

costseq

Paul Muir

January 21, 2016

Contents

1	Full economic cost [0/4]	2
1.1	[] Cost sources	2
1.1.1	Data generation	2
1.1.2	Data analysis	2
1.2	[] Cost information from YCGA (CMG Grant)	2
1.2.1	Dell PowerEdge R720 with 512GB of memory and 3.6TB of disk space for programming and data pro- cessing.	2
1.2.2	New technology development costs	2
1.2.3	In vitro functional studies	3
1.2.4	Acquisition of sample collections from other centers and NIH institutes.	3
1.2.5	Exome capture and sequencing	3
1.2.6	Whole genome sequencing	4
1.2.7	Notes:	5
1.2.8	Variant confirmation via Sanger Sequencing	5
1.3	[] Information from the NHGI	6
1.3.1	Email sent to Kris Wetterstrand.	6
1.4	[] Information from Illumina	6
1.4.1	Email sent to Crane Harris.	6

lo#-* mode: org -*

1 Full economic cost [0/4]

1.1 [] Cost sources

1.1.1 Data generation

1. labor
2. management
3. utilities
4. sequencing machines and other equipment
5. low level sequence informatics (e.g. machine level base calls)
6. data submission and storage
7. indirect costs

1.1.2 Data analysis

1. quality assesment
2. technology development
3. downstream bioinformatics analysis
4. data storage and analysis

1.2 [] Cost information from YCGA (CMG Grant)

1.2.1 Dell PowerEdge R720 with 512GB of memory and 3.6TB of disk space for programming and data processing.

1. \$10,000 (one time purchase in Year 1)

1.2.2 New technology development costs

1. \$44,800
 - (a) \$11,200 per year for:
 - i. PCR plates, pipette tips, 15 ml and 50 ml Falcon tubes, 1.5ml and 2ml Eppendorf tubes, 5-20ml serological pipettes, Pasteur pipettes, patient sample collection tubes, gloves, reagent reservoirs and other disposables.

2. \$115,688
 - (a) \$28,922 per year for:
 - i. agarose, tris, phenol, chloroform, molecular biology grade ethanol, proteinase K, PCR buffers, Taq polymerase various chemicals and R&D sequencing lanes.
3. Our total R&D budget constitutes ~5% of the total cost of the grant, which we believe is conservative.

1.2.3 In vitro functional studies

1. \$400,000
 - (a) \$100,000/year for years 1-4

1.2.4 Acquisition of sample collections from other centers and NIH institutes.

1. \$20,000 (over all 4 years)

1.2.5 Exome capture and sequencing

1. \$7,656,000 in direct costs is requested over the four year period to capture and sequence over 26,000 exomes.
 - (a) Year one: At the rate of \$330 in direct costs/exome, we request \$1,914,000 in direct costs towards capture and sequence analysis of 5,800 exomes.
 - (b) Year two: At the rate of \$305 in direct costs/exome, we request \$1,914,000 in direct costs towards capture and sequencing of 6,275 exomes.
 - (c) Year three: At the rate of \$291 in direct costs/exome, we request \$1,914,000 in direct costs to capture and sequence 6,577 exomes
 - (d) Year four: At a rate of \$254 in direct costs/exome, we request \$1,914,000 in direct costs to capture and sequence 7,535 exomes.

Table 1: Exome Sequencing Costs for the duration of the funding period*

Description	Year 1	Year 2	Year 3	Year 4
Capture and Library Preparation cost				
NimbleGen V2 Capture Oligo pools	\$36	\$34	\$32	\$32
Labor	\$57	\$54	\$51	\$51
Reagent and Supplies	\$50	\$48	\$45	\$43
Instrument Depreciation and Maintenance	\$12	\$12	\$12	\$0
Administration and IT etc. (3%)	\$5	\$4	\$4	\$4
Sub Total Library Prep	\$160	\$152	\$145	\$131
Sequencing Cost				
Labor	\$6	\$5	\$5	\$5
Sequencing Reagents & Supplies	\$118	\$105	\$100	\$100
Instrument Depreciation	\$24	\$22	\$22	\$0
HPC and IT	\$8	\$6	\$5	\$3
Maintenance contract	\$10	\$10	\$10	\$10
Administrative cost (3%)	\$5	\$4	\$4	\$4
Sub Total Sequencing	\$171	\$153	\$146	\$123
Total Direct cost (Library prep + Sequencing)	\$330	\$305	\$291	\$254
Indirect costs (Equipment not included)	\$196	\$180	\$171	\$169
Direct + Indirect Cost	\$526	\$485	\$463	\$423

