

## **Supplemental Materials**

### **Appendix 1**

#### **Policy issues**

- The G10K consortium members recognize this endeavor to be an open access, noncommercial research program without “reach-through” intellectual property (IP) agreements.
- All individual(s)/institute(s) contributing biomaterials for sequencing thus comply with the above mentioned noncommercial policy and acknowledge that contributed samples have been collected and distributed in accordance with applicable national and international laws and regulations.
- The reference materials used for genome sequencing should be clearly identified, accessible and where possible, adhere to best practices for biodiversity repositories (particularly concerning taxonomic identification of voucher specimens/tissues and genome size estimation (Hanner and Gregory 2007)). Reference documentation should include digital images/CT scans of the morphological voucher specimens, geospatial coordinates of the specimen collection site (Field 2008) and other emerging metadata standards proposed for genome sequences (Field et al., 2008).
- Deposition and publication will follow the Ft Lauderdale agreement (genome.gov 2003) in accord with prior best practices in genomic research, including release of

primary sequence data to the international databases within 24 hours and release of the full genome assemblies as soon as quality assurance is complete.

- The consortium is committed to building programs for training of highly qualified personnel, particularly in developing nations.

## Appendix 2

### Detailed table of classes, orders, families, genera and species

Groups	Orders	Families			Genera			Species		
		G10K samples available	Total	% of Total	G10K samples available	Total	% of Total	G10K samples available	Total	% of Total
Mammals	Afrosoricida	2	2	100%	12	19	63%	18	51	35%
Mammals	Carnivora	15	15	100%	98	126	77%	163	287	56%
Mammals	Cetartiodactyla	21	21	100%	120	129	93%	238	324	73%
Mammals	Chiroptera	18	18	100%	115	203	57%	316	1116	28%
Mammals	Cingulata	1	1	100%	4	9	44%	5	21	23%
Mammals	Dasyuromorphia	2	2	100%	19	23	83%	55	71	77%
Mammals	Dermoptera	1	1	100%	1	2	50%	1	2	50%
Mammals	Didelphimorphia	1	1	100%	12	17	70%	34	87	39%
Mammals	Diprotodontia	11	11	100%	34	39	87%	88	143	61%
Mammals	Eulipotyphla	4	4	100%	28	55	51%	54	452	11%
Mammals	Hyracoidea	1	1	100%	3	3	100%	4	4	100%
Mammals	Lagomorpha	2	2	100%	7	12	58%	21	91	23%
Mammals	Macroscelidea	1	1	100%	3	4	75%	4	15	26%
Mammals	Microbiotheria	1	1	100%	1	1	100%	1	1	100%
Mammals	Monotremata	2	2	100%	3	3	100%	4	5	80%
Mammals	Notoryctemorphia	1	1	100%	1	1	100%	2	2	100%
Mammals	Paucituberculata	1	1	100%	3	3	100%	3	6	50%
Mammals	Peramelemorphia	2	3	66%	6	8	75%	12	21	57%
Mammals	Perissodactyla	3	3	100%	6	6	100%	16	17	94%
Mammals	Pholidota	1	1	100%	1	1	100%	6	8	75%
Mammals	Pilosa	4	4	100%	5	5	100%	8	10	80%
Mammals	Primates	14	15	93%	54	69	78%	146	376	38%
Mammals	Proboscidea	1	1	100%	2	2	100%	3	3	100%
Mammals	Rodentia	30	33	90%	219	481	45%	612	2277	26%
Mammals	Scandentia	2	2	100%	3	5	60%	9	20	45%
Mammals	Sirenia	2	2	100%	2	3	66%	2	5	40%
Mammals	Tubulidentata	1	1	100%	1	1	100%	1	1	100%
<b>Mammals Count</b>	<b>27</b>	<b>145</b>	<b>150</b>	<b>97%</b>	<b>763</b>	<b>1230</b>	<b>62%</b>	<b>1826</b>	<b>5416</b>	<b>34%</b>
Birds	Accipitriformes	2	2	100%	51	72	71%	121	240	50%
Birds	Aegothelidae	1	1	100%	1	2	50%	2	9	22%
Birds	Anseriformes	3	3	100%	41	52	78%	130	162	80%

