Ilanit is the largest conference of all biological societies in Israel. It takes place every three years in Eilat, and is a big event: Four days, >2500 attendees, 8-10 parallel sessions, hundreds of posters, and several luminaries in the international communities as keynote speakers. The lectures span three different hotels. The 2020 conference took place on 17-20 February and several group members attended and presented there. Here are highlights according to each

David
I attended many microbiology and microbiome talks. They were almost all biological talks, with very few presenting primarily computational work. I will describe in brief the highlights of many of these talks to give a sense of current biological work in the field.

Microbial Ecology session:

*Mapping microbiome diversity in the age of big(gish) data* - Itai Sharon (Migal and Tel-Hai)
This was one of the few computational talks. The project looked at re-estimating the number of species in the overall human microbiome. As with the original estimate for the number of proteins encoded in the human genome, the number of bacterial species in the human microbiome has been corrected downward a number of times. The estimate of this project was even lower than the previous estimates, at around 2.5k. The bulk of the work was defining what is meant by a species by looking at Average Nucleotide Identities (ANI) of known reference species and marker genes in those references. The ANI has a bimodal distribution, with one mode in the mid 90% and the other distributed more broadly below 85%, with almost no pairs of genomes having ANI between the two modes.

*A globally abundant marine primary producer relies on microbial interactions to survive long-term stress* - Daniel Sher (U of Haifa)
This work studied the effects of nutrient stress and low-light conditions on *Prochlorococcus*, a very abundant phytoplankton and important primary producer in the ocean. They tested both isolates in lab conditions and in more “natural-like” environments and looked at the *Prochlorococcus* survival and viability both on its own and together with heterotrophic bacteria. They found that it survives much better when in co-culture and demonstrated that it undertakes surprisingly less photosynthesis at lower
depths than previously thought. A preprint describing most of the work in the talk can be found at: https://www.biorxiv.org/content/10.1101/657627v2.full.pdf

**Marine Microbial Interactions Across Scales - Einat Segev (Weizmann)**

Einat studies *Emiliania huxleyi*, the most common alga that is responsible for large algal blooms across the oceans. *E. huxleyi* creates calcite plates (coccoliths) that are passed through the membrane and arrayed around the outside of the cell. Einat discovered that these algae live in interaction with bacteria, where the bacteria fix themselves into the cell membrane of the algae. In this, initially symbiotic, but eventually parasitic, interaction, *E. Huxleyi* still creates the coccoliths but ejects them from the cell. The bacteria eventually act as pathogens, killing the algae, and their presence changes the algal cells’ metabolism. Importantly, the relative amounts of impurities found in the coccoliths - a critical marker in the geological climate record - can change, having profound implications for calibration of the climate record.

**Microbial simplicity for studying life’s complexity session:**

**Plasmids - from single cell dynamics to overall marine population - Shay Tal (National Center for Mariculture)**

This was very preliminary work in looking at the plasmids in seawater microbiome samples. The authors assembled the microbiomes and used Recycler, metaplasmidSPAdes, and HyASP to find plasmids. They had some initial characterizations of the genes, and noted that only in few cases were they able to identify known genes, even those that are associated with basic plasmid function such as replication. They were able to find some of the plasmids repeatedly across samples.

**Microbiology Adaptation & Evolution session:**

**Can microbiome functional profile elucidate hidden stress in its host? - Avihai Zolti (Hebrew U, Volcani)**

This was an interesting study in which the authors looked at plants growing in indeterminate stress conditions and were able to identify the causes of stress by looking at the metagenomes and metatranscriptomes of the root microbiome. Specifically, they compared plants watered using fresh water vs. treated wastewater (most of the irrigation in Israel uses treated recycled wastewater) and observed that the recycled water-irrigated plants were under stress (significantly lower yields). Differential expression analysis of the plants alone revealed elevated expression of general stress-response genes but no clues as to the specific sources or mechanisms of stress. In contrast, differential microbiome analysis revealed higher presence and elevated expression of genes specifically related to
high salinity, high pH, and lower oxygen levels. Overall, they presented this study as a proof-of-concept for using the microbiome to identify specific sources of stress or other phenotypic changes when differential expression is not revealing enough.

