encompassing a segmental duplication of three genes: luteinizing hormone subunit beta (LHB) involved in late stages of maturation, and two genes involved in cellular respiration; G-protein coupled receptor (GPR4) and Cytochrome C Oxidase Assembly Factor (COX20). Our results demonstrate how long-read pan-genomics can reveal previously hidden genomic variation and that these likely have consequences for fitness and adaptations in the wild. We anticipate the revealing of this previously hidden functional variation to inform future management of vulnerable wild populations of Atlantic salmon and contribute to sustainable aquaculture.

Laura Marian Valencia, University of Oslo

De-novo assembly and annotation of the crucian carp genome

The crucian carp (*Carassius carassius*) is a champion of anoxia tolerance and the physiological mechanisms are well documented. Still, there is a need for better understanding at the molecular and genetic level, and a fully sequenced, assembled and annotated genome will facilitate such studies. To obtain a high-quality genome assembly, we sequenced the crucian carp genome using methods such as chromosome conformation capture (HiC) and PacBio long-read sequencing, in addition to short-read sequencing for error correction. For annotation we are using full-length transcripts (IsoSeq data) as well as regular RNAseq data. The genome of crucian carp and other cyprinids consists of 50 chromosomes with a duplicated genome. The presence of duplicated genes offers the possibility for evolutionary changes through the mechanism of expression dominance: one paralog remains most expressed, while the other paralog is free to change without compromising the expression integrity of the biological systems required for normal function. We expect that paralogs of other important genes may have undergone similar changes, and aim to elucidate any differences that may have arisen between the crucian carp and the goldfish due to the long domestication history of the latter.

Halvor Knutsen, University of Agder

Combining population genomics with demographic analyses highlights habitat patchiness and larval dispersal as determinants of connectivity in coastal fish species

Geneflow, plays a significant role in shaping the genetic structure and potential for adaptation in local populations. In marine animals with non-migratory adults, the presence or absence of a pelagic larval stage is believed to greatly influence gene flow and population genetic structure. Additionally, the availability of suitable habitats can impact the distribution and connectivity patterns of populations. Using whole genome sequencing and reduced representation sequencing, we studied the demographic history and genetic structure of the broadnosed pipefish (*Syngnathus typhle*) in southern Norway. Its main habitat is eelgrass beds, which are

patchily distributed along the study area in southern Norway. Demographic connectivity among populations was inferred from long-term (~30-year) population counts that uncovered a rapid decline in spatial correlations in abundance with distance as short as ~2 km. These findings were contrasted with data for two other fish species that have a pelagic larval stage (corkwing wrasse, *Symphodus melops*; black goby, *Gobius niger*) that showed wider spatial scales of connectivity and weaker genetic isolation-by-distance patterns, except where there was a strong barrier to gene flow due to a lack of suitable habitat. Our results confirm that a fragmented habitat and the absence of a pelagic larval stage promote genetic structure, while the presence of a pelagic larval stage increases demographic connectivity and gene flow, except possibly over large areas lacking suitable habitat.

Aruna M. Shankregowda, Nord University

Contribution of epigenetic variation to local adaptation in two coexisting and phylogenetically related species

The capacity of species to respond to environmental change largely depends on evolutionary potential, which is rooted in the amount and structure of adaptive genetic variation. However, species may compensate for low evolutionary potential through phenotypic plasticity. Adaptive plastic responses to selective forces may be modulated by the regulation of gene expression, for instance via epigenetic modifications such as DNA methylation. In this study, we investigated the role of DNA methylation in local adaptation of coexisting populations of three-spined (*Gasterosteus aculeatus*; 3S) and nine-spined (*Pungitius pungitius*; 9S) stickleback across the Belgian-Dutch lowlands. We previously showed that the two species differ markedly in the strength and nature of local adaptation, where brackish water (BW) and freshwater (FW) populations show stronger phenotypic and genomic differentiation in 3S than in 9S. We hypothesize that differential DNA methylation is more pronounced in 9S, as this may compensate for the lack of genomic differentiation across populations from contrasting salinities. Correcting for sequence effort and number of mapped reads, we show similar proportions of methylated SNPs in 3S (0.0239) and 9S (0.0245). Yet, contrary to our expectation, the number of differentially methylated sites, especially among freshwater populations, was higher among 3S populations than among 9S populations. To shed further light on this issue, we plan to (1) correct the number of differential methylated sites for genetic differentiation; (2) identify the function of genes near differentially methylated sites; and (3) identify differentially methylated genomic regions involved in freshwater adaptation in the two species.

Ann-Christin Zinner, Nord University

Teleost genome architecture and the search for a gene loss - An argument for better genomes

We have previously reported a teleost specific distribution of intron sizes with a characteristic antimode at about 256 base pairs. This distribution is observed in the majority of annotated teleost genomes and is presumably a consequence of the processes underlying the evolution of genome architecture in teleosts. In most vertebrates, sites of meiotic recombination are largely determined by the activity of the PRDM9 protein. PRDM9 is absent from most known teleost genomes and we noted that this loss seems to be coupled to the appearance of the teleost intron size distribution. Here we present preliminary evidence of multiple independent losses of PRDM9 across the teleosts. All but one of these losses is associated with the teleost specific intron size distribution, suggesting that changes in recombination landscapes have contributed to the evolution of the teleost genome architecture. However, enumerating losses of genomic features is not straightforward as a failure to identify may reflect the assembly quality rather than true losses. This problem is particularly notable within the order of Gadiformes, where many apparent losses likely represent a failure of detection. Our work emphasizes the need for taxonomical diversification and quality improvements of genome assemblies as well as annotations to aid genome evolution research and argues for the importance of projects like the Earth BioGenome Project.

Sinan Ugur Umu, University of Oslo

MirMachine: accurate microRNA annotation of animal genomes using trained covariance models

While thousands of novel genomes are accessible now, no reliable methods exist to annotate them for microRNAs, an essential class of post-transcriptional regulators. Their prediction and annotation depend on the availability of transcriptomics data sets and hands-on expert knowledge, resulting in a massive disparity between the availability of novel genomes and high-quality microRNA complements. We constructed and trained covariance models for each conserved microRNA family using more than 16,000 microRNA entries from the manually curated microRNA gene database, MirGeneDB. MirMachine, our new pipeline for the automatic annotation of conserved microRNAs, now includes these models. We show that MirMachine can be used to accurately and precisely predict conserved microRNA complements from genome assemblies, properly identifying the number of paralogues. MirMachine, which was built and trained on representative metazoan microRNA complements, was used on a wide range of animal species, including those with very large genomes or additional genome duplications, as well as extinct species such as mammoths, where deep small RNA sequencing data will be difficult to produce. With precise predictions of conserved microRNAs, the MirMachine approach solves a long-standing gap in the microRNA field, which will not only facilitate automated genome annotation pipelines but can also serve as a model for future research.

February 9th, 2023 - European Reference Genome Atlas (ERGA) Open Meeting (1600-1700)

Session chairs: Rebekah Oomen (UiO) & Torsten Struck (NHMO)

Agenda:

1. Introduction to the European Reference Genome Atlas (ERGA)
2. Introduction to the Biodiversity Genomics Europe Consortium and its connection to ERGA
3. Call for membership to ERGA and those to form a core working group towards establishing ERGA-Norway
4. Notification of upcoming election for new ERGA Council Member for Norway
5. Any other business



February 10th, 2023 - Morning session (0900-1200)

Session chair: Helle Baalsrud

KEYNOTE: Inger Skrede, University of Oslo

Mating, divergence and speciation in wood decay fungi



Our research group investigate questions in the field of mycology, fungal evolutionary genomics, systematics, biodiversity of microorganisms in various habitats, and ecology. In this talk I will mainly focus on our research on fungal speciation, the evolution of reproductive barriers and fungal mating. We are using the genus *Trichaptum*, common wood decay fungi, as our model organism in this research. By using approaches as comparative genomics, population genomics and common garden experiments of an extensive collection from all over the northern hemisphere, we have detected an extensive mating type diversity, and a variety of responses of secondary contact after the last glaciations. Our research will provide deeper insight in how the immense fungal diversity originate and are maintained in a changing climate.