Groups	Orders	Families			Genera			Species		
		G10K samples available	Total	% of Total	G10K samples available	Total	% of Total	G10K samples available	Total	% of Total
Birds	Apodiformes	3	3	100%	99	125	79%	241	429	56%
Birds	Caprimulgiformes (Part 1, 2,3)	4	4	100%	16	20	80%	47	109	43%
Birds	Cariamidae	1	1	100%	1	2	50%	1	2	50%
Birds	Charadriiformes	14	17	82%	70	88	80	226	367	62%
Birds	Ciconiiformes	3	3	100%	32	39	82%	72	116	62%
Birds	Coliiformes	1	1	100%	2	2	100%	5	6	83%
Birds	Columbiformes	1	2	50%	31	44	70%	136	311	43%
Birds	Coraciiformes (Part 1, 2)	9	10	90%	36	50	72%	103	208	50%
Birds	Cuculiformes	1	1	100%	22	36	61%	57	138	41%
Birds	Falconiformes	1	1	100%	9	11	82%	40	64	63%
Birds	Galliformes	5	5	100%	58	80	72%	135	290	46%
Birds	Gaviiformes	1	1	100%	1	1	100%	3	5	60%
Birds	Gruiformes	7	7	100%	27	44	61%	72	165	44%
Birds	Leptosomatidae	0	1	0%	0	1	0%	0	1	0%
Birds	Mesitomithiformes	0	1	0%	0	2	0%	0	3	0%
Birds	Musophagiformes	1	1	100%	5	6	83%	16	23	69%
Birds	Opisthicomiformes	1	1	100%	1	1	100%	1	1	100%
Birds	Otididae	1	1	100%	4	11	36%	6	26	23%
Birds	Paleognathae	6	6	100%	14	15	93%	37	59	62%
Birds	Passeriformes	91	101	90%	873	1226	71%	2951	5756	51%
Birds	Pelecaniformes	7	7	100%	8	9	88%	34	62	54%
Birds	Phaethontiformes	1	1	100%	1	1	100%	3	3	100%
Birds	Phoenicopteriformes	1	1	100%	3	3	100%	4	5	80%
Birds	Piciformes	5	5	100%	53	68	77%	255	398	64%
Birds	Podicipediformes	1	1	100%	6	6	100%	13	22	59%
Birds	Procellariiformes	4	4	100%	20	27	74%	56	112	50%
Birds	Psittaciformes	1	1	100%	69	85	81%	201	364	55%
Birds	Pteroclideiformes	1	1	100%	1	2	50%	2	16	13%
Birds	Sphenisciformes	1	1	100%	6	6	100%	13	17	76%
Birds	Strigiformes	2	2	100%	21	29	72%	66	195	33%
Birds	Trogoniformes	1	1	100%	5	6	83%	25	39	64%
<b>Birds Count</b>	<b>34</b>	<b>182</b>	<b>199</b>	<b>91%</b>	<b>1587</b>	<b>2172</b>	<b>73%</b>	<b>5074</b>	<b>9723</b>	<b>52%</b>