*Conversion of Escherichia coli to generate all biomass carbon from CO2* - Shmuel Gleizer (Weizmann)

This was another very interesting work that succeeded in using genetic engineering and selective pressure to create autotrophic *E. coli* - it uses CO2 as its carbon source, although it still requires formate to supply energy for the reaction. The formate can be produced through an electrochemical reaction (i.e., it can be powered by a renewable energy source to make this a carbon-neutral process). This is a significant first step in carbon-neutral or even carbon-negative bioproduction.

After genetic engineering to modify the metabolic pathways to allow for autotrophy and disrupt the heterotrophic pathways, the *E. coli* were subjected to a gradual reduction in carbon sources over hundreds of generations. This allowed the autotrophic genotype to be fixed in the population (without this selective pressure, the autotrophic genes are not stably maintained). The work has been published in *Cell*: [https://www.sciencedirect.com/science/article/pii/S0092867419312309](https://www.sciencedirect.com/science/article/pii/S0092867419312309)

*A seep from the past: Isolation and characterization of live yeast cells from ancient beer vessels and analysis of the beverages they produce* - Ronen Hazan, HUJI (? - They switched the presenting author and I’m not sure who it was that spoke)

In this work yeast was isolated from ancient beer containers from a few sites throughout Israel. Significant archaeological and microbiological work was done to isolate and cultivate the yeast, and confirm that it was from the original beer and not a contaminant. Six strains of yeast were cultivated and used to brew a standard recipe of beer. The beers were analyzed in the lab and sampled by professional beer tasters, confirming that the strains have been selected for beverage making. The analysis of the yeast strains showed them to be close to strains used in beer making today, especially in more traditional brewing methods in various parts of Africa. The next steps planned in this work include brewing beers according to the ancient recipes corresponding to the sources of the yeast and commercializing the product.

In the latter part of the talk the speaker discussed yeast senescence and showed that yeast are able to survive for extremely long periods of time in a “survival mode” but that they still bud and reproduce, albeit infrequently, when in this mode.

The work was published in mBio: [https://mbio.asm.org/content/10/2/e00388-19.long](https://mbio.asm.org/content/10/2/e00388-19.long)
Nimrod

Tal Shay (Ben Gurion University) - Dissecting the sex and age effects in the immune system transcriptome

Tal Shay's group studies the immune system using computational methods. In her talk she discussed the effect of sex and age on gene expression in the immune system of mice. At the phenotype level, males have more frequent and severe infectious diseases, while females have more auto-immune diseases. These differences are evident even before sexual maturity. Shay's group measured RNA-seq (bulk) profiles from 11 cell types of the immune system (e.g. macrophages, CD8 T cells) as part of the Immunological Genome Project (ImmGen) consortium. A total of 92 females and 91 males were used in the whole study. Samples of the same cell type showed high correlation between the sexes. Only 14 genes showed differential expression between the sexes when all cell types were considered simultaneously, with 3 higher in females (including Xist, the regulator of X chromosome inactivation) and 11 in males. Additionally, all genes were sorted by differential expression, and gene set enrichment analysis (GSEA) found that interferon response genes were highly expressed in females. Next they looked at every cell type separately, and aside from one gene, all differentially expressed genes were in macrophages. 26 were upregulated in females, and 15 in males. 7 of the 26 were previously shown to be upregulated by interferon stimulation, while none of the 15 were. Genes known to be upregulated in response to interferon were significantly upregulated in females. Next Tal's group exposed cells to interferon alpha and performed RNA-seq, for both males and females. Known genes related to interferon response were more highly expressed in females compared to males both before and after stimulation, and the upregulation (ratio between the expression after and before interferon) was higher for females. They also looked at ATAC-seq data from macrophages, and found that differentially expressed genes between males and females had differential accessibility. Finally, to assess the relevance of these results to humans, the mouse data were compared to human data (from the ImmVar dataset), and genes highly expressed in human females compared to males were also more highly expressed in mouse females. This trend was stronger for genes with higher differential expression.

