

To: Christina Fan[chfan@stanford.edu]
Cc: Stephen Quake[quake@stanford.edu]
From: Yair Blumenfeld
Sent: Wed 5/28/2008 8:13:32 PM
Subject: Re: sequencing analysis

Awesome!

Quoting Christina Fan <chfan@stanford.edu>:

> Hi Steve, Yair,
>
> I am attaching a graph of the preliminary analysis of the full
> sequencing run.
> The run includes 4 T21 and 3 normal samples.
> The total number of reads per sample was about 9 million, of which
> about half (4.4 million) on average are useful for analysis.
> Pay attention to the distribution of chr21.
>
> Christina

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE plist PUBLIC "-//Apple//DTD PLIST 1.0//EN"
"http://www.apple.com/DTDs/PropertyList-1.0.dtd">
<plist version="1.0">
<dict>
<key>date-sent</key>
<real>1212005612</real>
<key>flags</key>
<integer>8590195969</integer>
<key>original-mailbox</key>
<string>imap://h.christina.fan@imap.gmail.com/INBOX</string>
<key>remote-id</key>
<string>3992</string>
<key>subject</key>
<string>Re: sequencing analysis</string>
</dict>
</plist>
```

STANFORD EXHIBIT 2127
SEQUENOM v. STANFORD
CASE IPR2013-00390

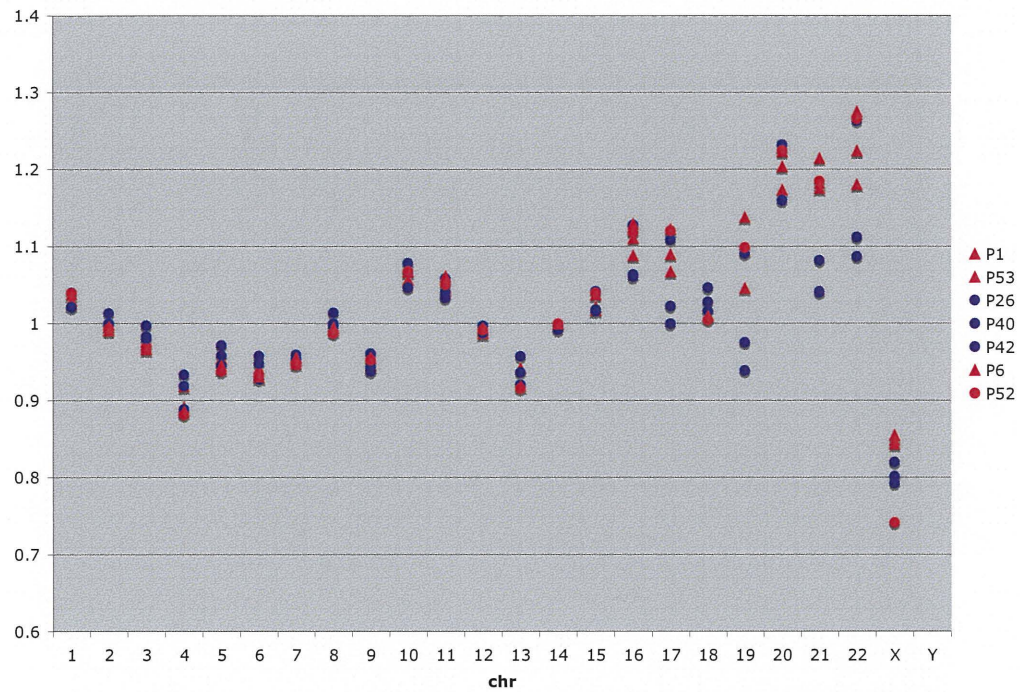
P1, P6, P53: 47 XX +21 (red triangle)

P52: 47 XY +21 (red circle)

P26, P40, P42: 46 XY (blue circle)

Coverage: # 25bp reads mapped to a chromosome / size of the chromosome

Average number of total useful reads per sample: 4.4 million



