Should I make it round? Suitability of circular and linear layouts for comparative tasks with matrix and connective data

E. Ståhlbom[1,2], J. Molin[2], A. Ynnerman[1] and C. Lundström[1,2,3]

[1]Department of Science and Technology, Linköping University, Sweden
[2]Sectra AB, Sweden
[3]Center for Medical Image Science and Visualization, Linköping University, Sweden

Abstract

Visual representations based on circular shapes are frequently used in visualization applications. One example are circos plots within bioinformatics, which bend graphs into a wheel of information with connective lines running through the center like spokes. The results are aesthetically appealing and impressive visualizations that fit long data sequences into a small quadratic space. However, the authors' experiences are that when asked, a visualization researcher would generally advise against making visualizations with radial layouts. Upon reviewing the literature we found that there is evidence that circular layouts are preferable in some cases, but we found no clear evidence for what layout is preferable for matrices and connective data in particular, which both are common data types in circos plots. In this work, we thus performed a user study to compare circular and linear layouts. The tasks are inspired by genomics data, but our results generalize to many other application areas, involving comparison and connective data. To build the prototype we utilized Gosling, a grammar for visualizing genomics data. We contribute empirical evidence on the suitedness of linear versus circular layouts, adding to the specific and general knowledge concerning perception of circular graphs. In addition, we contribute a case study evaluation of the grammar Gosling as a rapid prototyping language, confirming its utility and providing guidance on suitable areas for future development.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [Information interfaces and presentation]: User interfaces—Evaluation/Methodology

CCS Concepts
• ***Human-centered computing*** → ***Empirical studies in visualization; Visual analytics;*** *Visualization toolkits;* • ***Applied computing*** → *Genomics;*

1. Introduction

Circular graph layouts can be attractive for visualization designers. Bending the main axis into a circular shape is convenient for fitting long data sequences into a quadratic space. Genomics and bioinformatics data often consist of long sequences, which could be one explanation why circular plots are common in these domains. The high screen-space density is further articulated for connective data, also common in genomics, which can be represented as connective lines inside the circle in a spoke-like manner. Other reasons for choosing circular layouts could be that they are considered aesthetically appealing and innovative compared to a traditional linear alternative. However, visualization designs should primarily be evaluated on how well they support the analysis task at hand. Circular layout are used for a multitude of datatypes, including matrix data, which consists of rows and columns of scalar values, and connective data, which consists of a sequence of positions and interconnections between the positions. While comparisons between circular and linear plots have been made for other settings, we have

not found such studies for connective data and matrix data, which are two data types of particular relevance for genomics.

The main contribution of our presented work is an empirical study on the suitability of circular and linear layouts for comparative tasks with matrix and connective data. While the visualization challenge initially comes from genomics applications, the results are expected to be informative for these data types also in other settings. The tested visualization designs were developed in the Gosling framework [LWLG22], and the results include a review of our experiences from using Gosling for rapid prototyping.

2. Related work

The presented work is a foundational part of our visualization research agenda in genomics. This domain is nicely mapped by the survey and taxonomy definition provided by Nusrat et al. [NHG19]. Of particular relevance is their proposed genomics visualization layout categories: linear, circular, space-filling, and spatial. Drilling

down further, the arrangement within the plots can be categorized as parallel, serial, or orthogonal.

A key resource for genomics visualization designers is the Gosling visualization framework [LWLG22], which we have employed as our main prototyping tool. There are many examples of proposed visualization solutions for specific genomics analysis tasks (e.g., [vdBJvW*23, LKD*23, RFT*13, FNM13, RZZ*13]), as well as explorations of the genomics visualization design space [LG22, PLW*22].

A tool of particular importance for our work is the plotting tool *Circos* created by Krzywinski et al. [KSB*09]. Circos is popular in genomics publications [GET*13, HHA*09, DSX*23] but also utilized in other domains such as transportation and migration [ZFAQ13, Sti17]. It primarily targets visualization of variation in genomic structure and relationships between genomic intervals and is presented as a tool for rapid deployment in data analysis and reporting.

The suitability of circular and linear plots in genomics has been discussed previously. The software GenoRec [PLW*22] recommends visualizations for genomics data using a rule-based recommendation scheme. The authors recommend both linear and circular layouts for overview tasks searching for trends or patterns. There are indications that circular plots do not work well for tasks requiring zooming and panning. Moreover, for comparisons of length and position, the authors conclude that circular layouts are not useful citing Waldner et al. [WDG*20]. However, the relevance of the Waldner et al. study for genomics can be questioned, as it targets temporal (and thus cyclical) data and primarily focuses on rose charts, which is not what is typically used in genomics contexts. The circular layouts in genomics data are typically one or more tracks arranged in parallel or serially, often with interconnections running through the circle's center [NHG19], creating a chord diagram.

Outside of genomics-specific work, circular visualizations [DLR09] (also called *radial*) have attracted substantial research attention. Radial charts are often used for temporal data, due to their cyclical nature [WDG*20, BCPR22, CDBM22, MM18, BLIC19, BW14].

Some previous efforts have compared the two layouts, from different angles. One example is a comparison of smartwatch visualization designs [BBB*19] for regular bar charts, donut charts, and radial bar charts, finding that the radial visualization performed the worst. Performance of an area estimation task has been shown to be similar for circles and rectangles [HB10]. Goldberg and Helfman [GH11] evaluated linear versus radial bar, line, area, and scatterplots, outlining how scanning patterns differed and recommended that not too many concentric rings be used in a radial graph due to risk of confusion. Further details on these insights were made in the genomics setting: when adding more tracks to a circular layout the individual tracks become very small and if tracks are long part of the circle ends up out of the viewport [LG22].

Dielh et al. [DBB10] evaluated the performance of radial and linear layouts of matrix data for a number of tasks requiring navigation and comparison of elements. The results indicated that it is easier to remember positions in a radial coordinate system than

a cartesian one and that there are reading direction effects for the cartesian layout. They also found that task completion times for correct answers were significantly faster for the cartesian layout.

Our study complements the linear vs circular comparisons by investigating previously unchartered territory. We provide empirical evidence for the suitability of linear and circular layouts for two data types, matrix data (multiple stacked tracks) and connective data.

The suitability of visualization designs is of course dependent on the type of task at hand. In our experiments, the tasks were designed to resemble future real-world situations in clinical genomics. The taxonomy by Nusrat et al. [NHG19] concerns visualization tasks in genomics. This is an extension and adaptation of the more generic taxonomy of Brehmer et al. [BM13], where the motives from the means and the goal are disentangled by dividing the tasks into why, how, and what. Another relevant work is from Andrienko and Andrienko [AA06] where tasks are categorized as elementary or synoptic, where the synoptic tasks relate to trends and tendencies of a reference set while elementary tasks pertain only to one element.