Groups	Orders	Families			Genera			Species		
		G10K samples available	Total	% of Total	G10K samples available	Total	% of Total	G10K samples available	Total	% of Total
Amphibians	Anura	35	40	87%	248	408	60%	1609	5811	27%
Amphibians	Caudata	10	10	100%	40	69	57%	133	583	22%
Amphibians	Gymnophiona	5	6	83%	13	33	39%	18	176	10%
<b>Amphibians Count</b>	<b>3</b>	<b>50</b>	<b>56</b>	<b>89%</b>	<b>301</b>	<b>510</b>	<b>59%</b>	<b>1760</b>	<b>6570</b>	<b>27%</b>
Reptiles	Crocodylia	3	3	100%	8	9	88%	20	23	86%
Reptiles	Rhynchocephalia	1	1	100%	1	1	100%	1	2	50%
Reptiles	Squamata	47	48	97%	672	983	68%	3136	8664	36%
Reptiles	Testudines	12	13	92%	70	94	74%	140	313	44%
<b>Nonavian Reptiles Count</b>	<b>4</b>	<b>63</b>	<b>65</b>	<b>97%</b>	<b>751</b>	<b>1087</b>	<b>69%</b>	<b>3297</b>	<b>9002</b>	<b>37%</b>
Fishes	Acipenseriformes	2	2	100%	4	6	66%	8	28	28%
Fishes	Albuliformes	1	1	100%	1	2	50%	2	11	18%
Fishes	Amiiformes	1	1	100%	1	1	100%	1	1	100%
Fishes	Anguilliformes	13	15	86%	40	153	26%	92	903	10%
Fishes	Ateleopodiformes	1	1	100%	2	4	50%	2	13	15%
Fishes	Atheriniformes	7	10	70%	19	51	37%	54	326	16%
Fishes	Aulopiformes	13	16	81%	27	44	61%	60	256	23%
Fishes	Batrachoidiformes	1	1	100%	6	23	26%	10	80	12%
Fishes	Beloniformes	5	5	100%	21	34	61%	55	244	22%
Fishes	Beryciformes	7	7	100%	18	30	60%	52	158	32%
Fishes	Carcharhiniformes	6	8	75%	32	50	64%	96	277	34%
Fishes	Ceratodontiformes	1	1	100%	1	1	100%	1	1	100%
Fishes	Cetomimiformes	1	2	50%	5	18	27%	7	33	21%
Fishes	Characiformes	14	18	77%	85	278	30%	184	1898	9%
Fishes	Chimaeriformes	3	3	100%	5	6	83%	17	46	36%
Fishes	Clupeiformes	4	6	66%	24	84	28%	48	393	12%
Fishes	Coelacanthiformes	1	1	100%	1	1	100%	1	2	50%
Fishes	Cypriniformes	4	6	66%	47	451	10%	96	3943	2%
Fishes	Cyprinodontiformes	9	10	90%	33	118	27%	101	1172	8%
Fishes	Elopiformes	2	2	100%	2	2	100%	5	8	62%
Fishes	Esociformes	2	2	100%	4	4	100%	6	12	50%
Fishes	Gadiformes	9	10	90%	49	81	60%	107	610	17%
Fishes	Gasterosteiformes	5	5	100%	10	11	90%	11	29	37%
Fishes	Gobiesociformes	1	1	100%	4	46	8%	6	154	3%
Fishes	Gonorynchiformes	2	4	50%	2	7	28%	2	37	5%
Fishes	Gymnotiformes	5	5	100%	8	32	25%	11	158	6%

Groups	Orders	Families			Genera			Species		
		G10K samples available	Total	% of Total	G10K samples available	Total	% of Total	G10K samples available	Total	% of Total
Fishes	Heterodontiformes	1	1	100%	1	1	100%	4	9	44%
Fishes	Hexanchiformes	2	2	100%	4	4	100%	5	5	100%
Fishes	Lamniformes	6	7	85%	8	10	80%	11	17	64%
Fishes	Lampriformes	6	7	85%	10	12	83%	12	27	44%
Fishes	Lepidosireniformes	2	2	100%	2	2	100%	2	5	40%
Fishes	Lepisosteiformes	1	1	100%	2	2	100%	4	7	57%
Fishes	Lophiiformes	12	18	66%	25	67	37%	41	328	12%
Fishes	Myctophiformes	2	2	100%	18	36	50%	37	262	14%
Fishes	Myxiniformes	1	1	100%	2	5	40%	4	71	5%
Fishes	Notacanthiformes	2	2	100%	5	6	83%	9	26	34%
Fishes	Ophidiiformes	3	5	60%	25	116	21%	43	501	8%
Fishes	Orectolobiformes	6	7	85%	9	13	69%	20	42	47%
Fishes	Osmeriformes	11	13	84%	37	77	48%	68	315	21%
Fishes	Osteoglossiformes	7	7	100%	10	32	31%	11	219	5%
Fishes	Perciformes	129	164	78%	684	1722	39%	1892	10663	17%
Fishes	Percopsiformes	3	3	100%	4	7	57%	5	9	55%
Fishes	Petromyzontiformes	1	1	100%	4	8	50%	5	43	11%
Fishes	Pleuronectiformes	11	11	100%	80	136	58%	152	793	19%
Fishes	Polymixiiformes	1	1	100%	1	1	100%	3	10	30%
Fishes	Polypteriformes	1	1	100%	1	2	50%	1	12	8%
Fishes	Pristiformes	1	1	100%	2	2	100%	4	7	57%
Fishes	Pristiophoriformes	1	1	100%	1	2	50%	2	6	33%
Fishes	Rajiformes	12	12	100%	43	70	61%	143	574	24%
Fishes	Saccopharyngiformes	4	4	100%	4	5	80%	5	28	17%
Fishes	Salmoniformes	1	1	100%	6	11	54%	12	210	5%
Fishes	Scorpaeniformes	25	35	71%	107	294	36%	257	1573	16%
Fishes	Siluriformes	18	38	47%	68	478	14%	125	3400	3%
Fishes	Squaliformes	4	4	100%	12	23	52%	40	126	31%
Fishes	Squatiformes	1	1	100%	1	1	100%	6	22	27%
Fishes	Stephanoberyciformes	2	4	50%	5	11	45%	6	55	10%
Fishes	Stomiiformes	4	4	100%	38	53	71%	69	423	16%
Fishes	Synbranchiformes	2	3	66%	5	13	38%	7	110	6%
Fishes	Syngnathiformes	4	5	80%	18	64	28%	39	336	11%
Fishes	Tetraodontiformes	10	10	100%	72	104	69%	143	436	32%
Fishes	Torpediniformes	2	4	50%	2	12	16%	8	68	11%
Fishes	Zeiformes	5	6	83%	10	16	62%	16	33	48%