Sara Orellana, University of Oslo

Yeast bioprospecting and genomics for a food circular bioeconomy

The fungus kingdom encompasses an enormous phylogenetic, phenotypic, and genomic diversity suggesting distinct ecological roles. However, despite their importance and ubiquity, the distribution of most fungi species remains poorly understood. In particular, relatively little attention has been paid to fungal species ecology in Scandinavia. Here we address the identification of fungal species belonging to different areas in Norway for evaluating the effects of temperature and substrate on their distribution. Samples of substrates (soil, plants, insects...) were collected in natural areas from Svålbard (north) and Oslo (south), although the inclusion of other Nordic intermediate regions is expected further. Isolated fungi were classified according to their morphology (color or microscopic features) or physiology traits (e.g., growth at a specific isolation temperature). Taxonomy was assigned through amplification and sequencing of the ITS region of the rDNA locus. Finally, we examined the diversity of species associated with different substrate types to better understand where fungus communities are expected to be especially complex. Preliminary data suggest a strong effect of location (north or south), probably due to temperature, on the phyla of fungi recovered. Remarkably, we found some taxa

non-previously isolated from sampled regions of Norway as a sign that it is a poorly studied niche with potentially valuable species for the development of bio-industrial capabilities. The sequencing and characterization of the newly isolated species (with a special interest in yeasts) can shed light on the improvement of phenotypic traits associated with industrial bioprocesses, especially in the production of healthy, safe, and sustainable foods.

Dabao Lu, University of Oslo

*Population genomics reveal multiple lineages with leaky intersterility barriers in the widespread forest fungus *Trichaptum abietinum**

The majority of species descriptions in fungi are based on fruit body morphology, and in recent years the species status of many of these morphospecies have been called into question by DNA based surveys which have revealed distinct phylogenetic and geographically restricted lineages within them. In many cases, crossing studies have also detected reproductive barriers within fungal morphospecies. In this study we use population genomic analyses together with in vitro crosses of the widespread wood decay fungus *Trichaptum abietinum* to determine the genetic structure within this morphospecies and infer the processes that have shaped it. To this end, we have whole genome sequenced 350 *T. abietinum* samples from Asia, Europe and North America. Our phylogeographic analyses show six major lineages, one in Asia, two in Europe, and three in North America. The two lineages present in Europe are interfertile and admixing, whereas our crosses show that the North American lineages are reproductively isolated. In Asia a more complex pattern appears, with partial intersterility between multiple sub-lineages. We found reduced wood decay for in vitro hybrids to be strongly correlated with the genomic divergence between parents, consistent with the accumulation of Bateson-Dobzhansky-Muller incompatibilities. On the other hand, pre-mating barriers were moderately correlated with genomic divergence, suggesting that other mechanisms have been involved in generating these barriers in North America and Asia. Demographic modelling and phylogenetic network analyses support that these pre-mating barriers in *T. abietinum* were reinforced upon secondary contact between lineages that diverged in allopatry during the Pleistocene glacial cycles.

Lik Theng Ho, University of Agder

Microbial summer community of a South Norwegian fjord

Fjord systems along the southeast Norwegian coastline are influenced by multiple stressors including increasing temperature, eutrophication, acidification and coastal darkening. Although marine microbes play major roles in the functioning of coastal ecosystems, our understanding of their community structure and response to environmental change is often limited. This study aims to establish a baseline for characterising bacterioplankton and pelagic protist summer communities in Spindsfjorden, Farsund, South Norway. Seawater was sampled bi-weekly from the surface, 15m and the deep (90m) from end of May to August 2019, and was filtered through

Sterivex filters to collect organisms > 0.2µm. DNA was extracted from the filters for metabarcoding of the bacterioplankton and protist communities using partial 16S and 18S nrDNA, respectively. Metabarcoding data were processed using the dada2 bioinformatics pipeline. Net hauls were collected from 10-0m using a 10µm net and protists were identified using microscopy. Vertical profiles of temperature, salinity, oxygen and fluorescence were obtained using a CTD with oxygen and fluorescence sensors. Preliminary analyses suggested that while the deep bacterial communities were distinctly different from the shallower samples (surface and 15m), the protist communities showed a differentiation also between the surface and 15m samples. Both microscopic and metabarcoding data suggested that the diatoms *Skeletonema* sp., *Cerataulina* sp. and *Tripos* spp. were important members of the summer phytoplankton community. Here we discuss the summer physical conditions and its influence on the bacterial and protist communities of a South Norwegian fjord in 2019. This is the first study from a recently established coastal time series in Farsund, South Norway.

Solomon Rorellien, University of Oslo

Accumulation and effects of microplastics in a sediment community

There is a general assumption that plastics in the oceans to a large extent will end up in sediments. This study aimed to clarify whether sediment-dwelling organisms at different trophic levels will accumulate microplastics (MPs) and also whether the presence of microplastics will affect the species composition. The experiment was conducted using 24 mesocosms with sediment collected in the outer Oslofjord. Continuous water supply and air pumps were supplied to each mesocosm throughout stabilization and exposures. The mesocosms were divided in four treatment groups and dosed with 1 µm and 6 µm fluorescent polystyrene beads, as well as beach plastics from the Adriatic Sea (100-200 µm). Following termination, sediment core samples were taken for microfauna and meiofauna analysis. The remaining sediment in the boxes was sieved in cold seawater to extract all visible organisms. Samples were kept at 20°C to keep them alive until further analysis, and the sediment for barcoding were kept in freezer (-80°C). Bacterial eDNA is extracted for DNA metabarcoding from sediments to analyze possible differences in species composition in the different treatments. For meiofaunal community, the more time sensitive eRNA extraction will be employed to ensure analysis of DNA of only live meiofauna at the time of termination. Fluorescence microscopy along with dissection is used to visually identify and semi-quantitatively assess MP accumulation in organisms at different trophic levels. Preliminary results of fluorescent microscopy indicate that organisms in meiofauna have accumulated microplastics.

James Fleming, Natural History Museum Oslo

Scoutknife: A naïve, whole genome informed phylogenetic robusticity metric

The genomics revolution completely altered our understanding of phylogeny - the study of the relationships between animals. The identification of new cryptic species through phylogenetic methods that combine molecular and morphological data has led to an appreciation of phylogenetic relatedness as a biodiversity concept that can inform and shape further sampling, conservation priorities and environmental policy. Currently, however, phylogenetics has adopted a conservative stance towards this deluge of “big data”, selecting genes that fit within our models of evolution. This has the benefit of evading artifactual topologies that violate these models, but may present us with hypotheses of evolution preselected to our own biases, or cause us to adopt confidence in hypotheses that are not as well supported by the data as it first appears. This can have a powerful impact on our understanding of the relationships between taxa, which in turn could bias future sampling or conservation efforts. From this perspective, new methods optimized to make the most of “big data” may be of critical importance: here we present Scoutknife, a new method for establishing phylogenetic support. Based on the traditional bootstrap approach, Scoutknife discards the creation of artificial pseudoreplicates to instead use large multi-gene inputs to create true replicates. It produces comparable results to selecting the “best” genes from a phylogeny on a number of classical datasets and proves to be more granular in its assessment of topological reliability than traditional bootstrap values, allowing researchers to be more informed about their phylogenetic hypotheses than ever before.

Joost Raeymaekers, Nord University

COST Action CA18134 - Genomic Biodiversity Knowledge for Resilient Ecosystems (G-BIKE)

G-BiKE is a network of researchers and practitioners promoting the use of genomic diversity knowledge in nature and conservation management. G-BiKE enables standard and routine tools for assessing, monitoring and managing the genetic resilience and related adaptive potential of wild and captive populations. Although genetic data can be obtained for most organisms, the standardization of protocols for detecting and monitoring species' genetic diversity, and their potential for adaptation, is still lacking. G-BiKE therefore assists scientists and practitioners across the EU and particularly in COST Inclusiveness Target Countries to integrate genetic and evolutionary knowledge into conservation planning policies, and to promote cross-border management and long term monitoring programs of evolutionary potential in order to ensure persistence of populations and species, and ultimately the continued supply of nature-based ecosystem services. Specifically, G-BiKE aims to bring together scientists and practitioners from a diversity of countries to: (1) clearly articulate for managers how genetic diversity can support ecosystems; 2) develop and test best practice protocols for monitoring genetic diversity in time and space; 3) provide an online forum on emerging tools; 4) connect all stakeholders through networking and training opportunities; 4) build a network of conservation genetics labs; 5) build a foundation for long term impact. G-BiKE disseminates its achievements as easy-to-read summaries for nature managers and outreach to the general public at Natura 2000 sites,

botanic gardens and zoos, as well as in scientific publications, different media outlets, and social platforms.

February 10th, 2023 - Afternoon session (1300-1600)

Session chair: Thomas Marcussen

KEYNOTE: Paul Kersey, Royal Botanical Gardens, Kew

Natural history collections in the digital age



I began my career as a molecular biologist working on fission yeast before turning to bioinformatics. After working at the European Bioinformatics Institute on various databases including Ensembl, I moved to the Royal Botanic Gardens, Kew, where I am Deputy Director of Science, responsible for our research programme and currently overseeing the digitisation of Kew's eight million biological specimens. My personal interests focus on the application of genomics to biodiversity, for example, through the PAFTOL and Darwin Tree of Life projects. In this talk I will discuss how genomics and digital technologies are revitalising natural history collections and reinforcing their relationship to cutting edge scientific research.