In their work on visualizations for comparison, Gleicher et al. [Gle18] state that the difficulty of a comparison task depends on the number of items, item size and complexity, and the size and complexity of the relationship between items. They outline three strategies for comparison, which are scanning sequentially, scanning subsets, and summarizing somehow. They further identify that layout of the visualization can affect its effectiveness and identify this as an area for future research. L'Yi et al. [LJS21] recently conducted a survey on comparative visualization arrangements and suggested design principles for using juxtaposition, superposition, and explicit encoding for visualizations aimed at comparative tasks. Ondov et al. [OJEF19] compared different arrangements for data analysis, finding they were not able to extract specific guidelines per datatype, and called for more empirical studies on different datasets and tasks. To conclude, both data and task type affect what layout choices are appropriate, and layout in turn needs to be considered to ensure the effectiveness of the visualization.

3. Methods

In this section, we present the methods used within the study, starting with our prototyping approach followed by the building blocks of the user study.

3.1. Prototyping approach

We used the grammar for genomics data visualization Gosling to create the data visualizations, starting from the Gosling for React example provided on the Gosling GitHub. In the prototyping and system design processes the system developer kept a design diary. The visualizations were designed to best isolate the effect from having a circular or linear layout, rather than being designed for optimal performance of the study tasks. For this reason, we opted not to introduce interactive selection features such as brushing. The visualization designs are introduced in connection with the task descriptions below.

Biology basics

There is DNA in more or less every cell in the human body. It is built up of four bases: A, C, T, and G, attached to a strand made of sugar.

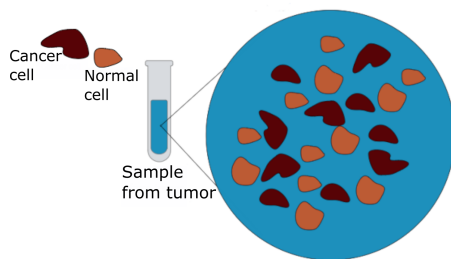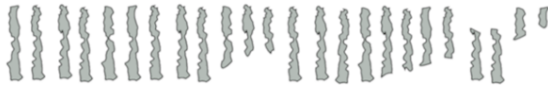The DNA of humans is divided into 23 DNA strands, called chromosomes.

Figure 1: *Introduction to the biological concepts of the study, which each participant was presented with at the beginning of the study. In addition, a text reading as follows: A specific base is referred to by its chromosome and the position it has on that chromosome. For example, Chr3:310478 refers to the 310478th base on chromosome three. Humans have mostly the same bases in the same positions but there are variations. When a base is different than the reference genome, this is called a variant, or in layman's terms, a mutation. Variants are frequent in cancer and since a cancer sample has both tumor cells and healthy cells, only a percentage of the sample will have a variation that is specific to the tumor cells.*

Figure 2: *Datapoints and missing values from the study sessions.*

Figure 3: *Contamination of one sample by another. This figure was used to introduce the contamination task, explaining how processing multiple samples in parallel can lead to contamination events.*

3.2. User study overview

The empirical results were elicited through a user study with 22 participants. To achieve consistency, the study sessions were conducted by the same session manager and the experiments used the same monitor. Participants were guided through the study through a web-based interface, built-in typescript with React, and a Python backend for file serving.

A study session started with the participant being provided with a brief introduction to the study and then a short introduction to the basic biology knowledge needed to understand the context of the tasks (see Figure 1). The participant was then introduced to one of the two tasks, with some background for the task and a simple example to illustrate the idea. Then the participant practiced the task with one of the conditions, with the possibility to see the correct answer for each case. The participants mostly discovered the interactions, but, if asked, the session manager detailed which were available. Next, the test was started where the participant worked through 9 cases. Practice and test for the second condition then ensued, followed by the same procedure for the second task. Task
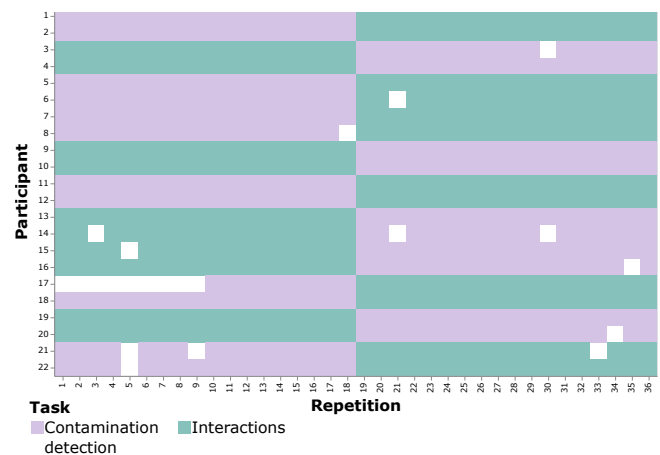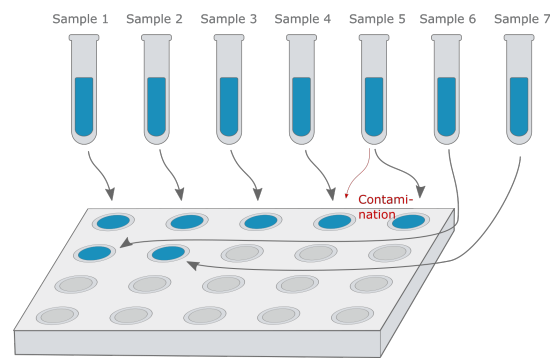
order and condition order were varied between participants to compensate for learning and fatigue effects.

Participants were instructed to take approximately thirty seconds to answer each task and to focus on their best estimation and general impression rather than being exact, but this was not a hard limit and they were allowed to take longer than thirty seconds. On a few occasions, the study session manager had to prompt the participant to make a choice and move on, after spending about 90-120 seconds on a single case.

Some data points had to be excluded due to technical issues causing repetition of the same case for several participants, two participants clicking past one case without realizing it, and the visualization tracks losing alignment for one case and participant. Additionally, participant 17 did not understand the task for the first condition, why those data points were excluded. Figure 2 displays the data points included in the analysis.

3.3. Task 1: Contamination detection

One of the experiment tasks simulated detection of contamination events, which is a quality control step in some genomics labs.
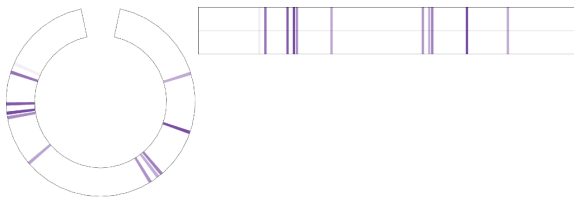
Figure 4: *The genomic variants of one sample shown in a circular and a linear layout. Genomic variants are indicated by the purple bars, with opacity encoding for concentration.*

As multiple samples are typically processed in parallel, spilling or other mistakes can lead to one sample contaminating another (see Figure 3). In the event of a contamination, the genomic variants from the contaminating sample will also occur in the contaminated sample, but at a lower concentration.

The variants of one sample can be visualized as a vector, as in Figure 4, where each purple bar indicates the presence of a variant, encoding the concentration by opacity. By stacking the variant vectors of all samples processed in parallel, it is possible to detect similarities and thus contamination events. The corresponding visualization designs for this task in the user study are shown in Figures 5 and 6. Users could pan and zoom using mouse interactions, and display a tooltip on hover.