This research was published in Nature Communications:

https://www.nature.com/articles/s41467-019-12348-6

Tal briefly mentioned a similar project they have on aging, where they analyze changes in gene expression in mouse aging. So far, no such changes were found, but this is work in progress.
Keren Yizhak (Technion, formerly Gadi Getz's lab at the Broad) - RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues.

Somatic mutations cause cancer, and the role of mutations in cancer is heavily studied. Recent research showed that specific mutations are also overrepresented in non-cancerous tissue. This was shown for blood cells as well as for human skin and esophagus. To investigate this phenomenon more comprehensively, Keren leveraged the GTEx dataset, which contains RNA-seq profiles from thousands of people in dozens of tissues. She developed an algorithm, RNA-MuTect, that identifies somatic mutations from RNA-seq data. Briefly, the algorithm gets the genotype of a person (measured from blood or an adjacent tissue) and RNA-seq data, and identifies prevalent somatic mutations in the RNA data compared to that person's reference (while accounting for different confounders, e.g. RNA editing). She tested the method on TCGA data, where ground truth mutation calling was available from tumor DNA, and detected 82% of the mutations that had sufficient RNA expression. She next applied RNA-MuTect to ~6700 samples and ~500 samples from GTEx. Overall ~9000 mutations were detected, covering ~95% of the samples. The tissues with the highest number of mutations were skin, lung and esophagus. Sun-exposed skin had more mutations than non-sun-exposed in Caucasians, but a similar number of mutations in African Americans. Somatic mutations were more frequent in older individuals, and this association was stronger in highly proliferative tissues. The set of mutated genes was significantly enriched for known cancer genes in multiple tissues. Cancer related genes also exhibited high ratio of nonsynonymous to synonymous substitutions. The most frequently mutated genes in the data were P53 and NOTCH1, known cancer genes, and they were found only in skin and esophagus. By analyzing the expression patterns of all genes in specific chromosomal arms, eight esophagus samples were detected with allelic imbalance in 9q.

Keren is opening a lab at the Technion, and intends to continue investigating somatic mutations. Research directions include somatic cells in single cells, and detecting somatic mutations without a reference genome from a different tissue.

The study is available here: https://science-sciencemag-org/content/364/6444/eaaw0726

Since the publication of this study, another study performing the same analysis was published by Hunter Fraser's lab in Stanford: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-019-1919-5. Keren's study was published while this one was under review. The authors of this study claim that they detected more than a hundred times more mutations at higher confidence (34% FDR compared to ~80% for Keren's study). This higher number of detected mutations allowed them to perform additional analyses, such as linking somatic mutations to gene expression and chromatin state.
Lianrong

Moran Yassour (HUJ) - The establishment of the infant gut microbiome in the first months of life

Moran Yassour’s group did a lot of work on gut microbiome. They investigate the establishment of the infant gut microbiome and the mother-to-child bacterial transmission. They also study the difference of gut microbiome between infants born by different delivery modes. In this talk, Moran presented two works:

(1) Analysis of the mother-to-child bacterial transmission pattern. They used a cohort of 44 mothers and their newborns, got their stool samples and used metagenomics sequencing methods to investigate the difference of microbiome communities within or across families. They developed a single-nucleotide variation profiling approach to compare the shared bacterial species between mother and her child. Previous studies usually only examined the dominant strain (most abundant haplotype detected by metagenomics sequencing) within a microbial species, while this study analyzed both dominant and “secondary” (less abundant) strains. They found that in most cases the mother’s dominant strain was transmitted to her child, while in some cases her secondary strain was transmitted. Further investigation of the secondary strain transmission suggested that the secondary strain may harbor specific genes that are missing in the mother’s dominant strain and advantageous in the infant gut.