*Costs are determined for 2x100 bp sequencing using Illumina HiSeq 4000

1.2.6 Whole genome sequencing

1. \$1,404,000 in direct costs is requested over the four year period to library preparation and sequencing of 672 genomes.
 - (a) Year one: At the rate of \$2,340 in direct costs/genome, we request \$351,000 in direct costs towards library preparation and sequence analysis of 150 genomes.
 - (b) Year two: At the rate of \$2,226 in direct costs/genome, we request \$351,000 in direct costs to capture and sequence 157 genomes.
 - (c) Year three: At the rate of \$2,121 in direct costs/genome, we request \$351,000 in direct costs to capture and sequence 165 genomes.
 - (d) Year four: At a rate of \$1,754 in direct costs/genome, we ask \$351,000 in direct costs to capture and sequence 200 genomes.

Table 2 : Whole-Genome Sequencing cost for the entire funding period

Description	Year 1	Year 2	Year 3	Year 4
Library Preparation Costs				
Labor	\$70	\$56	\$45	\$36
Reagent and Supplies	\$75	\$71	\$68	\$64
Instrument Depreciation	\$10	\$10	\$10	\$0
Service Contract IT	\$8	\$8	\$8	\$8
Administration (3%)	\$5	\$4	\$4	\$3
Sub Total Library Prep	\$168	\$149	\$134	\$111
Sequencing Costs (1 genome/lane 2x150 bp)				
Labor	\$82	\$78	\$74	\$70
Sequencing Reagents and Supplies	\$1767	\$1679	\$1595	\$1515
Instrument Depreciation	\$175	\$175	\$175	\$0
HPC IT	\$75	\$75	\$75	\$0
Maintenance Contract	\$10	\$10	\$10	\$10
Administrative cost (3%)	\$63	\$60	\$58	\$48
Sub Total Sequencing	\$2172	\$2077	\$1987	\$1643
Total Direct Cost (Library prep + Sequencing)	\$2340	\$2226	\$2121	\$1754
Indirect Cost (Equipment cost is not included)	\$1433	\$1358	\$1287	\$1167
Direct + Indirect costs	\$3773	\$3584	\$3408	\$2921

*Costs are determined for 2x150 bp sequencing using Illumina HiSeq 4000

1.2.7 Notes:

1. Labor costs for WGS projected to drop while exome sequencing labor costs level off.
2. Decrease in WGS sequencing prep is primarily from labor costs and to a lesser extent reagent costs.

1.2.8 Variant confirmation via Sanger Sequencing

1. \$60,000
 - (a) \$3.25 for the primer pair and \$1.75 for the actual Sanger sequencing. In total, we will have 12,000 variants from 1,500 exome sequencing variants.

1.3 [] Information from the NHGI

1.3.1 Email sent to Kris Wetterstrand.

1. She said she is happy to share whatever information she has but is busy and will get back to us in a few days.

1.4 [] Information from Illumina

1.4.1 Email sent to Crane Harris.

1. He shared this link: <http://www.molecularecologist.com/next-gen-fieldguide-2014/>
2. "Illumina is running three regional sales meetings around the world over the next two and a half weeks, and the folks who might be able to help here are consumed with prep and travel for those. If as you get into your analysis there is a call for more system-specific information – high throughput/low throughput, clinical/research, depth of coverage, human/other, targeted/whole genome – let me know and we can look for something more targeted."