The task consisted of determining if there was a contamination in the case and if so, to select the two samples involved. Thus, the abstraction of the task is pattern matching between rows, more specifically to find a row that where a fraction of another row's values has been added. According to the task taxonomy by Andrienko and Andrienko [AA06], the task is a comparison of the behavior of the same attribute between reference sets, and in the taxonomy of Nusrat et al. [NHG19] a *Summarize* task for multi-feature sets over multiple loci.

For this study, the data set size was chosen to make the task to be realistic and reasonably challenging: 30 samples and between 12 and 909 variants. The matrix was set to initially display all samples and 200 variants. The data simulation employed a 50% probability for a case to be contaminated, and the participants were informed about this probability prior to the test. This statistic is not very realistic but was selected to create a meaningful study task.

The data was simulated using numpy and pandas, see Figure 7. To generate the samples, integers between 0 and 100 were sampled from a uniform distribution, to create vectors of variants with different concentrations. A mask was created by randomly drawing positions to exclude from a uniform distribution for the number of variants. Contamination events were simulated by mixing 0.7 of the target sample with 0.3 of the contaminant and adding random noise to the result. Finally, only variants that occurred in exactly two samples were kept in the dataset. The choice of simulation parameters was derived through pilot studies aimed at a balanced difficulty, where the tasks would challenge participants but be possible to solve.

3.4. Task 2: Interaction comparison

The second task was to determine for a specified chromosome which (other) chromosome it had the most interactions with. Interactions can represent a number of biological features. The example provided to the study participants was that two regions are located in close proximity when the DNA is unwound in the cell between cell divisions.

In the visualization the chromosomes are spread along the main axis, and each chromosome is represented by a color-coded segment whose length represents its size (see Figures 8 and 9). Interaction between two positions was encoded with a blue arc connecting the positions. On hover, the hovered connections were highlighted and a tooltip showed which chromosomes they connected, and the visualization supported panning and zooming. Since we wanted to evaluate whether useful information could be extracted from the graph rather than the exact answer, we accepted the top three most interacting chromosomes, or more if there was a tie.

In terms of task abstraction, this constitutes a somewhat modified case of Andrienko and Andrienko's [AA06] task type comparison between attribute behavior over a specified reference subset (the specified chromosome) and attribute behaviors over other reference subsets (the rest of the genome). In Nusrat et al.'s taxonomy [NHG19] this interaction assessment corresponds to a multi-locus *Compare* task for a single feature set.

The data for this task is connective, meaning that each data point consists of two within-chromosome positions that are connected. In our data, each case had approximately 150-350 connections. The data was generated according to the process outlined in Figure 10. Each chromosome was assigned a number of connections from a normal distribution, and all negative values were capped to zero. Then for each connection, the target chromosome was drawn from a uniform distribution, as well as the origin and end positions within the chromosomes. We tuned the distribution parameters for the number of connections until the difficulty of the tasks would challenge participants but still be possible to solve, as for task 1. We ensured that all chromosomes asked about in the study task were chromosomes with many connections.

3.5. Differences between layout conditions

Some framework limitations resulted in slight differences between the two conditions. A vertical cursor line and row numbers were shown only for the linear condition. The mouse-over effect on the circular layout interactions visualization required holding down the Alt-key. In the linear layout, when hovering just on the edge of the interaction visualization, the mouse-over highlighted some random connections.

For comparison fairness, we designed the sizes of the two layouts so that they took up approximately the same amount of screen space. This resulted in the contamination visualization being rather wide for the linear layout.

3.6. Statistical analysis

No formal power analysis was performed, instead, the number of participants was set to a number similar to other two-condition
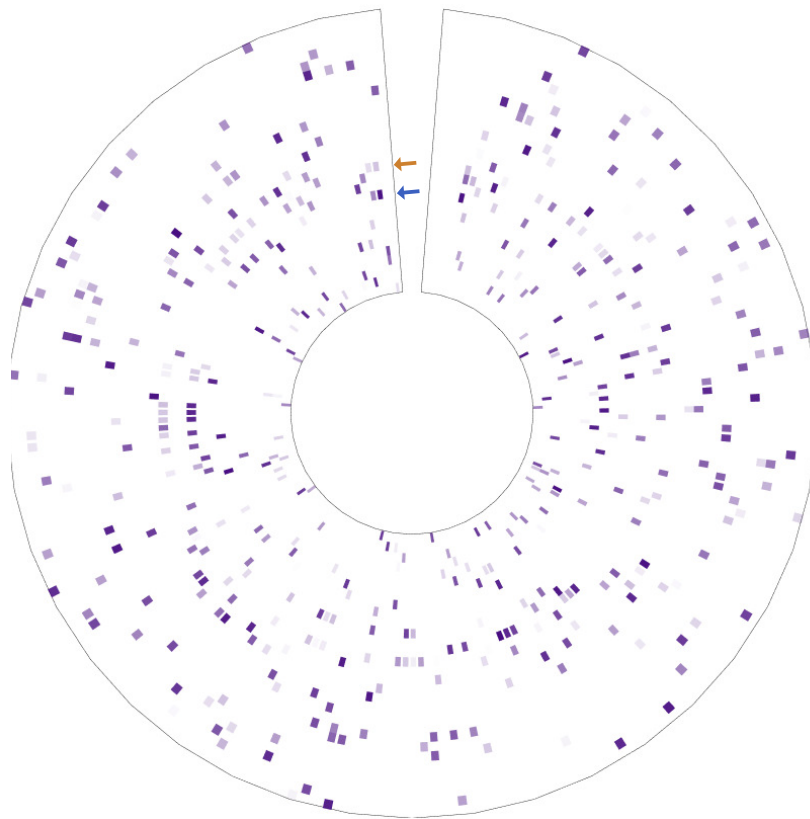
Figure 5: *Variants from multiple samples stacked in a circular layout. A contamination can be seen between two samples which for illustration purposes here are denoted by the arrows.*
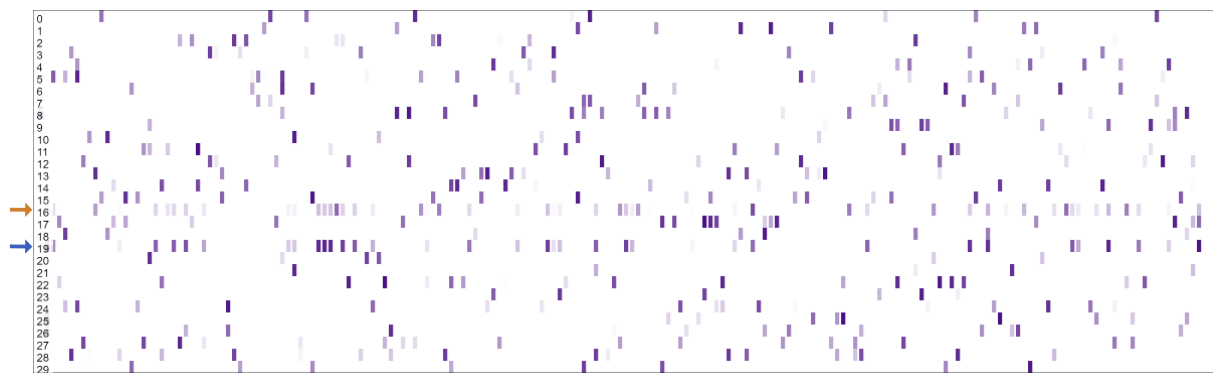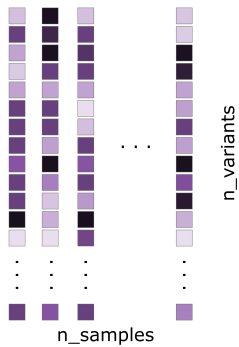