This work was published on Cell Host & Microbe: [https://www.ncbi.nlm.nih.gov/pubmed/30001517](https://www.ncbi.nlm.nih.gov/pubmed/30001517)

(2) Investigation of the role of delivery mode on the infant gut microbiome. Multiple studies have reported that children born by C-section tend to lack Bacteroides species in their gut microbiome in the first few months of life, in contrast to vaginally born children. This talk focused on a special type of children that are born by non-elective C-sections, where the newborn was exposed to microbes in the birth canal, but was eventually extracted by section. They found that the microbiome community of non-elective C-sections children is similar to children born by C-section, namely with low Bacteroides species in the first few months. This suggests that the microbial signature is not driven by exposure to the maternal vaginal microbiome. Further analysis of mothers with and without antibiotic treatment shows that the microbial signature is mainly determined by the delivery mode, regardless of taking antibiotic treatment or not.
I attended the Virus and Phage Biology session since I am working on metaviome recently. The talks in this session focused mainly on biological discovery and had little about computational methods. Whatever, most of the topics are new to me and it is good to know about them. I describe one of the talks briefly in the following.

**Anat Herskovits (TAU) – Bacterial phage cooperation during mammalian infection**

Bacterial pathogens often carry multiple prophages and other phage-derived elements within their genomes, some of which can turn into deadly particles under stress. This talk introduced the *Lm* strain 10403S that carries two phage elements, one producing infective virions and the other producing bacteriocins (monocins). This talk mainly presented how these two phage elements are synchronized and adapted to the life style of their host. Previous studies have shown that genomic excision of virion serves as a molecular switch of *comK* gene expression during *Lm* infection of macrophage cells. Note that *comK* gene plays a role in the escape of *Lm* from the phagosomes and hence promotes *Lm* intracellular growth. Anat’s study revealed that monocin element plays a critical role in virion genomic excision and thus affects the expression of *comK* gene. Another new discovery in Anat’s study is that a metalloprotease, encoded by monocins, can serve as a global coordinator of the lytic induction of cohabiting phage elements. They also studied another *Lm* strain, *L. innocua* CLIP 11262 strain, and a similar mechanism was verified.

This study was published in Nature Communications: [https://www.nature.com/articles/s41467-019-13296-x](https://www.nature.com/articles/s41467-019-13296-x)

**Hagai**

**Elucidating disease mechanisms via their tissue-selectivity - Esti Yeger-Lotem, BGU**

The author discussed manifestation of heritable diseases in specific tissues. The germline genetic aberrations that lead to the disease exist across all tissues. However, the disease phenotype is manifested only in a single or a few tissues. For example, cystic fibrosis manifest mainly in the lungs. This raises the question: which mechanisms are responsible for regulating the manifestation of symptoms caused by the aberrated genes? The authors describe four families of such mechanisms:

(A.) Expression based mechanism: (1) expression-based mechanism that control the RNA expression of the aberrated gene. (2) tissue specific post-transcriptional mechanisms that
controls the likelihood that a transcribed RNA would be translated to a protein (e.g. by affecting its stability) (3) existing homologs that are differently activated in a specific tissue.

(B.) Regulatory mechanisms: (1) tissue-specific chromatin organization (e.g. tissue-specific open chromatin region). (2) tissue specific eQTLs (i.e., elements whose polymorphism correlates with variation in the expression level the genes). (3) tissue-specific chromatin contact (i.e., topologically associated domains (TADs) that can disrupt enhancer-promoter interaction.

(C.) Tissue-disrupted networks: (1) different pathways in which a broadly distributed gene is involved are activated in different tissues. (2) network’s gene modules that are active only in a specific tissue.

(D) Non-cell-autonomous mechanisms: (1) different microenvironment around the tissue (e.g., high hormone concentration around mammalian gland cells). (2) tissue specific differential receptor expression over the cell membrane that alters cell response for extracellular signal.

Following these insights, the authors designed a method to prioritize which gene are causal for the particular disease. The authors used multi-omics information to construct features for the manifestation mechanisms and GWAS data for heritable variants. The authors found that known causal gene were ranked high in the tissue the disease is manifested but not in others.

The paper discussing the Mechanisms was published in Nature Reviews Genetics: https://www.nature.com/articles/s41576-019-0200-9?proof=trueSouthampton. The paper about the method will be published in the next month.

