**Mutation Method Selection**

\*\*\*

( Methods and materials)

To find the balance between exploration and exploitation 4 methods were checked:

1. Exploitation first method - Select random neighbor from the first neighbors for the first 100 generations and select random neighbor from second neighbors after 100 generations
2. Exploration first method - Select random neighbor from the second neighbors for the first 100 generations and select random neighbor first neighbors after 100 generations
3. Combined method - