

# Survey of Species Covered by DNA Barcoding Data in BOLD and GenBank for Integration of Data for Museomics



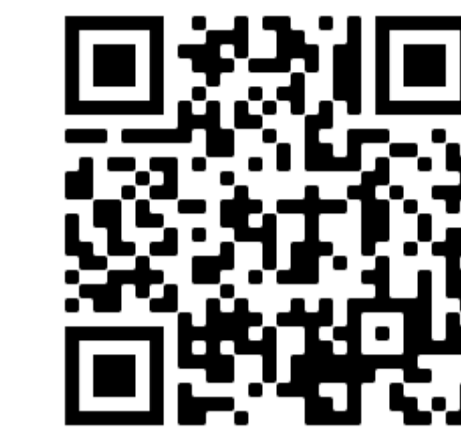
**Takeru Nakazato**

Database Center for Life Science (DBCLS),  
Research Organization of Information and Systems (ROIS),  
JAPAN

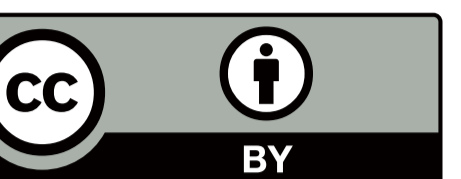
[nakazato@dbcls.rois.ac.jp](mailto:nakazato@dbcls.rois.ac.jp)

@tkr\_nak

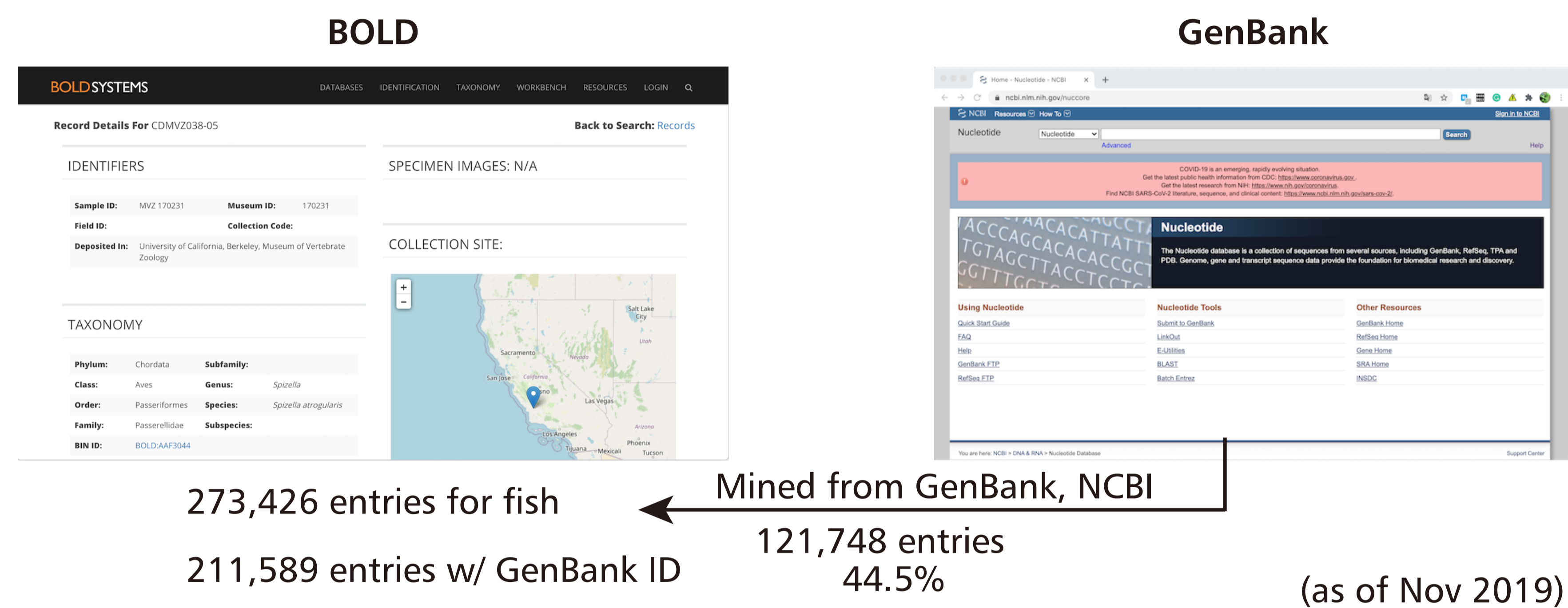
<https://orcid.org/0000-0002-0706-2867>



<https://doi.org/10.3897/biss.4.59065>

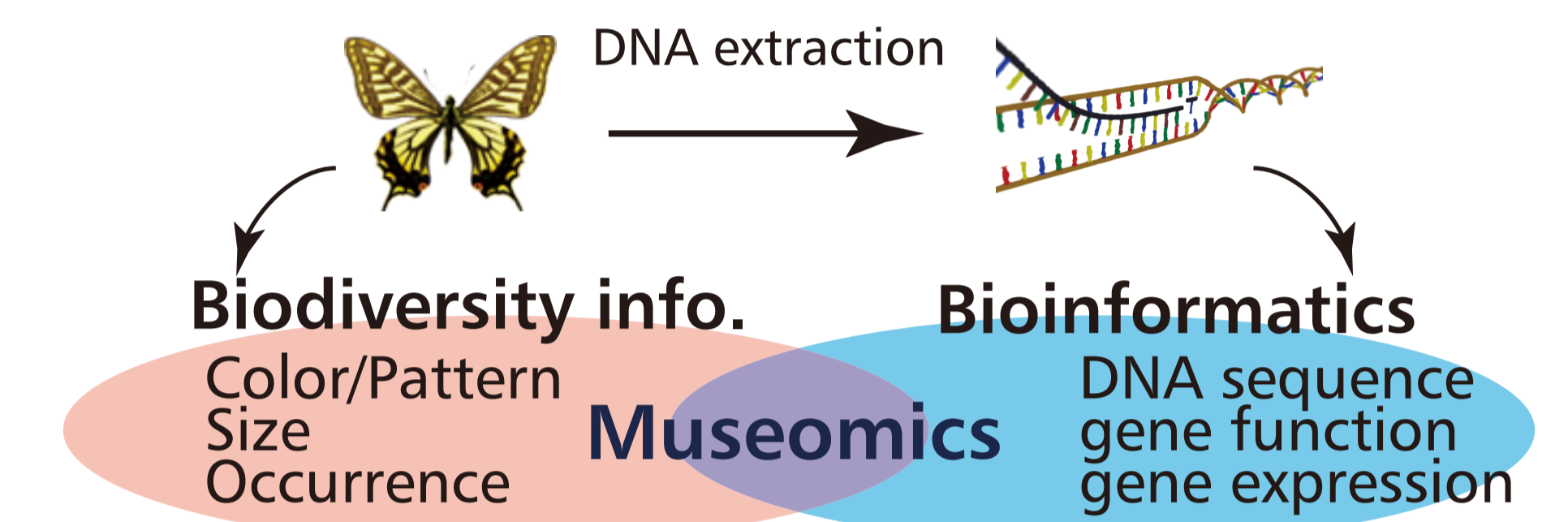


## Two databases for DNA Barcoding



## Summary

- ▶ We, DBCLS, attempt to integrate and standardize sequence data with Semantic Web technology.
- ▶ DNA Barcoding data is registered separately for BOLD and GenBank. It would be convenient if these data could be used together.
- ▶ In the future, we will be working with a variety of DNA, including genomes from museum specimens as well as DNA barcoding. This will require more integrated use of biodiversity information and DNA sequence data.



## Barcoding/voucher descriptions in GenBank

```

LOCUS      DQ433192          599 bp    DNA     linear   VRT 14-JUL-2016
DEFINITION Spizella atrogularis voucher MVZ:Bird:170231 cytochrome oxidase
            subunit 1 (COI) gene, partial cds; mitochondrial.
ACCESSION  DQ433192
VERSION    DQ433192.1
KEYWORDS   BARCODE.
SOURCE     mitochondrion Spizella atrogularis (Black-chinned sparrow)
ORGANISM   Spizella atrogularis
...
FEATURES   Location/Qualifiers
            source          1..599
                        /organism="Spizella atrogularis"
                        /organelle="mitochondrion"
                        /mol_type="genomic DNA"
                        /specimen_voucher="MVZ:Bird:170231"
                        /db_xref="BOLD:CDMVZ038-05"
                        /db_xref="taxon:40208"
                        /country="USA: California"
                        /lat_lon="35.3578 N 120.305 W"
                        /collection_date="20-May-1984"
                        /collected_by="K. Ned"
                        /PCR_primers="fwd_seq: ttctccaaccacaaagacattggcac,
                        rev_seq: acgtgggagataattccaaatcctg"
            gene            <1..599
                        /gene="COI"
            CDS             <1..599
                        /gene="COI"
                        /codon_start=1
                        /transl_table=2
                        /product="cytochrome oxidase subunit 1"
                        /protein_id="ABK29354.1"
                        /translation="MVG TALS..."
ORIGIN
1 atagtaggta ccgcctaag ctcctcatt cgagcagaac taggccaacc cggagccctc
61 ...
    
```

## How many species does BOLD/GenBank cover?