Groups	Orders	Families			Genera			Species		
		G10K samples available	Total	% of Total	G10K samples available	Total	% of Total	G10K samples available	Total	% of Total
Fishes Count	62	424	532	80%	1777	4956	36%	4246	31564	13%

## **Appendix 3**

### **Data Analysis**

The G10K project intends to produce assembled whole chromosomes with high enough quality to support bioinformatics analysis, including whole-genome multiple alignment, determination of lines of descent for segments of DNA with sizes ranging from single bases to multi-mega-base chromosomal regions, as well as gene finding and the identification of other functional elements by patterns of selection. This will require the coordination of a network of computational centers to assemble, align, store, and disseminate via on-line genome browsers the vast amount of sequencing and annotation information. A central sample tracking database would also track samples and their quality and link taxonomic data with important phenotypic attributes (Appendix 4). A cloud computing infrastructure could support the availability of algorithms for efficient large-scale analysis of these data.

To put the ambition of G10K into perspective, we present the following test calculation for the sample throughput and processing requirements for the sequencing and assembly stages. A five-year project requires 2000 genomes to be sequenced and assembled every year. It seems reasonable to distribute the workload over 20 sites, resulting in 100 samples (genomes) to be completed by one sequencing site per year. This translates to a required output of two genomes per week for each sequencing site, including both sequencing and draft assembly.

Distributing and sequencing 10,000 sequence-ready samples across 20 sequencing sites at a rate of 100 per site per year for five years necessitates a highly structured workflow that minimizes delays in sequencing. We foresee a distributed network of sequencing facilities, each represented by a coordinator that will communicate with a large data center (Figure 1 of Appendix 3).



Assembly of vertebrate genomes using new sequencing technologies is still an active research topic. While we anticipate that assembly methods will improve in the next few years, with current technology a full assembly requires the use of a large memory machine (~1.5Tb RAM), 128 CPUs and 2 weeks to complete. It is likely that new sequencing technologies will take advantage of some derivative of cloud computing technology, or equivalent large scale distributed computing resources, for assembly and initial analysis of data from a sequencing run. This activity could be coordinated remotely by the data center. A primary driver here is the fact that the amount of storage for a vertebrate sequencing run can reach many Tb of disk. To transfer all raw data to a central location daily would exceed the capacity of most networks; therefore, initial steps in analysis of each sequencing run must be conducted in a distributed fashion.

We envision the output of the initial analysis of a sequencing run to take the form of a 'minimal' assembly of all reads into (possibly small) contigs. The central data center will then download this assembly in a format taking up a much smaller amount of storage space. For example, the files could be communicated in FASTA-formatted sequences with quality scores (fastq format), or some more expressive variant. This 'minimal' assembly will be archived, accessioned and made publically available. The current protocol for data release is that every contig over 1000 bases is made public. As the precise nature of the data release depends on the properties of the technology used, the spirit of timely public data release should be maintained.

The completeness of each species' assembly will depend on the sequencing technology employed. The most useful assembly for annotation and comparative genomics is a "complete" assembly, which includes deep enough coverage for whole-chromosome assemblies, i.e., sequence coverage that is sufficient to (ideally) create

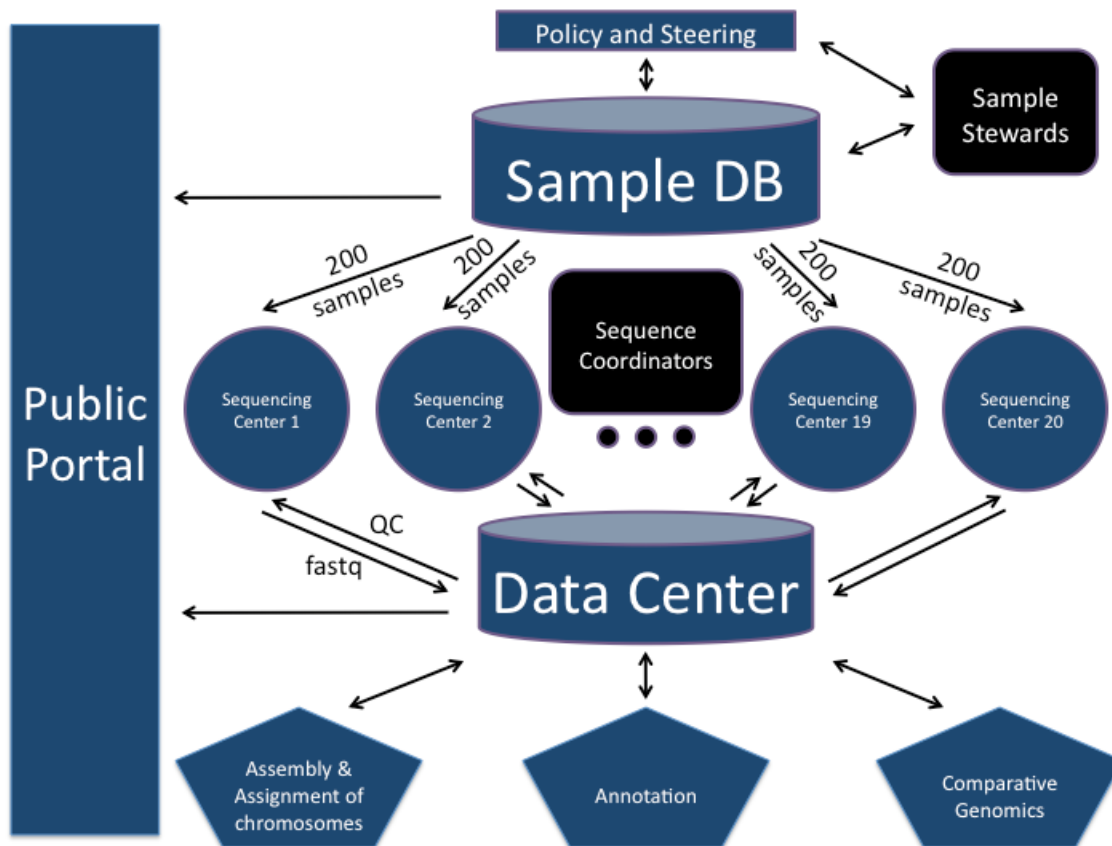
one sequence contig per chromosome, and identify DNA polymorphisms within and between species. This is a seminal goal of the project.

## **Analysis and Annotation**

Initial computational analysis, annotation and visualization of the Genome 10K data is a very open-ended aspect of the project, but can be anticipated to demand at least as much computational effort as sequence assembly. Most of this will be performed at the data center. It will critically rely on whole-genome multiple alignments, currently one of the most computation-intensive steps in vertebrate genome analysis, requiring large clusters of machines. It is unclear how difficult this problem will be, as whole-genome multiple alignments for more than 50 vertebrate genomes have never been built. We suggest a computational pipeline for analysis, annotation and visualization be specified soon, and put in place and tested prior to the start of sequencing. This pipeline would serve as an evaluation of assembly, analysis and annotation methods prior to their deployment.

