## Definition of a “landscape”

The concept of landscape was first introduced by Sewell Wright, as a metaphor to visualize a scale or score across a biological entity, such as a chromosome (Figure 1). In a linear biological entity, such as a chromosome; the biological entity is represented on the horizontal axis and the score, such as GC content, is represented on the vertical axis. This result in a line graph representing the GC content across the entire chromosome, which can then be used to identify GC content fluctuations across a chromosome.

### Intrinsic landscapes

In this section, 10 landscapes will be reviewed; namely, copy number landscape, mutational landscape, genetic/gene/genomic landscape, gene regulation/regulomic landscape, transcriptomic landscape, proteomic landscape, fluxomic landscape, metabolic/metabolomic landscape, molecular landscape, and phenotypic/phenomic landscape. The concept of landscape, as referred to the biological entity on the horizontal axis, tends to be more relevant in the first five landscapes; namely, copy number, mutational, genomic, regulomic, and transcriptomic landscapes; thus, true landscapes. In the other five landscapes (proteome, fluxome, metabolome, molecule, and phenome), nominal values (such as, different organelles) are represented by the horizontal axis; thus, quasi-landscapes. Despite so, it must be noted that value of “landscape” is enhanced when the horizontal axis is ordinal (such as, across a chromosome) rather than nominal (such as, different samples) as spatial distribution is implied in the former case.

### Copy number landscape

Copy number landscape often refers to copy number variation (CNV), which is the differences between the numbers of copies of one or more nucleotides within individuals in the same species. The repeats can range from a few bases; such as, polyglutamine repeats; to entire genes; such as alpha-amylase and neutrophil cytosolic factor. CNV can be a result when sections of a genome are duplicated/ repeated or deleted; which exists in both higher eukaryotes, such as humans, and prokaryotes, such as Escherichia coli; and is implicated in several disease phenotypes. For example, CNV in neutrophil cytosolic factor 1 has been implicated in rheumatoid arthritis and CNV in protease has been implicated in anthropogenic toxins in arthropods as differences in copy number may have an
Back-of-the-envelope guide (a tutorial) to 10 intracellular landscapes

Mutational landscape refers to the frequency of mutations and/or type of mutations across a biological entity, which may include CNVs and single nucleotide polymorphisms or SNPs. As mutation can be broadly defined as a permanent change in the nucleotide sequence, mutational landscape can be seen as the superset of all measurable nucleotide variations between individuals of the same species or across species. However, the term “mutation” tends to invoke a negative impression of “disease” and “errors”. Hence, mutational landscape often refers to sequence variations leading to diseases; such as, with reference to cancer. On the other hand, nucleotide sequence variations consistent with positive impressions; such as with respect to adaptation and tolerance; or neutral implications such as allelic variations within species; tends to be referred to as genomic landscape. However, the opposite may also be true as mutational landscape has also been used to describe natural variations between species. As such, these terms are synonymous.

Here, five studies on genomic/mutational landscapes are reviewed. Firstly, Xu et al. studied the genomic differences between two important Chinese indigenous cattle breeds with distinct phenotypes—Nanyang (Bos indicus) and Qinchuan (Bos taurus). The genomes of four Nanyang and four Qinchuan cattle were sequenced with 10 to 12-fold on average of 97.86% and 98.98% coverage of genomes, respectively. Using Bos taurus UMD 3.1 reference assembly as standard, 9010996 and 6965062 SNPs were identified for Nanyang and Qinchuan cattle, respectively. Of which, 51% of the SNPs in Nanyang cattle and 29% of the SNPs in Qinchuan cattle were novel. In addition, 154934 indels (1-3 bases) were found in Nanyang genome while 115032 indels were found Qinchuan genomes. These results suggest that Nanyang cattle showed more genetic diversity as compared to Qinchuan cattle. By analyzing the CNVs, the copy number of leptin receptor is significantly higher in Qinchuan cattle compared to Nanyang cattle, which may contribute to the higher fat deposition in muscles as observed in Qinchuan cattle. Secondly, Mata et al. sequenced 36 genes in 57 clinical samples from histologically confirmed classic Hodgkin lymphoma patients for mutations and found that 4 genes (CSF2RB, EP300, STAT6, and BTK) had mutation rates of more than 10%. These 4 genes are functionally related to B-cell receptor signaling pathway. This suggests B-cell receptor signaling as potential therapeutic targets. Thirdly, Babenko et al. used reference genomes in UCSC Genome Browser to generate a genomic landscape of CpG-rich elements in human and found strong correlation (r=0.97; P<1.1E-188) between open chromatin, using DNase hypersensitivity essay, and number of CpG islands. This further demonstrates the close association between CpG islands and chromatin states. Fourthly, Ruiz et al. used random mutagenesis to map out a landscape of essential genes in Bifidobacterium breve UCC2003, a commensal bacterium in human gut with probiotic properties. The essential genes in B. breve include (a) housekeeping genes; such as, DNA replication and transcription, manufacture of cell envelope components, the Sec-dependent protein translocation pathway; (b) genes in central pathways for energy acquisition and conversion; such as central glycolysis, a specific bifid-shunt, pentose phosphate pathway, and pathways crucial for production of energy and precursors for other metabolic routes; (c) genes encoding subunits of ATP synthase; (d) genes encoding enzymes involved in purine and pyrimidine synthesis; and (e) genes encoding proteins involved in ion transport and redox homeostasis. Lastly, Gneeczi-Rusconce et al. successfully genotyped 59 buccal swab from 4 communities near the Tibetan plateau. By analyzing the SNPs, a model to describe the migration and colonization of Southern Himalayan slopes from East Asia ancestry is proposed. However, this ideological segregation is not absolute as “genome landscape” is also been used to refer to diseases such as cancers.

Regulome refers to the whole set of regulatory components and network in a cell, including gene expression regulation; and protein and enzyme activation, inhibition, and degradation. Considering that the genome is the blueprint of the cell, regulome is equivalent to its operation manual. Hence, regulome is an integral part of all aspect of the workings of the cell, including transcriptome and metabolome; except and perhaps, the genome. Therefore, regulomic landscape covers an extensively wide area and is equivalent to deciphering the cell’s operation manual. Often, gene expressions (the transcriptome) are analyzed and grouped into sets of correlated expressions, known as metagenes. After which, the regulation of each metagenes were identified. The most common method is to analyze the promoters to identify the transcription factors for each gene as common transcription factor use have a tendency to result in correlated gene expressions; thereby, finding common transcription factors for each metagene. For example, Zhang et al. performed a 15-timepoint (from previtellogenesis to choriogenesis) NGS analysis of domestic silkworm (Bombyx mori) oogenesis. Six stages of oogenesis were found using correlated gene expressions, with each stage demonstrating a specific set of activities. Next, 761 transcription factors were identified and mapped onto each stage, resulting in a regulatory landscape for each of the 6 stages of oogenesis.

Transcriptomic landscape refers to the expression pattern across a chromosome or genome to examine the spatial distribution of expression, to elucidate genomic regions of specific level of transcriptional activities. Earlier studies tend to use “transcriptomic landscape” or “transcriptional landscape” synonymously with “transcriptome”. In certain cases where different transcriptomes were analyzed, such as transcriptomes of different tissues; “transcriptome landscape” may refer to transcriptional pattern across biological samples.

Here, three studies on transcriptomic landscapes are reviewed: (a) Severe alcoholic hepatitis are often treated with glucocorticoids and prednisolone but a portion of the patients do not respond within 7 days, with higher mortality rates in non-responders compared to responders. Sharma et al. studied the pre-therapy liver transcriptomes of Indian and Frech patients with severe alcoholic hepatitis using both microarrays and NGS. 1202 differentially expressed genes (1,106 upregulated genes and 96 downregulated in non-respondents relative to respondents) between livers from Indian patients. The upregulated genes from Indian population did not exhibit Gene Ontology enrichment but downregulated genes are under-represented.
in hepatocyte-specific metabolic pathways for xenobiotics by cytochrome P450, drugs, retinol and all-trans-retinoic acid, amino acids, steroids. In French patients, 207 differentially expressed genes were found (65 upregulated genes and 142 downregulated genes in non-respondents relative to respondents). The upregulated genes were enriched in the following Gene Ontology terms: “Hemoglobin complex”, “Oxygen transport”, “Heme binding”; while downregulated genes were enriched in Gene Ontology terms and KEGG pathways related to mitotic cell cycle, G1/S transition to mitotic cell cycle, and DNA replication. This suggests that genetic variations between French and Indians. However, differentially expressed genes were not spatially analyzed. (b) Ovariectomy has been used on livestock goats to increase mutton production but the molecular mechanisms are unknown. Zhang et al. studied the longissimus dorsi muscle transcriptome of ovariectomized Boer hybrid goats (Boer male goat × Guanzhong dairy female goat) 50 days after ovariectomization using NGS.62 By comparing against un-ovariectomized control goats, 376773 SNPs and 1612 differentially expressed genes (718 upregulated genes and 894 downregulated gene in ovariectomized goats) but no analysis was performed using spatial approach. These differentially expressed genes were of neuromuscular junction class, synapse assembly class, and sulfiredoxin activity; in terms of Gene Ontology; and mapped onto development and reproduction associated pathways in KEGG. Thus, providing a molecular basis for increased muscle growth in ovariectomized goats. (c) Li et al.,63 performed NGS on Mycobacterium smegmatis mc2155 and mapped its transcriptional activities onto its genome map; thus, producing the first transcriptomic landscape for Mycobacterium smegmatis mc2155. Using this, 2139 transcriptional start sites with 2233 independent mononicronic or polycistronic mRNAs within the operon/sub-operon structures were found. In addition, 8 highly active promoters were found and experimentally validated with β-galactosidase assays.

**Proteomic landscape** generally refers to the overview of protein activities; such as, protein-protein interactions and protein-molecule interactions, collectively known as protein interactomes;64 with or without reference to spatial organization within the cell; such as, intracellular locations or organelles. Hence, proteome/interactomic landscape tends to be represented as networks by the first group that coined the term “interactome”.65 For example, a study examining the proteomic landscape of cyclin E considered the interactomics of cyclin E across different tissues as proteomic landscape.66 Chiang et al.,67 examined the proteomic differences in 8-12weeks C57BL6/J mice’s suprachiasmatic nucleus, the master circadian pacemaker in mammals, across 6 regular time points (every 4hour intervals) within a 24-hour circadian cycle and found 20% (421 of 2112 accurately quantified proteins) demonstrating time-of-day-dependent expression profile. Hence, timing of the circadian cycle can be considered as the horizontal axis and each protein can be mapped as relative abundance on the vertical axis. When comparing to the corresponding time-dependent transcriptomes, Chiang et al.,67 found that more than 40% of each proteome was encoded by non-rhythmic transcripts. This suggests that the time-dependent transcriptome profiles are poorly indicated by the corresponding transcriptome profiles. Kole et al.,68 referred to “proteomic landscape” synonymously with “proteome” as they are concerned with differential protein abundance in primary somatosensory cortex across different treatments of sensory deprivation.69

Metabolomics and fluxomics are closely related. Metabolomics is concerned with the relative abundance of metabolites in a biological sample,70 putting metabolomics to be the proteomics and transcriptomics counterpart for metabolites. The advantage of metabolomics over transcriptomics or proteomics is the reduction of data elements – metabolomics has few components compared to transcriptomics or proteomics.71 Fluxomics is the rate of metabolic reactions.72 While metabolomics provides a snapshot of metabolism,73 fluxomics provides the activity of metabolism. In another words, metabolomics provides a photograph while fluxomics provides a video, and metabolomics is an integral part of fluxomics. A video is comprised of multiple “photographs” arranged in precise time intervals and it is always possible to extract a photographic image from a video. Hence, **metabolomic landscape** and fluxomic landscape are inter-connected and refers to the relative abundance of metabolites and the distribution of metabolic rates, respectively; with or without reference to spatial organization within the cell.

For example, Lien et al.,74 quantified the flux differences of wild-type *Pseudomonas fluorescens*, which does not produce alginate; and a double MucA and AlgC knockout as MucA knockout *P. fluorescens* is able to produce alginate75 but double MucA and AlgC knockout mutants are unable to produce alginate. The fluxomic landscapes were generated from multiple metabolomes of radioactive carbon labelling experiments, each representing a metabolomic landscape. Lien et al.,74 found that MucA-AlgC double knockout mutants are able to reorganize carbon flux without going through alginate production. Ahn et al.,76 combined liquid chromatography/mass spectrometry (LC-MS) isotope tracing with metabolite flux modeling to create a fluxomic landscape of tricarboxylic acid (TCA) cycle of HeLa cell line across cell cycle; and found that the concentration of central metabolites, such as cytidine triphosphate, oscillates across cell cycle. This is likely to be a result metabolite entry from glycolysis into TCA cycle is cell cycle phase dependent, resulting in oscillating metabolite concentrations in the entire central metabolism. Sabra et al. examined the fluxomic landscape of *Yarrowia lipolytica*, oleaginous yeast, for citric acid production at different substrate and dissolved oxygen concentrations in a fermenter.77 It was found that oxygen requirements differed based on different carbon sources resulting from differential fluxes towards TCA cycle and pentose phosphate pathway, which highlights the importance of oxygen demand based on different feed stocks.