Youri Lammers, University of Tromsø

The PhyloNorway project: a Norwegian vascular plant genome reference set and applications

Large vascular plant genomic reference datasets are invaluable for biodiversity assessments and environmental studies. This is especially true for the Polar regions, where the warming conditions are causing significant changes in the Arctic plant communities. Here we present the PhyloNorway vascular plant reference dataset, consisting out of 2000+ specimens of 1900 vascular plant taxa native to Norway and the Polar regions and which covers the majority of taxa present. The reference dataset is based on our low coverage genome skimming efforts of collection material from the Tromsø museum herbarium. Each sample was sequenced deep enough for the subsequent assembly of the high copy regions, such as the nuclear ribosomal DNA and chloroplast genome. The PhyloNorway reference set has proven to be an useful resource for the identification of palaeoecological data such as sedimentary ancient DNA (sedaDNA) records. DNA barcoding regions extracted from the assemblies have been used for the identification of metabarcoding material from Norway, while the assembled genomes have drastically improved the identifications of vascular plants from shotgun sedaDNA. Furthermore,

the available genetic material allows for the development of additional metabarcoding markers or target capture probe sets to further improve the taxonomic resolution of key taxonomic groups.

Miguel Naranjo-Ortiz, Natural History Museum Oslo

Museomics as a way to anchor reference databases with taxonomic knowledge

Museum collections preserve and categorize biological material that serves as taxonomic reference, or type materials, for the description of the world's biodiversity. In the age of genomics, species identification is often done by matching query sequences derived from a sample versus previously generated references in a database. The combined work of researchers across the world have created the enormous databases we nowadays can publicly access. However, a lot of work remains to be done in order to link the biological samples from where these sequences came to taxonomic type materials. Type specimens contain DNA that can be extracted and analyzed, although many are decades, or even centuries old. In this talk I summarize recent advances in our plan to generate reference sequence data for all specimens in the Herbarium and Fungarium of the University of Oslo and discuss future prospects in our daunting quest.

Adam Vivian-Smith, Norwegian Institute of Bioeconomy Research

Forest genetics and biodiversity from a long-read perspective

Terrestrial forest ecosystems provide rich biodiversity niches. They are highly heterogeneous and can harbor substantial red list species, but these environments may also host many species which are data deficient or cryptic. Some boreal forests are also unmatched in soil carbon depositions, especially when considering the above ground standing biomass and productivity. Genomic research is required to understand effective population sizes, their historical associations and their evolutionary ecology. This research is absolutely necessary to manage and mitigate increasing pressures of forestry related activities. We are using Oxford Nanopore long-read sequencing to resolve some genomes of red and black list species, their congenitors, as well as poorly understood forest pathogens. One particular challenge is extracting high molecular weight DNAs from diverse, and sometimes limited samples, and then sequencing these accurately. Here we provide an example of telomere to telomere assemblies (chromosomal scale) from a few species using Nanopore, and/or hybrid assembly methods (fungi, lichen and plants). Reads >200 kb and read length N50 values of 30-50 kb, are often provided by Nanopore data. Success depends on the choice of the DNA extraction method and short read elimination. Recent improvements to Nanopore sequencing chemistry provide high read N50 values and more accurate reads that greatly assist genome assembly and quality. Having accelerated access to genomic data will answer a broad array of questions. Overall this

will reduce uncertainty and help risk assessments, and lay out a management path or tools for protecting vulnerable forest species and mitigate the effects of invasive species. Insights are also offered into the modes of habitat fragmentation, sexual isolation and mating systems for entire groups of species.