Figure 6: *Variants from multiple samples stacked in the linear layout. There is a contamination between samples 16 and 19 which for illustration purposes here are denoted by the arrows.*

Samples are generated
Variant concentrations are drawn from a uniform distribution from 0 to 100.

For each sample a mask is generated
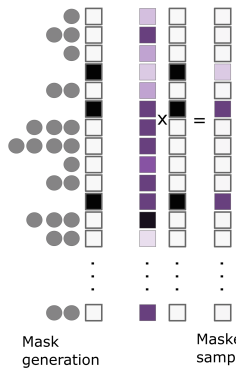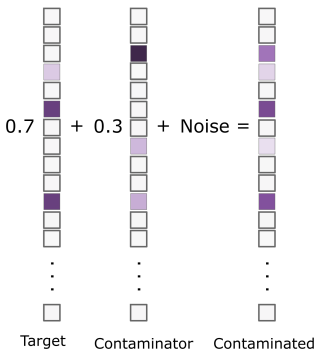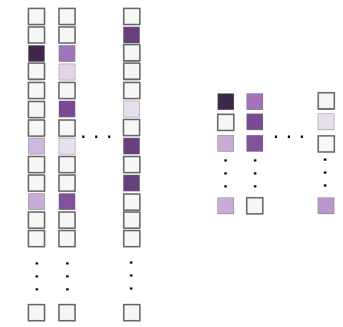Positions to remove are drawn from a uniform distribution.

Calculation of a contaminated sample

Only variants that occur exactly two times are kept

Figure 7: *The data generation process for the contamination detection task. Shades of purple represent variant concentration, with white denoting full absence. Masks are in black (value 1) and white (value 0).*
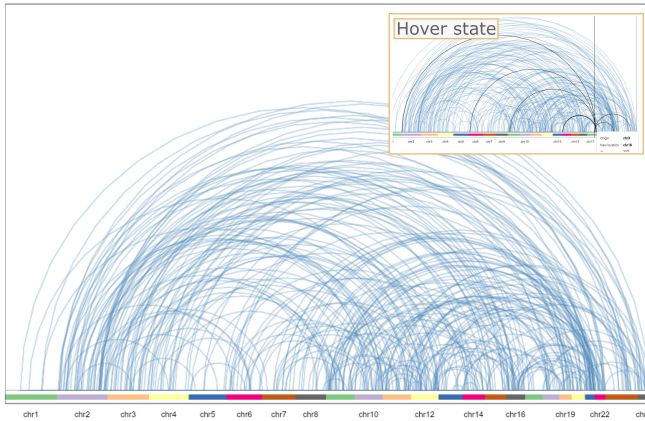

Figure 8: *The linear layout for the connective data as visualized in the user study. Chromosomes span the horizontal axis. The task is to select one of three chromosomes that have the most interactions with chr2. Correct answers are chr19, chr17, chr14, or chr13. There are four answers due to a tie in the number of interactions.*
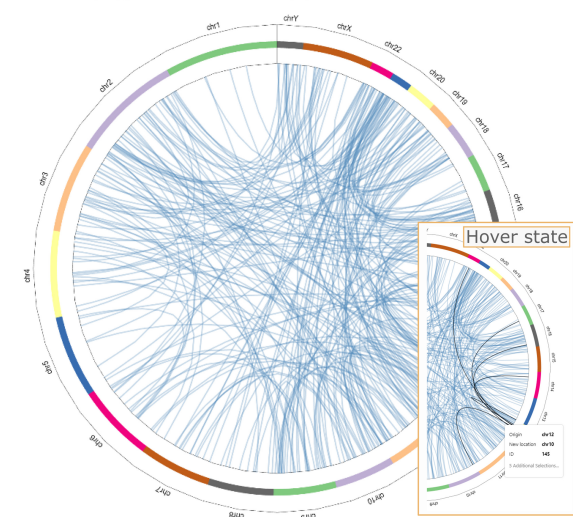

Figure 9: *The circular layout for the connective data as visualized in the user study. Chromosomes span the outer rim.*

studies and the number of repetitions was set to the maximum possible within the study time. We calculated the task completion time for each case and participant. The two tasks were analyzed separately and treated as independent datasets. Since we had multiple repetitions for each condition we applied linear mixed effects models and then used informed maximal random-effect structures for hypothesis testing using ANOVA. For each task, we fitted a linear mixed effects model using random slopes per participant and random intercept per case to model the effect the condition had on task completion time. The analysis was done using *R* and *lme4*.

For the accuracy, we fitted a general linear mixed effects model for the binomial distribution. For the contamination detection task,

we used random intercepts per case and random slopes per participant. For the interactions task the above parameters generated a singularity warning, leading to us excluding the intercept per case to ensure stability of the model. Both settings yielded the same significance level.

3.7. User study participants

Participants were recruited from one IT company and one university division, with a majority working in software development or visualization research. Though the tasks were inspired by genomics, the underlying tasks can be performed on data from other
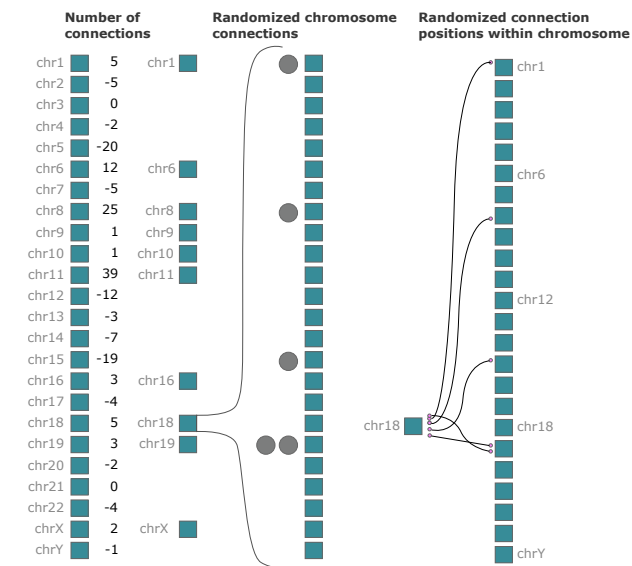
Figure 10: *Simulation of connection data. The number of connections is randomized from a normal distribution, and only positive values are kept. For each connection, a target chromosome is drawn from a uniform distribution, and then the positions within the origin chromosome and target chromosome are pulled from uniform distributions. This is done for all chromosomes in the sample.*

Figure 11: *Mean task completion time and mean accuracy. There is a clear difference in accuracy between the conditions for the contamination detection task and a trend towards linear layouts resulting in shorter completion times.*

Figure 12: *Point estimations and 95%-confidence intervals for the difference between linear and circular layouts in accuracy and time spent on each case. The contamination detection has a higher accuracy value for linear representations, leading to a difference in accuracy that is above zero. For the time spent on each task, both point estimates indicate a slightly faster completion time for linear layouts, but the results are inconclusive.*

domains. By sampling a population that was not geneticists, we strengthen the claim to generalization outside of genomics. We also opted for non-geneticists since that allowed us to recruit more participants to the study within a shorter time frame, strengthening the statistical power of the results. Finally, participants' unfamiliarity with the subject allowed us to introduce the data and task with minimal previous knowledge of the domain influencing their mental model, which might not have been the case for domain experts.

4. Results

In this section, we present quantitative and qualitative results from the user study. We also report on using Gosling for rapid prototyping in general and for our case in particular.

4.1. Task performance comparison

The mean accuracy and task completion time are presented in Figure 11. It shows that the linear layout yields more accurate results than the circular for the contamination detection task. For the interaction task, the circular yielded slightly more accurate results. As for time, the two conditions are similar in completion time for both tasks, but with the linear being slightly faster in both cases.

For the contamination detection task, the mean accuracy was 0.690 for the circular layout and 0.851 for the linear layout. The difference was statistically significant ($\chi^2(1) = 6.86, p < .05$). For the interactions task, the mean accuracy was 0.792 for the circular layout and 0.764 for the linear layout, with no statistical significance. The mean completion time for the contamination detection
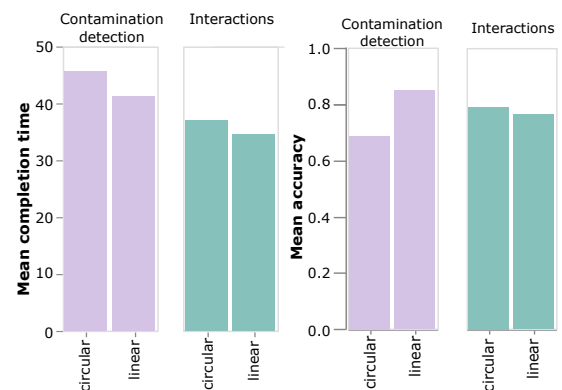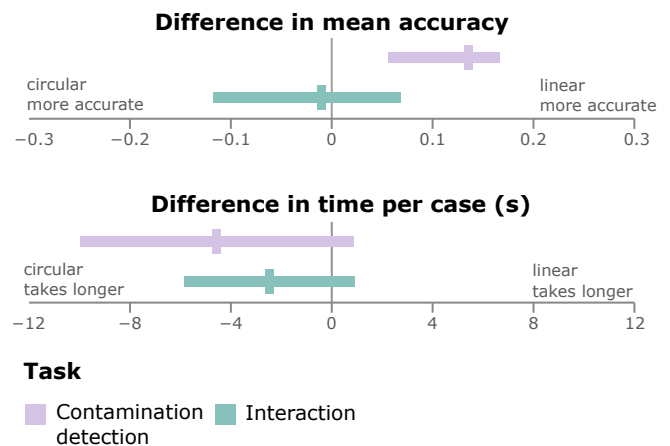
task was 45.7 seconds for the circular layout and 37.0 seconds for the linear layout, with no statistical significance. For the interactions task the mean completion time was 41.3 seconds for the circular layout and 34.5 seconds for the linear layout. The difference was not statistically significant. The estimated difference in intercept from a model including conditions and a model without conditions are displayed in Figure 12. The 95% confidence intervals are plotted along with the point estimates.

4.2. Task-related feedback and observations

The participants' answers to the questionnaire filled out after the study are presented in Figure 13. Note that participants' experience with visualization is self-reported, and we observed a skewness that
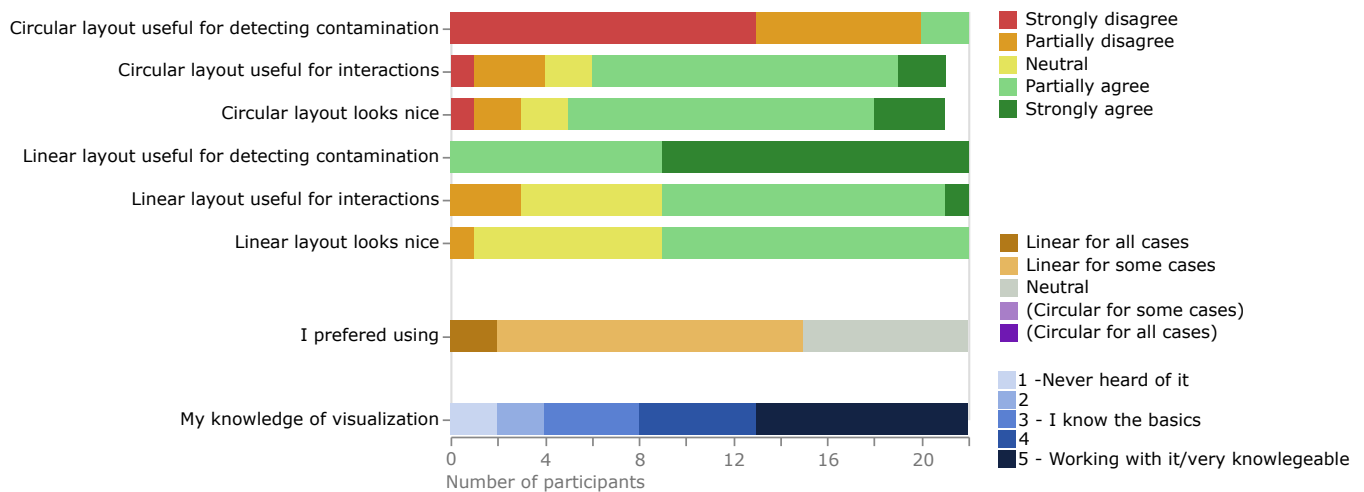
Figure 13: *Answers to the post-study questionnaire. There are two missing datapoints in the second and third row respectively due to responses missing. Alternatives selected by no participant are omitted from the color scale.*

participants were unwilling to select a high score. Below we describe insights generated from answers to the questionnaire free text questions and notes from the study sessions of observations and participant comments.

Participants strongly preferred the linear layout to the circular one for the contamination task. Reasons cited were difficulty following a row in the circular layout, difficulty following a column in the circular layout, neck pain from twisting the head, and difficulty comparing samples close to the center with samples far from the center due to differences in size. Some participants also had difficulties interacting with the circular layout through panning and zooming.

Multiple participants questioned the appropriateness of making this visualization circular. We observed different strategies for interaction, including using no interaction, panning back and forth, zooming in and out, or panning through the data and only looking at one part of the graph. For many participants, we observed difficulties in clicking the correct elements when selecting a contamination event. For the linear layout, a majority of participants appeared to scan the visualization from left to right, which was not observed as many times or as clearly for the circular layout. Participants also reported they could not scan all rows from left to right in one go, but needed to vertically subdivide the graph, making it more difficult to spot contamination events between samples on rows far apart. In addition to sequential scanning, a common strategy was to identify suspect samples that had many variants with low opacity or to look for clusters of variants to use as a starting point for match searching.

For the interactions task, there was not such a clear preference among participants. Some preferred one layout over the other, but many also saw different strengths and weaknesses with the two. There was a slight preference for the circular layout but generally expressed with less emphasis than the contamination task preferences. The difficulty of finding the correct chromosome was mentioned by some participants, both due to not all being labeled and

due to the labels being upside down on the lower half of the circle. Multiple participants mentioned using a strategy of trying to gauge in which direction most connections were going and then looking closer into that general area. Some participants thought it was easier in the circular layouts since directions were more varied, while one expressed that the edge chromosomes in the linear layout were easiest since there were fewer back-and-forth eye movements. Several participants also stated that the circular layout was nicer to look at, one calling it more "playful" with eye movements going not just back and forth but in all directions.

Many participants relied heavily on the highlighting on hover and the tooltip to make their decision. They moved the mouse back and forth over the chromosome and tried to see where most connections were by detecting "where it rained", as one participant put it. Several participants complained about heavy memory workload for this task, trying to remember where one had seen most highlights and not being able to focus on the entire genome at once. It was also difficult to make sure the mouse only hovered within the specified chromosome, and bundles of connections sometimes connected to the border between two chromosomes. For the linear layout, some participants thought that connections going to far away chromosomes were easier to see than others, causing skewing in their perception of the data.

Almost every participant described the tasks as very difficult, in line with the intention of the study design to result in far from perfect accuracy. However, several participants expressed displeasure with their own performance, indicated that they were not sure of their answers, and expressed fatigue due to the heavy workload on their working memory. When being shown their answers, most participants were surprised to see they performed better than assumed, and several expressed disbelief that anyone would be working with the visualizations they were shown. We believe this indicates that the tasks were just shy of being too difficult for participants to make an honest effort to solve.

4.3. System improvement feedback

Participants suggested improvements to our visualizations that would simplify the task. Most commonly suggested was highlighting clicked rows in the contamination detection visualization. This was especially requested for the circular layout since it was difficult to follow the rows when bent. Similarly, the vertical cursor line in the linear layout was requested for the circular as well, and grid lines to help with navigation. Making sure you stayed on the same row or in the same column was one of the most cited difficulties for the contamination detection task with a circular layout, both due to the bending of the rows and the area distortion along the columns (radii).

For the interactions task the most common request was brushing or filtering to highlight all connections originating from a specific chromosome. Many participants also requested for the tooltip to always show the two connected chromosomes with the closest one first, to facilitate quick scanning.

Another theme for potential improvements was better solutions to handle data sets extending outside the view. For the interactions task, some users expected that when zooming in, connections should remain visible even if one endpoint moved outside of the view, which was not the case for the circular layout. The panning interaction could also be improved since several participants struggled with using it effectively. For both layouts during the contamination detection task, some participants expressed disappointment that the zooming only scaled along the rows and that zooming out did not create a dense heatmap but rather resulted in very narrow bars.

4.4. Rapid prototyping experiences using Gosling

In this work, we used the genomic grammar-based toolkit, Gosling, as it provides a domain-specific rapid prototyping environment needed to support the study. Our intention was to document the experiences of using Gosling and thus a detailed diary, from the developer's perspective, was kept. In this section, we present a thematic analysis of the entries in this design diary. In this, our goal was to define best practices for ourselves and for future users, as well as provision of specific feedback for improvement of Gosling, to create added value for the community of Gosling users. The results of the analysis described consist of both the derived themes as such, potentially relevant for rapid visualization prototyping in general, and feedback specifically regarding Gosling and the genomics context. The four themes are time spent, comprehensiveness, usability, and data handling.

Time spent: Rapid prototyping, by definition, requires that the developer can accomplish their visions quickly. The design diary entries reflect a rapid initial pace, decreasing as the prototype matured. Towards the end of the design process, entries in the diary point out the spending of long time periods on making refinements, attempting to add features, and doing debugging, perhaps indicating that the design reached the limits of the grammar's specificity.

Comprehensiveness: The design diary showed that while prototyping is not the same as developing a fully working solution, the extensiveness and functions of the prototyping tool affect the developer's experience. Particularly towards the end of the design process, some expected features were not supported yet, or only supported for linear layouts. Difficulties making the system reactive were mentioned in the design diary on multiple occasions and it was difficult to infer from the documentation which features were supported. In addition, the library is rather new and has not yet gathered a community online with questions and answers in forums to provide guidance. On the other hand, a Gosling developer engaged when we posted a question on the GitHub forum, a responsiveness particularly important in the absence of a larger community. To summarize, while a prototyping tool cannot be expected to be entirely comprehensive, missing features and a limited set of help resources do affect the experience of using it.

Usability: It was clear from the design diary that a well-crafted prototyping grammar can be very effective, especially when examples are provided. It was possible to design using the grammar without constantly referencing the documentation after a day of using it. Overall, the grammar was intuitive, and by using the provided examples as starting points, visualizations could be produced in a short time.

Data: Handling of data was one of the most mentioned problems at the beginning of the design diary. Firstly, loading a local dataset, extra important for sensitive genomics data, required setting up a local server. Handling of large data was also mentioned, although Gosling provides a solution for this using HiGlass [KAL*]. Formatting the data correctly was challenging, but simultaneously enforced reasonable formatting. When the data was correctly formatted, the data reading worked well. The main unsolved problem related to data transformation such as filtering. This is supported within Gosling but only worked for some data fields. Moreover, though covered in the taxonomy by Nusrat et al. [NHG19], region abstraction was not supported. Thus, data curation and formatting remain challenging for prototyping genomics visualizations.

5. Discussion

Our results show that performance for row-wise comparisons is better for linear layouts than circular ones. While not a surprising conclusion, our results broaden and strengthen the generalizability of this claim, by showing it holds for comparison tasks in matrix data. This is in addition to previous insights regarding small displays [BBB*19] and radial bar, line, area, and scatter plots [GH11]. The importance of relying on empirical evidence is underlined by the existence of earlier research findings in favor of circular layouts. For example, it has been argued that it is easier to remember the position for circular layouts [DBB10], which indicates that circular layouts might be better for comparison than linear layouts since comparison requires remembering and matching patterns. Our results oppose this conclusion, thus adding nuance to the understanding of the utility of circular layouts.

The contamination detection task involved comparing a rather high number of large items (rows) to find a moderately complex relationship. According to Gleicher et al. [Gle18] this comparison task should be rather difficult, which is in line with the participants' descriptions of the tasks. In addition, it was difficult to compare all parts of all items at once, leading to sequential or subset

scanning of the dataset. By finding one suspect sample and searching for matches participants reduced the task to a series of two-element comparisons, which yields other opportunities for visualization [Gle18], for example, a difference log, sorting, or highlighting of the selected sample as suggested by some participants. A possibility is to order rows by clustering algorithms and apply the guidelines by L'Yi et al. [LJS21] for comparative visualizations.

In genomics applications, it is common to stack many tracks in a circos plot and add interconnections in the middle [GET*13, LKO*13]. Our results indicate that stacking of tracks might make it difficult to compare them, and we thus echo the advice by L'Yi and Gehlenborg [LG22] to not add too many tracks to circular plots. Navigation along a row was mentioned to be difficult in the circular layout for the contamination detection task, in line with the findings by Diehl et al. [DBB10] that remembering the row was challenging for circular layouts. While they found good performance for remembering an angular position, our study participants expressed that it was difficult to follow one column in the circular layout, indicating that angular discrimination might be difficult for detailed data. Navigation was mentioned also for the interactions task, concerning finding the chromosome specified, which was perceived as difficult, especially for the lower half. It appears that when navigation is necessary, circular layouts perform worse, extending the findings of Goldberg and Helfman [GH11] to two new graph types.

For the interactions task, our results indicate that the accuracy was similar for linear and circular layouts. Though not statistically significant, the task completion times were slightly longer for the circular layout for both tasks. The number of participants in the study permitted detection of larger differences in time per case and accuracy but would be unable to detect smaller differences as can be seen by the confidence intervals in Figure 12. Thus, there is a weak indication that linear layouts should be used if speed is of the essence. However, comments from participants that the circular layout looked better, indicate an aesthetic value in circular layouts. Participants described it as more playful, and allowing nicer eye movements than the linear one. Since the accuracy does not seem to be affected by the layout, selecting a circular layout is reasonable for contexts where the aesthetics of the graph are important.

Our experience is that Gosling generally worked well for rapid prototyping, and we expect it to become more useful as more features are supported and as the community grows. One area that we would suggest prioritizing is the circular layout, completing the feature support, and improving the interaction design. We believe the key characteristics making Gosling effective as a rapid prototyping tool were the intuitive grammar and availability of examples, in combination with the ease of modifying layout and encodings, hinting that these characteristics could be important to consider also for other rapid prototyping tools.

We acknowledge that the participant sample imposes limitations on the study. Any general effect on performance caused by the participant's backgrounds should be accounted for by the paired experiments, however, there might be biases within this population for or against the conditions. The visualization researchers in particular could have a negative view of the circular condition due to popular opinion within the field. Though we took care to minimize the background information needed to complete the tasks, it is pos-

sible that exposure to a new field and datatype caused fatigue or confusion about the task. The time restriction limits the relevance of the results to more superficial review scenarios. Our effort to ensure generalization outside of genomics naturally results in less specific claims within genomics, and we acknowledge the possibility that a sample from a population of geneticists could yield a different result. However, the study is aimed at the perception of these types of visualizations, and therefore our opinion is that the findings likely generalize to geneticists too.

The limitations of the Gosling framework also impacted the study. Adding a cursor line to the circular layout as requested by participants, might have improved performance, however, we believe the cursor in the linear layout did not affect the performance much. The difficulties interacting with the circular layout also highlight the challenge of designing intuitive and well-working solutions for circular graphs. We decided to use Gosling as the limitations were deemed not to invalidate the experiments, and due to the strong benefits of building on existing visualization knowledge represented by the framework. Moreover, starting from existing software is practicing responsible use of research resources.

A suggestion for future work is to investigate a combination of matrix and connective data. Circos plots with multiple stacked tracks and connections within are not uncommon. It could be interesting to evaluate how completion time and accuracy are affected by linear versus circular layouts in that case. Other design alternatives to evaluate include adding external connections to the circular layouts, in comparison to two-sided connections in the linear case. Another evaluation could be to investigate the performance of a more interactive tool with all the features suggested by participants.

Finally, within this study, we briefly touched on the possibility of displaying genomics data without having genomic position on any of the axes. Our contamination visualization with variant IDs on one axis independent of their genomic position, was a first step in this direction, and we believe there are interesting visualization opportunities to be found there.

6. Conclusions

We performed a user study comparing linear and circular layouts for two comparative tasks inspired by genomics. For matrix data, we found the linear layout to lead to higher accuracy, while there was no difference for interconnection data. For both datatypes, the results weakly indicate that linear layout resulted in shorter task completion times. We contribute a strengthened and expanded understanding of when linear layouts are preferable over circular ones. In addition, we contribute an evaluation of the Gosling grammar as a tool for rapid prototyping.

7. Acknowledgements

References

[AA06] ANDRIENKO N., ANDRIENKO G.: *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006, pp. 47–161. doi:10.1007/3-540-31190-4_3. 2, 4

[BBB*19] BLASCHECK T., BESANÇON L., BEZERIANOS A., LEE B., ISENBERG P.: Glanceable visualization: Studies of data comparison performance on smartwatches. 630–640. doi:10.1109/TVCG.2018.2865142. 2, 9

[BCPR22] BASTARDO R., CASTRO M., PAVÃO J., RAMOS L.: A timeline model for clinical events: empowering data. 1053–1058. doi:10.1016/j.procs.2021.12.109. 2

[BLIC19] BREHMER M., LEE B., ISENBERG P., CHOE E. K.: Visualizing ranges over time on mobile phones: A task-based crowdsourced evaluation. 619–629. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi:10.1109/TVCG.2018.2865234. 2

[BM13] BREHMER M., MUNZNER T.: A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics 19* (12 2013), 2376–85. doi:10.1109/TVCG.2013.124. 2

[BW14] BURCH M., WEISKOPF D.: On the benefits and drawbacks of radial diagrams. In *Handbook of Human Centric Visualization*, Huang W., (Ed.). Springer, 2014, pp. 429–451. doi:10.1007/978-1-4614-7485-2_17. 2

[CDBM22] CHEN K.-T., DWYER T., BACH B., MARRIOTT K.: Rotate or wrap? interactive visualisations of cyclical data on cylindrical or toroidal topologies. 727–736. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi:10.1109/TVCG.2021.3114693. 2

[DBB10] DIEHL S., BECK F., BURCH M.: Uncovering strengths and weaknesses of radial visualizations—an empirical approach. 935–942. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi:10.1109/TVCG.2010.209. 2, 9, 10

[DLR09] DRAPER G., LIVNAT Y., RIESENFELD R.: A survey of radial methods for information visualization. 759–776. doi:10.1109/TVCG.2009.23. 2