Metabolomic landscapes, on the other hand, usually focus on samples rather than time course. For example, Heiland et al.,78 performed nuclear magnetic resonance (NMR) spectrometry on sonicated tissue extracts from 48 patients with primary glioblastoma multiforme WHO grade IV; thus, generating 48 metabolomic landscapes, and found 3 patient clusters by unsupervised clustering on the NMR results. Rabinowitz et al.,79 first mapped out the transcriptomic and proteomic landscapes along the proximodistal axis (from body to extremity) of Zebrafish caudal fin. Mass spectrometry data was used to generate a 3-position (proximal, middle, and distal) metabolic landscape of the causal fin where 26 proximally enriched metabolites and 16 distally enriched metabolites were found. Wound healing and cell proliferation amino acids;80 such as glutamate, arginine, and leucine were enriched at proximal end, suggesting that the proximal end of the fin is metabolically primed for rapid cell proliferation in event of injury. Fuhrer et al.,81 collected wild-type *E. coli* K-12 and 4320 single gene deletion mutants from KEIO knockout collection82 to perform mass spectrometry and generated 4321 metabolomic landscapes in order to decipher the metabolic effects of each gene knockout. One of the important conclusions was that the metabolic perturbations are significant within 2 metabolic steps from the deleted enzyme with diminishing significance beyond 3 metabolic steps. In a
way, the 4320 metabolomic landscapes are spatially oriented as they can be mapped onto the E. coli K-12 genome landscape.

Both metabolome and fluxome are the results of genome, transcriptome, proteome, and regulome; metabolome and fluxome, especially fluxome, can be seen as the layer directly underneath phenotype. As all various compound-based omics are molecules in nature, molecular landscape becomes the de facto umbrella term for genomic/mutational, transcriptomic, proteomic, regulomic, metabolomic, and fluxomic landscapes. It is also common to use the term “molecular landscape” as an abbreviation to represent the interconnectedness and inter-relationship across multiple landscapes. For example, Medico et al. used the term “molecular landscape” to represent genomic-transcriptomic-proteomic-metabolic landscape in human colorectal cancer. Schafer et al. used “molecular landscape” to represent genomic-transcriptomic-regulomic landscape in disease phenotypes. van de Vondervoort et al. used “molecular landscape” to represent genomic-proteomic landscape in obsessive-compulsive disorder. Neely et al. used “molecular landscape” to represent transcriptomic-proteomic landscape in clear-cell renal cell carcinoma. However, it is also possible to use “molecular landscape” to refer to a dominant landscape in a study rather than the variety of landscapes. For example, both Grimwade et al. and Bolouri et al. used “molecular landscape” to refer to primarily genomic landscape in acute myeloid leukemia.

Phenomic or phenotypic landscape refers to the phenotype observations on the vertical axis with respect to either changes in molecular landscape or changes to external stimuli on the horizontal axis. If the horizontal axis is ordinal or interval, the phenotypic landscape can offer an overview of the changes. Median lethal dose (LD₅₀) or minimum inhibitory concentration (MIC) graph are examples of phenotypic landscape as they show growth (the phenotypic observation) on the vertical axis again dosage on the horizontal axis. Hence, we are likely to see more phenotypic landscapes than we would have noticed. For example, Nichols et al. measured the growth (by colony size) of 3979 mutant strains from KEIO knockout collection against 324 conditions accounting for 114 unique stresses and found 179 genes that are essential for cell survival under specific conditions, which are known as conditionally essential genes. Mukherjee et al. isolated 232 non-Saccharomyces yeast strains from a wide variety of sources and evaluate their growth based on different stresses; namely, osmotic stress (glucose, fructose, and sorbitol), salt stress (sodium chloride, potassium chloride, and lithium chloride), ethanol stress, heat stress, fosfomycin derivative stress, and heavy metal stress (zinc chloride, copper sulfate, and cadmium sulfate); and found that wild Pichia kudriavzevii (VMU139) strain is able to tolerate most of the tested fermentation conditions.

Concluding remarks

Landscape provides a view of an ordinal biological entity, such as a genome or chromosome; or an ordinal condition, such as levels of chemical stress; or a nominal condition, such as different samples or tissues; yet, does not limit the resolution of the view. In this omics era, landscapes can provide a useful scaffold to collate available information on an organism. It has been widely known that one of the major usefulness of UCSC Genome Browser (https://genome.ucsc.edu/) is its ability to view multiple tracks at the same time. In this case, a track in UCSC Genome Browser is equivalent to a landscape. The value ahead is to be able to compare and interrogate across multiple landscapes.

Acknowledgements

None.

Conflict of interest

No conflict of interests results from the publishing of this article.

References


Citation: Ling MHT. Back-of-the-envelope guide (a tutorial) to 10 intracellular landscapes. MOJ Proteomics Bioinform. 2018;7(1):31–36.
DOI: 10.15406/mojpb.2018.07.00209