```

% look at the distribution of reads across each chromosome
% ignore golden path

chrname={'chr1','chr2','chr3','chr4','chr5','chr6','chr7','chr8',
',','chr9','chr10','chr11','chr12','chr13','chr14','...',
',','chr15','chr16','chr17','chr18','chr19','chr20','chr21','c',
hr22','chrX','chrY'};

chrlength=[247249719
242951149
199501827
191273063
180857866
170899992
158821424
146274826
140273252
135374737
134452384
132349534
114142980
106368585
100338915
88827254
78774742
76117153
63811651
62435964
46944323
49691432
154913754
57772954];

load gapregions

%noreadregioncount=cell(24,1);
%noreadregion=cell(24,1);

readcount=cell(24,1);
%X=cell(24,1);

for i=1:length(chrname)
    filename=sprintf('%s_coord_U0U1',chrname{i}(1:(end-1)));
    load (filename);
    bin=[0:50000:chrlength(i)];
    bar(bin, histc(chr_coord, bin));
    title(chrname{i}(1:(end-1)));
    xlabel('chr coordinate');
    ylabel('frequency of reads');
    figurefilename=sprintf('%s_distrib.bmp',chrname{i}(1:(end-1)));

```

```

saveas(gcf, figurefilename)
readcount=[];
bin=[0:20000:gapbegin{i}(1)];
readcount=[readcount, histc(chr_coord, bin)];
for j=1:(length(gapbegin{i})-1)
    bin=[gapend{i}(j):20000:gapbegin{i}(j+1)];
    readcount=[readcount; histc(chr_coord, bin)];
end
n{i}=readcount;
bar([0:1:100], histc(readcount, [0:1:100]));
title(chrname{i}(1:(end-1)));
xlabel('chr coordinate');
ylabel('# reads per 20kb');
figurefilename=sprintf('%s_hist.bmp', chrname{i}(1:(end-1)));
saveas(gcf, figurefilename)
%X{i}=[1:50000:chrlength(i)];
%readcount{i}=histc(read_chr, X{i});
%distrib=histc(read_chr, X);
%noread=find(distrib==0);
%noreadcoord=X(noread);
%Y=[1:50000:chrlength(i)];
%[noreadregioncount{i} noreadregion{i}]=hist(noreadcoord, Y);
%figure, hist(noreadcoord, Y);
%title(chrname{i}(1:(end-1)));
%xlabel('chr coordinate');
%ylabel('frequency of reads');
end

for i=1:24
    nmed(i)=median(n{i});
end

save readdistrib n

% for i=1:length(chrname)
%     tolerance=3*sqrt(median(readcount{i}));
%     [b c]=sort(readcount{i});
%     a=bin{i}(c);
%     outlier_region_U{i}=a(find(b>median(b)+tolerance));
%     outlier_region_L{i}=a(find(b<median(b)-tolerance));
%     outlier_count_L{i}=b(find(b<median(b)-tolerance));
%     outlier_count_U{i}=b(find(b>median(b)+tolerance));
%     mostoutlierU_count{i}=flipud(outlier_count_U{i}((end-9):end));
%     mostoutlierU_region{i}=flipud(outlier_region_U{i}
((end-9):end)');
% end
%
```

```

%% save outlierU outlier_region_U outlier_count_U mostoutlierU_count
mostoutlierU_region
%% save outlierL outlier_region_L outlier_count_L
%
%
%% write coordinates of underrepresented regions to file
% fid=fopen('underrepresentedregion.txt','wt');
%
%
% for i=1:length(chrname)-1
%
%     outlier_region_up=outlier_region_L{i}+50000;
%
%     for j=1:length(outlier_region_up)
%         if outlier_region_L{i}(j)<chrlength(i)
%
%             if outlier_region_up(j)>=chrlength(i)
%                 outlier_region_up(j)=chrlength(i)-1;
%             end
%
%             fprintf(fid, '%s\t',chrname{i}(1:(end-1)));
%             fprintf(fid, '%d\t', outlier_region_L{i}(j));
%             fprintf(fid, '%d\t', outlier_region_up(j));
%             fprintf(fid, '%d\n', outlier_count_L{i}(j));
%         end
%     end
% end
%
% fclose(fid);
%
%
%
%
% write coordinates of underrepresented regions (==0) to file
% fid=fopen('underrepresentedregion_0.txt','wt');
%
%
% for i=1:length(chrname)-1
%     outlier_0=find(outlier_count_L{i}==0);
%     outlier_region_0=outlier_region_L{i}(outlier_0);
%     outlier_region_0_up=outlier_region_0+50000;
%
%     outlier_count_0=outlier_count_L{i}(outlier_0);
%     for j=1:length(outlier_region_0)
%         if outlier_region_0(j)<chrlength(i)
%
%             if outlier_region_0_up(j)>=chrlength(i)

```

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.