

1 Hello from Illumina

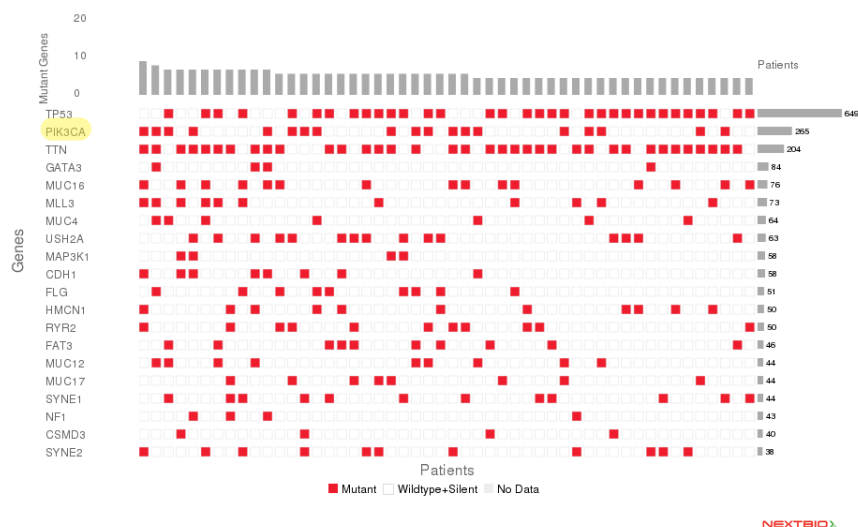
Thanks for paying Illumina a visit!

First, check out:

- the Illumina BaseSpace application platform @ basespace.com
- the Illumina NextBio platform @ nextbio.com

Second... **Ready for a challenge?**

2 CHALLENGE: Visualize patient data



2.1 Goal: make sense of clinical data

Build a web application which:

- Accepts patient data as flat text files (see next section for public example of TCGA data)
- Generates **visualizations and/or reports** of cohort (group) and individual patient data

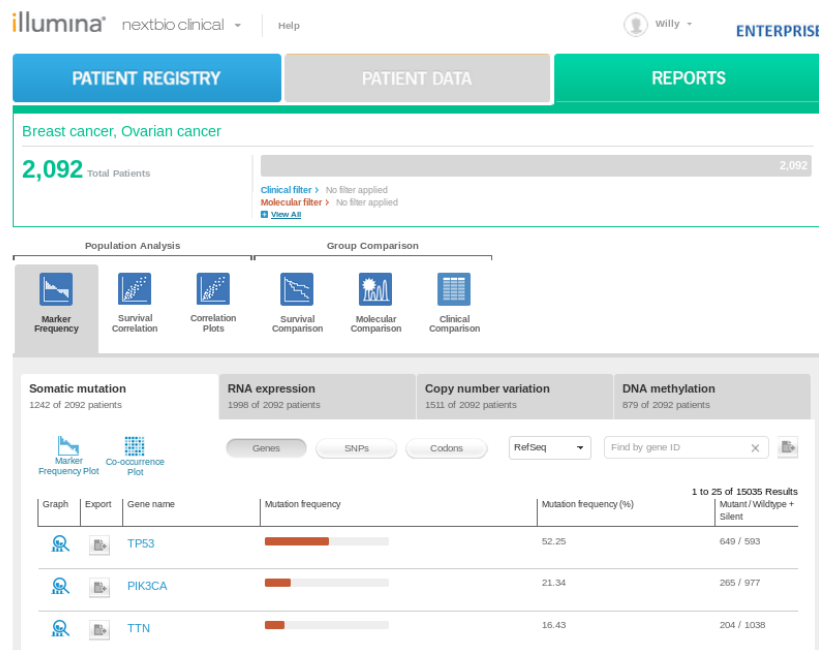
2.2 Data: Public patient data from Cbioportal

- Cbioportal.org is a great source of Clinical and Molecular data

- Many APIs can be found at http://www.cbioportal.org/web_api.jsp
- Specific example: you can find Clinical Attribute data for cancer patients at http://www.cbioportal.org/webservice.do?cmd=getClinicalData&case_set_id=ov_tcga_all

If you can accept such a file, or the URL to it, and show you any visualization you feel like building, e.g. a scatterplot, or an interactive view of a patient "graph" where links are similarity across attributes, PCA, etc. We encourage you to go wild – surprise us with the reporting/mashup of data.