## **Timing**

Due to the anticipated rapid improvement in technology, some decisions should be postponed as late as possible to allow for a better estimate of precise parameters (e.g. exact sample size and assembly method). However, it is imperative that the sample tracking database (Appendix 4) be operational as soon as possible in order to store the thousands of existing samples.



**Figure 1 of Appendix 3**

This is an overview of the sample and data flow through the multiple sequencing centers. A central tracking database (Sample DB) will store the progress of all species through the sequencing and analysis process and will be publicly available.

## **Appendix 4**

### **Genome 10K Sample Tracking Database Design**

#### **1. Introduction**

This document describes the design of the Genome 10K sample tracking database. The goal of this database is to catalog samples available to the Genome 10K program. It will provide sufficient information to allow the project to prioritize and select samples for sequencing. Tracking of samples during the sequencing and assembly processes will be supported.

#### **2. Taxonomies**

Tracking taxonomy for animals is the key to evaluating the state of the Genome 10K project and to prioritize obtaining and sequencing samples. This is greatly complicated by the fact that there is no one agreed upon taxonomy covering all vertebrates. Different clade-specific taxonomic trees are preferred by experts in each of the species groups. While the NCBI taxonomy (Benson et al., 2008; Wheeler et al., 2008) covers all vertebrates, it often disagrees with the preferred clade-specific trees. Mapping between the taxonomies is not straightforward and is often not possible due to conflicting classifications or the incompleteness of taxonomic branches.

The Genome 10K project will not attempt to dictate a particular taxonomy for the entire project. Instead groups may use clade-specific taxonomies for their samples and these will be recorded in the database. A G10K hybrid taxonomy will be constructed for reporting purposes by attaching the clade-specific trees to a vertebrate wide taxonomy tree “backbone”. Species will also be mapped to the leaves of alternate taxonomy trees when these are available. This will permit different analyses to use different trees.

Taxonomies are not static and undergo periodic updates. The NCBI taxonomy database is updated daily, while others undergo slower rates of change. The G10K database will store multiple versions of each taxonomy, supporting migration to new versions. We will not attempt to track every version of NCBI's taxonomy; it will be updated periodically or as needed by the project.

The following taxonomies will be imported into the Genome 10K database if possible:

- NCBI taxonomy database (Benson et al., 2008; Wheeler et al., 2008)  
(<http://www.ncbi.nlm.nih.gov/Taxonomy/>)
- Wilson and Reeder: Mammal Species of the World (Wilson and Reeder 2005)  
(<http://www.bucknell.edu/MSW3/>)
- Amphibian Species of the World: an Online Reference (Frost 2009)  
(<http://research.amnh.org/herpetology/amphibia/>)
- TIGR Reptile Database (Uetz et al., 2007) (<http://www.reptile-database.org/>)
- Catalog of Fishes (Eschmeyer 1998)  
(<http://research.calacademy.org/research/Ichthyology/Catalog/fishcatmain.asp>)
- Sibley Bird database (Monroe and Sibley 1993; Sibley and Ahlquist 1990; Sibley and Monroe 1993) (<http://www.scricciolo.com/classificazione/cover.html>)
- Tree of Life project (Maddison and Schulz 2007) (<http://tolweb.org>)

### **3. Database organization**

This section describes the logical organization of the G10K sample tracking database. It is a high-level description of the data that will be managed for the project. This attempts to be comprehensive while omitting the details of the SQL schema that implements the database.

### **3.1 Species record**

The species records provide the link between the externally defined taxonomy trees and the species names in the database, along with any species-specific data tracked by the project.