[DSX*23] DIESH C., STEVENS G. J., XIE P., DE JESUS MARTINEZ T., HERSHBERG E. A., LEUNG A., GUO E., DIDER S., ZHANG J., BRIDGE C., HOGUE G., DUNCAN A., MORGAN M., FLORES T., BIMBER B. N., HAW R., CAIN S., BUELS R. M., STEIN L. D., HOLMES I. H.: JBrowse 2: a modular genome browser with views of synteny and structural variation. 1–21. Number: 1 Publisher: BioMed Central. doi:10.1186/s13059-023-02914-z. 2

[FNM13] FERSTAY J. A., NIELSEN C. B., MUNZNER T.: Variant view: visualizing sequence variants in their gene context. 2546–2555. doi:10.1109/TVCG.2013.214. 2

[GET*13] GOECKS J., EBERHARD C., TOO T., THE GALAXY TEAM, NEKRUTENKO A., TAYLOR J.: Web-based visual analysis for high-throughput genomics. 397. doi:10.1186/1471-2164-14-397. 2, 10

[GH11] GOLDBERG J., HELFMAN J.: Eye tracking for visualization evaluation: Reading values on linear versus radial graphs. 182–195. doi:10.1177/1473871611406623. 2, 9, 10

[Gle18] GLEICHER M.: Considerations for visualizing comparison. 413–423. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi:10.1109/TVCG.2017.2744199. 2, 9, 10

[HB10] HEER J., BOSTOCK M.: Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2010), CHI '10, Association for Computing Machinery, pp. 203–212. doi:10.1145/1753326.1753357. 2

[HHA*09] HEINKEN A., HERTEL J., ACHARYA G., RAVCHEEV D. A., NYGA M., OKPALA O. E., HOGAN M., MAGNÚSDÓTTIR S., MARTINELLI F., NAP B., PRECIAT G., EDIRISINGHE J. N., HENRY C. S.,

FLEMING R. M. T., THIELE I.: Genome-scale metabolic reconstruction of 7,302 human microorganisms for personalized medicine. 1320–1331. Number: 9 Publisher: Nature Publishing Group. doi:10.1038/s41587-022-01628-0. 2

[KAL*] KERPEDJIEV P., ABDENNUR N., LEKSCHAS F., MCCALLUM C., DINKLA K., STROBELT H., LUBER J. M., OUELLETTE S. B., AZHIR A., KUMAR N., HWANG J., LEE S., ALVER B. H., PFISTER H., MIRNY L. A., PARK P. J., GEHLENBORG N.: HiGlass: web-based visual exploration and analysis of genome interaction maps. 125. doi:10.1186/s13059-018-1486-1. 9

[KSB*09] KRZYWINSKI M., SCHEIN J., BIROL , CONNORS J., GASCOYNE R., HORSMAN D., JONES S. J., MARRA M. A.: Circos: An information aesthetic for comparative genomics. 1639–1645. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab. doi:10.1101/gr.092759.109. 2

[LG22] L'YI S., GEHLENBORG N.: Multi-view design patterns and responsive visualization for genomics data. 1–11. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi:10.1109/TVCG.2022.3209398. 2, 10

[LJS21] LYI S., JO J., SEO J.: Comparative layouts revisited: Design space, guidelines, and future directions. 1525–1535. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi:10.1109/TVCG.2020.3030419. 2, 10

[LKD*23] L'YI S., KELLER M. S., DANDAWATE A., TAING L., CHEN C.-H., BROWN M., MEYER C. A., GEHLENBORG N.: Cistrome Explorer: an interactive visual analysis tool for large-scale epigenomic data. *Bioinformatics 39*, 2 (01 2023). btad018. arXiv:https://academic.oup.com/bioinformatics/article-pdf/39/2/btad018/49096012/btad018.pdf, doi:10.1093/bioinformatics/btad018. 2

[LKO*13] LEMIEUX J. E., KYES S. A., OTTO T. D., FELLER A. I., EASTMAN R. T., PINCHES R. A., BERRIMAN M., SU X.-Z., NEWBOLD C. I.: Genome-wide profiling of chromosome interactions in plasmodium falciparum characterizes nuclear architecture and reconfigurations associated with antigenic variation. 519–537. doi:10.1111/mmi.12381. 10

[LWLG22] L'YI S., WANG Q., LEKSCHAS F., GEHLENBORG N.: Gosling: A grammar-based toolkit for scalable and interactive genomics data visualization. *IEEE Transactions on Visualization and Computer Graphics 28*, 1 (2022), 140–150. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi:10.1109/TVCG.2021.3114876. 1, 2

[MM18] MAÇÃS C., MACHADO P.: Radial calendar of consumption. In *2018 22nd International Conference Information Visualisation (IV)* (2018), pp. 96–102. ISSN: 2375-0138. doi:10.1109/iV.2018.00027. 2

[NHG19] NUSRAT S., HARBIG T., GEHLENBORG N.: Tasks, techniques, and tools for genomic data visualization. 781–805. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.13727. doi:10.1111/cgf.13727. 1, 2, 4, 9

[OJEF19] ONDOV B., JARDINE N., ELMQVIST N., FRANCONERI S.: Face to face: Evaluating visual comparison. 861–871. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi:10.1109/TVCG.2018.2864884. 2

[PLW*22] PANDEY A., L'YI S., WANG Q., BORKIN M., GEHLENBORG N.: GenoREC: A recommendation system for interactive genomics data visualization. doi:10.31219/osf.io/rscb4. 2

[RFT*13] RUDDLE R. A., FATEEN W., TREANOR D., SONDERGELD P., QUIRKE P.: Leveraging wall-sized high-resolution displays for comparative genomics analyses of copy number variation. IEEE, pp. 89–96. doi:10.1109/BioVis.2013.6664351. 2

[RZZ*13] REBER S., ZHAO Y., ZHANG L., ORLOFF M., ENG C.: Visual analysis of tracts of homozygosity in human genome. 5 pages.

Artwork Size: 5 pages ISBN: 9783905674552 Publisher: The Eurographics Association. `doi:10.2312/PE.EUROVAST.EUROVA13.037-041`. 2

[Sti17] STILLWELL J.: *The Routledge Handbook of Census Resources, Methods and Applications: Unlocking the UK 2011 Census*. Routledge, 2017. Google-Books-ID: 5GcPEAAAQBAJ. 2

[vdBJvW*23] VAN DEN BRANDT A., JONKHEER E. M., VAN WORKUM D.-J. M., VAN DE WETERING H., SMIT S., VILANOVA A.: Panva: Pangenomic variant analysis. *IEEE Transactions on Visualization and Computer Graphics* (2023), 1–15. `doi:10.1109/TVCG.2023.3282364`. 2

[WDG*20] WALDNER M., DIEHL A., GRAČANIN D., SPLECHTNA R., DELRIEUX C., MATKOVIĆ K.: A comparison of radial and linear charts for visualizing daily patterns. 1033–1042. Conference Name: IEEE Transactions on Visualization and Computer Graphics. `doi:10.1109/TVCG.2019.2934784`. 2

[ZFAQ13] ZENG W., FU C.-W., ARISONA S. M., QU H.: Visualizing interchange patterns in massive movement data. 271–280. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.12114. `doi:10.1111/cgf.12114`. 2