Screening Human Embryos for Polygenic Traits Has Limited Utility - Shai Carmi, HUJI

This work evaluates the possibility to utilize genotyping of embryos for screening for traits. The authors used polygenic score (PS) as predictors for two traits, height and IQ. The authors evaluated the benefit of selecting the best scoring embryo by comparing the phenotype values of that embryo to the the average phenotype values. Annoyingly, the speaker did not mention at all the moral and ethical aspects of such selection, which smells very strongly of eugenics.
The author measured the improvement obtained by the selection in three ways: using a quantitative genetic model, by generating simulated embryo genotypes based on real data on genotypes and phenotypes of individuals, and by assessing the phenotypes of real individuals in a dataset of nuclear ultra-orthodox families with large numbers of offspring (10 on average).

The authors showed the expected gap between the phenotype values of the best scoring embryo and the mean of phenotype over the embryo increases slowly when increasing the number of embryos, and rapidly when the proportion of variance explained by PS ($r^2$) increases. Using the simulations over real dataset, assuming choice of the best out of ten embryos, the authors reported a gain of 2.5cm when selecting for height, and a gain of 2.5 points when selecting for IQ. On the large families’ data, the authors tested whether the high scoring individual is the one with the highest phenotype value. They reported that in only 25% of the cases the highest scoring individual by PS indeed had the highest phenotypic value.


**Tom**

**Genomic retargeting of p53 and CTCF in Oncogenesis - Ofir Hakim, Bar Ilan University**

Ofir’s goal is to identify transcription factors (TFs) that underlie transcription reprogramming during tumorigenic transformation. There are three regulatory layers of transcription: (1) chromatin accessibility, (2) TF binding, and (3) enhancer-promoter (EP) interactions. At first, Ofir questioned how we can associate between regulatory regions and their target genes. He presented the 4C technique to identify locus specific EP links. For example, focusing on the Zfp36l promoter (right), 4C finds all regulatory elements (Res) interacting with the promoter. The intensity of any such interaction is presented as a contact frequency. A Glucocorticoid Receptor (GR) ChIP-seq is done to show that regions involved in the contact are regulated by GR protein.

Then, Ofir focused on HiC technique, aimed for capturing all 3D contacts. Most of these contacts lie within topologically associated domains (TADs), which are mainly conserved across cell types ([Rao et al., 2014](#); left figure). Here, Ofir compared between two cell types, MCF10A (fibrocystic breast cell line) and G12V (glioblastoma cell line). For all tested genes above, the 4C contacts were mainly
conserved between the two cell types and were within TADs. Next, Ofir showed a decrease in SDC1 gene expression in G12V compared to MCF10A when the chromatin accessibility in G12V's RE is lost (right figure).

Testing the hypothesis that there is association between coordinated variation in chromatin accessibility and gene expression (GE) within TADs yielded a strong association.

Then, Ofir tried to find which TFs enriched in REs are associated with MCF10A-specific genes. Ofir showed that p53 TF is highly enriched in MCF10A-specific REs compared to G12V cells. These REs in MCF10A bound by p53 also showed high accessibility in ATAC-seq signal (bottom left figure).

Ofir questioned what happened in G12V REs and said that AP1 TFs may mediate p35 binding along with GR TF. A summary of the talk is available in the right figure.

**Gene regulation during neural differentiation - Anat Kreimer, Yosef and Ahituv labs, Berkeley and UCSF.**

Papers: (1) Identification and Massively Parallel Characterization of Regulatory Elements Driving Neural Induction, (2) Meta-analysis of massively parallel reporter assays enables prediction of regulatory function across cell types

Anat combined time series (0-72 hrs) datasets from epigenome, transcriptome, and functional assays to study neural differentiation. The model system is human embryonic stem cell (hESCs) to neural stem cell (NSC) differentiation. The differentiation from hESC to NSC is done by dual SMAD inhibition:
The questions asked, the tools used, and the output are described in the following figure:

Massively parallel reporter assays (MPRAs) were used to validate whether a candidate enhancer is active and functional (see paper 2):