Lluís Garcia Mir, Real Jardín Botánico - CSIC

Genome-wide analysis and floral expression of CRABS CLAW (YABBY TF family) in Malva-lineage species with contrasting pollination syndromes

The NEXTPOL project aims at the study of the genomic bases of floral traits associated to pollination. In the *Malvaceae*, floral reward to pollinators mostly relies on nectar production to ensure the mutualistic bind. Within the *Malva* lineage, entomophily is the predominant pollination syndrome, but *Navaea phoenicea*, a bird-pollinated threatened endemic to Tenerife, shows distinctive floral nectaries with a higher nectar output. Here, by analyzing genome-wide and floral transcriptome data of *N. phoenicea*, *Malva wigandii* and *M. canariensis*, we identify and annotate the YABBY gene family of transcription factors, responsible for many essential roles in flower development. Our study sheds light on the sequences and expression of CRABS-CLAW (CRC) candidates, a YABBY gene of conserved function throughout eudicots, whose expression is known to determine the nectary identity and development in *Arabidopsis* and *Gossypium*. In each of the species studied, we have found a YABBY gene family with up to 24 genes, of which two are CRC members in *N. phoenicea* and *M. canariensis* and three in *M. wigandii*. One CRC is ortholog to all *Malvaceae*, while the other one is only shared with *Hibiscus*. Our genomic results reveal an exon/intron structure, domain conservation and sequence diversity of CRC maintained across *Malvaceae*. Transcriptomics data covering all stages of flower development suggests in both entomophilous and ornitophilous species that CRC expression accompanies nectary formation in the calyx early developmental stage. Yet, this interpretation needs validation by RNA in-situ hybridization to further address the role of CRC in the inception of the nectary.

Dario Ojeda, Norwegian Institute of Bioeconomy Research

Asymmetric niche lability across lineages and timescales underlies the evolution of rainforest legume trees in Africa

The Guineo-Congolian region in Africa constitutes the second largest area of tropical rainforest (TRF). It covered an estimated 15-22 million km² during the late Miocene (55-11 Ma) and it has experienced a declining trend, currently reaching 3.4 million km². The overall contraction in TRF area over time is associated to increasing aridification and the replacement of TRF by savanna habitats (e.g., miombo). Here we aim to examine whether rainforest area contraction led to a decrease in net diversification rates linked to increasing extinction, or if it is associated to increasing opportunities for allopatric or ecological speciation during periods of forest

fragmentation. We used a target enrichment approach combined with a complete data set representing all 16 genera within the Berlinia clade (75% species) combined with niche modelling of habitat types. We recovered fully resolved and well supported relationships. Most genera (87.5%) diverged before the Pleistocene and we inferred seven independent shifts to savanna, but with an asymmetric distribution, suggesting a degree of lability to shift habitat in some genera. This indicates the Berlinia clade displays an overall trend of accumulation of species over evolutionary time, suggesting that the reduction of TRF area has not decreased net diversification rates. Most habitat shifts to savanna occurred in the Miocene, with no major habitat shifts during the most recent phases of forest expansion-contraction in the Pleistocene. Shifts in habitat from lowland forest to savanna did not trigger diversification rates, but habitat fragmentation might have increased diversification rates through allopatric speciation.

José Cerca, University of Oslo

Separating subgenomes in allopolyploid assemblies using fossile transposable elements and conserved orthologs

In only 5 years, the standard for large genome assemblies has moved from low-quality and fragmented to chromosome-resolved. Despite this advance, high-quality assemblies of non-model polyploid organisms, which often have large genomes and are characterized by high repeat-content, are still rare. In this contribution, I present a practical method to separate subgenomes in an allopolyploid chromosome-resolved assembly for non-model organisms where information on ancestral species is often lacking. This involves: 1) the identification of homeolog chromosomes by identifying orthologs and mapping them to their chromosome location; 2) subgenome characterization by the identification and quantification of transposable elements that were active before the polyploidization event (i.e. when the subgenomes were separated) – termed ‘fossil transposable elements’. I apply this method to our recent assembly of the *Scalesia atractylloides* (Asteraceae, Galápagos adaptive radiation) genome, with a size of 3.2 Gbp and with 76% repeat content. After separation of subgenomes, our analysis uncovered large inter-subgenome inversions and similar numbers of genes, isoforms and transposable elements on both subgenomes, indicating a lack of subgenome dominance. I estimated the speciation event to be around 4.2 million years ago, and the polyploidization event to have occurred 3.6 million years ago. The timing of polyploidization is concordant with the geological age of the Galápagos, and I speculate that a lack of subgenome dominance benefits plants’ colonization of remote islands, where increased genetic variation is needed to survive and establish.

February 10th, 2023 - Earth BioGenome Project Norway (EBP-Nor) Meeting (1600-1700) *Professorboligen*

Session chair: Kjetill Jakobsen

This meeting is invite-only.