The following fields are defined:

- G10K species ID
- class
- order
- family
- genus
- species
- subspecies
- common name
- G10K taxonomy tree - taxonomy tree and version used to define the taxonomy for this species in the G10K project
- species ID in G10K taxonomy tree
- list of other taxonomies and corresponding species IDs for this species in those other taxonomies
- USESA (U.S. Endangered Species Act status)
- CITES (Goldsmith 1978) appendix status

### **3.2 Animal record**

An animal record is kept for each individual animal tracked by the database. Biological samples from an animal are tracked in separate records, allowing multiple samples to be tracked for a given individual animal.

The following fields are defined:

- G10K animal ID - ID number assigned by G10K project
- species ID - reference to the G10K species record
- barcode - BOLD (Ratnasingham and Hebert 2007) barcode accession
- sex
- steward – steward for the animal or animal tissues
- steward's animal ID - identifier for animal used by the steward
- date of birth
- location of birth - longitude, latitude, depth
- date of death
- date of collection
- location of collection - longitude, latitude, depth
- collector
- voucher ID number
- voucher location
- steward holds permit for collection (yes, no, unknown, NA)
- ACUC permit status - steward has permit of their institution's Animal Care and Use Committee (confirmed, pending, unknown, NA)
- USESA permit status (confirmed, pending, unknown, NA)
- CITES permit status (confirmed, pending, unknown, NA)

### **3.3 Sample record**

A sample record describes a biological sample from an animal. Multiple samples may be created from the same or different tissues from the same animal, including cell lines.

The following fields are defined:

- G10K sample ID - ID number assigned by G10K project
- animal ID - reference to G10K animal record
- steward – steward for the sample
- steward's sample ID – sample identifier for specimen used by the steward
- tissue from which sample is taken
- sample type (DNA, blood, other tissue, cell line)
- sample quantity
- sample quality
- storage location
- preservative type
- fields to track progress of sample in G10K pipeline (to be specified)

### **3.4 Steward record**

The steward records track the individuals and institutions responsible for each animal and sample in the database. Animals and samples may have different stewards.

The following fields are defined:

- G10K steward ID
- name
- institution
- address
- email
- telephone
- fax



## References:

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL, 2008. GenBank. *Nucleic Acids Res.* 36: D25-D30.

Eschmeyer WN, 1998. *Catalog of fishes*. California Academy of Sciences: San Francisco, 3 v. (2905 p.)pp.

Field D, 2008. Working together to put molecules on the map. *Nature.* 453: 978.

Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV et al., 2008. The minimum information about a genome sequence (MIGS) specification. *Nat. Biotechnol.* 26: 541-547.

Frost DR. 2009. Version 5.3 (12 February, 2009). *Amphibian species of the world: an online reference*. Available from: <http://research.amnh.org/herpetology/amphibia/>

genome.gov. 2003. Reaffirmation and extension of NHGRI rapid data release policies: large-scale sequencing and other community resource projects. [cited 2009 June 28]. Available from: <http://www.genome.gov/10506537>.

Goldsmith EI, 1978. *The Convention on International Trade in Endangered Species of Wild Fauna and Flora*. *J Med Primatol.* 7: 122-124.

Hanner RH, Gregory TR, 2007. Genomic diversity research and the role of biorepositories. *Cell Preserv. Technol.* 5: 93-103.

Maddison DR, Schulz K-S. 2007. *The Tree of Life Web Project*. Available from: <http://tolweb.org>.

Monroe BL, Sibley CG, 1993. *A world checklist of birds*. Yale University Press: New Haven, xix, 393 p.pp.

Ratnasingham S, Hebert PD, 2007. bold: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Mol Ecol Notes.* 7: 355-364.

Sibley CG, Ahlquist JE, 1990. *Phylogeny and classification of birds : a study in molecular evolution*. Yale University Press: New Haven, xxiii, 976 p.pp.

Sibley CG, Monroe BL, 1993. A supplement to distribution and taxonomy of birds of the world. Yale University Press: New Haven, vi, 108 p.pp.

Uetz P, Goll J, Hallermann J, 2007. Die TIGR-Reptiliendatenbank. Elaphe. 15: 16-19.

Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Edgar R, Federhen S et al., 2008. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 36: D13-D21.

Wilson DE, Reeder DM, 2005. Mammal species of the world : a taxonomic and geographic reference. 3rd edn. Johns Hopkins University Press: Baltimore, 2 v. (xxxv, 2142 p.)pp.