Anat first analyzed the RNA-seq time series data (7 time points) after dual SMAD inhibition to induce neural differentiation. She applied k-means clustering (k=6) to find temporal changes across the time points. Various pluripotent TFs were found to be expressed in early time points (NANOG, POU5F1, SMAD7, ID1). Neural TFs were found to be expressed in later stages (SOX1, PAX6, OTX2, LHX5, IRF3, MAP2, ROR2).
Later, Anat did the same clustering procedure on H3K27ac ChIP-seq data, denoting active enhancers and promoters, and on chromatin accessibility ATAC-seq data, both on 7 time points (TPs). Anat showed a high overlap (by computing pairwise overlaps between each two data types – see figure below) between cluster 4 of ATAC-seq, clusters 5-6 of H3K27ac, and cluster 6 of RNA-seq GE. This might indicate that chromatin accessibility is acquired first, then H3K27ac are recruited to the open regions, and later followed by an increase in GE:

Next, Anat used MPRAs to scan functional enhancers involved in neural differentiation. She used the H3K27ac data to define the candidate regulatory elements (REs) for scanning enhancers (left figure below). Since enhancers are mostly involved at early stages, Anat clustered the MPRA enhancer signals into four clusters instead of six as previously done. The same was done for the other omics (right figure). Cluster overlaps between each two omics showed that temporal enhancers (67%) are consistent with endogenous temporal profiles.

Anat presented a method to calculate a TP-TF activity score. As shown below, motif scanning was performed on H3K27ac open regions. Then, regions with TF hits were intersected with MPRA enhancer regions, and regions with TF hits and MPRA signal showing an induced nearest gene
expression change from TP=t-1 to TP=t were counted and divided by the total number of open regions across all TPs:

Clustering the TF activity scores using k-means (k=4) revealed 26 TF candidates predicted to act after stem cell state (neural clusters 2-4).

For summary:

---

Additional notable talks are listed below without a summary (two were summarized above by Hagai). A link is given to slide shots taken.

- **Explicit probabilistic models of the 3D genome** - Noam Kaplan, Technion  
  Slides: [https://www.dropbox.com/s/s1w051m3tow9ecc/Noam_Kaplan.docx?dl=0](https://www.dropbox.com/s/s1w051m3tow9ecc/Noam_Kaplan.docx?dl=0)

- **Plenary talk: Towards global cooperation** - Martin Nowak, Harvard  
  Slides: [https://www.dropbox.com/s/d1mlscp5sf5uz1/Martin_Nowak.docx?dl=0](https://www.dropbox.com/s/d1mlscp5sf5uz1/Martin_Nowak.docx?dl=0)

- **Big data in health and disease: Challenges and opportunities** - Varda Shalev, Maccabi  
  Slides: [https://www.dropbox.com/s/fyqnci0xwtwr84/Varda_Shalev.docx?dl=0](https://www.dropbox.com/s/fyqnci0xwtwr84/Varda_Shalev.docx?dl=0)

- **Turning basic science into precision medicine offering with millions of people** - Yaniv Erlich, MyHeritage  
  Slides: [https://www.dropbox.com/s/6tp8l62lw2b9rv/Yaniv_erlich.docx?dl=0](https://www.dropbox.com/s/6tp8l62lw2b9rv/Yaniv_erlich.docx?dl=0)

- **Elucidating disease mechanisms via their tissue-selectivity** - Esti Yeger-Lotem, BGU  
- **Towards predicting drug effects using neural networks** - **Noam Shental**, The Open University
  Slides: [https://www.dropbox.com/s/pnw1ll2c4blih4/Noam_Shental.docx?dl=0](https://www.dropbox.com/s/pnw1ll2c4blih4/Noam_Shental.docx?dl=0)

- **Selecting embryos for traits: Is it possible?** - **Shai Carmi**, HUJI
  Slides: [https://www.dropbox.com/s/345vnhim2rlumpl/Shai_Carmi.docx?dl=0](https://www.dropbox.com/s/345vnhim2rlumpl/Shai_Carmi.docx?dl=0)