3 Context: NextBio Clinical



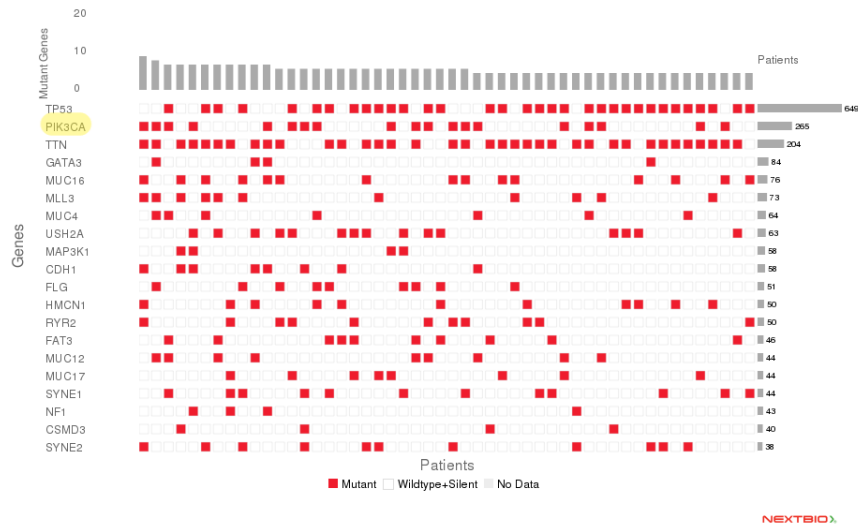
We're trying to make sense of this data too!

Here's a STORY told with data visualization (using NextBio Clinical, www.nextbio.com visualizations)

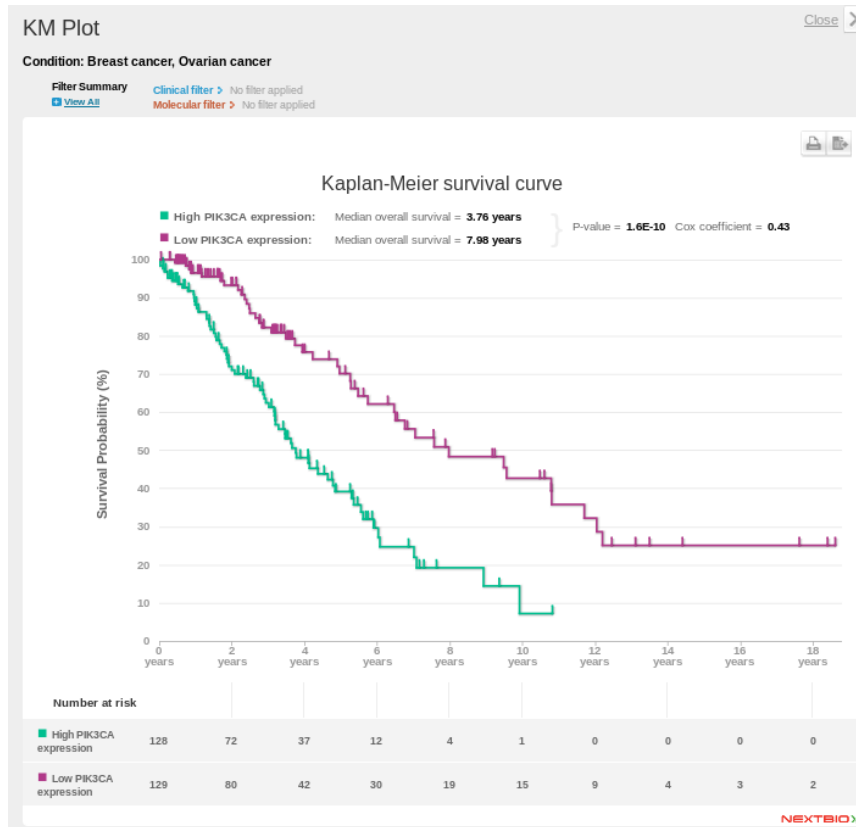
3.1 PIK3CA in Ovarian and Breast cancer

- Breast cancer and ovarian cancer afflict a huge number of women

- NextBio Clinical has a huge repository of data including public ovarian/breast cancer data
- You can easily filter public data by condition (Show Basic Patient finder) to focus on these two
- And generate reports and visualizations to answer questions, e.g.:
- What are the most mutated genes in our Ovarian/Breast cancer cohort? And how many genes are mutated per patient?

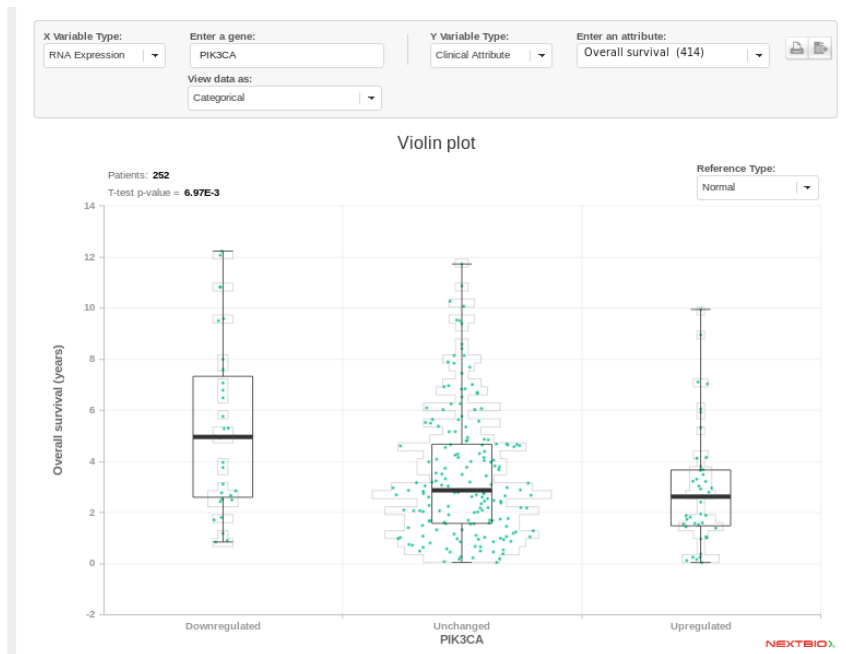


- PIK3CA looks interesting. We can ask more questions: Does the level of PIK3CA expression* predict survival? Why look at expression rather than mutation? We can't change our genome easily, but may be able to modulate expression

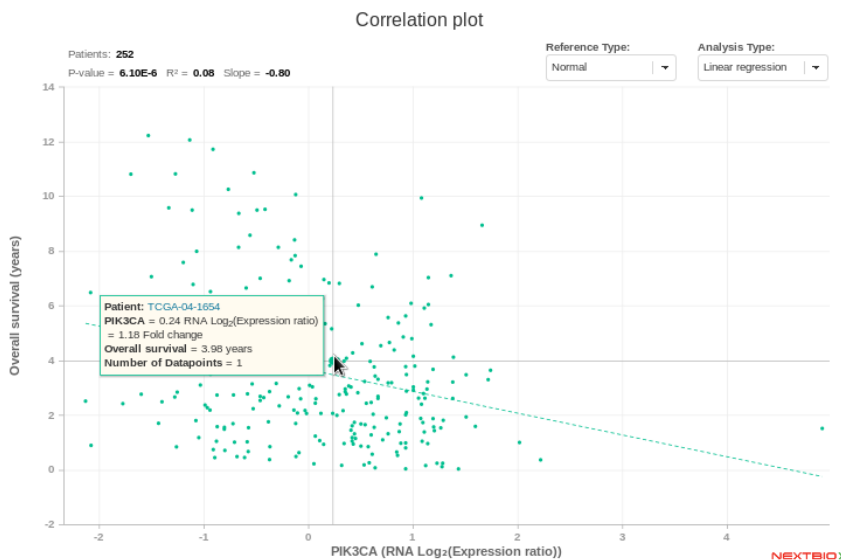


Clearly, **YES** - patients with high expression live about HALF as long.

- We can dive down into the correlation between expression and survival (or any other clinical or molecular attribute of the cohort) at the patient level with a "correlation scatterplot":
- Looking at the survival of patients grouped by whether PIK3CA is downregulated, unchanged, or upregulated reinforces the association found earlier:



- But clearly our categories have overlapping survival distributions. A plain scatterplot may be better, and perhaps the correlation will lead to a significant linear association:



- Yes

- And we can ask what this gene is:

PIK3CA Summary [View BodyAtlas for PIK3CA](#) Close X

Gene Symbol: **PIK3CA** Homo sapiens

Full name: phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha

Synonyms: PI3-kinase p110 subunit alpha, PI3K, PI3K-alpha, phosphatidylinositol 3-kinase, catalytic, 110-KD, alpha, phosphatidylinositol 3-kinase, catalytic, alpha polypeptide, phosphatidylinositol 4,5-bisphosphate 3-kinase 110 kDa catalytic subunit alpha, phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform, phosphatidylinositol-4,5-bisphosphate 3-kinase 110 kDa catalytic subunit alpha, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit, alpha isoform, phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha, phosphoinositide-3-kinase, catalytic, alpha polypeptide, p110-alpha, MGC142161, MGC142163, CLOVE, PtdIns-3-kinase p110, ptdIns-3-kinase subunit p110-alpha, serine/threonine protein kinase PIK3CA

Genomic Location: Chr 3: 180349005-180435194

Description: Phosphatidylinositol 3-kinase is composed of an 85 kDa regulatory subunit and a 110 kDa catalytic subunit. The protein encoded by this gene represents the catalytic subunit, which uses ATP to phosphorylate PtdIns, PtdIns4P and PtdIns(4,5)P2. This gene has been found to be oncogenic and has been implicated in cervical cancers. [provided by RefSeq]

Cytogenetic Map: 3q26.3

But we need your (their) help to do more, to produce more informative visualizations, to compute more powerful statistics, and to load information from more data providers and formats.

4 Data / APIs

http://www.cbioportal.org/web_api.jsp

Example flat data files:

Clinical data for Ovarian cancer: http://www.cbioportal.org/webservice.do?cmd=getClinicalData&case_set_id=ov_tcga_all

Mutation data for BRCA1 in TCGA Ovarian+Breast: http://www.cbioportal.org/webservice.do?cmd=getMutationData&genetic_profile_id=ov_tcga_mutations+ucec_tcga_mutations&gene_list=BRCA1

Thanks, Willy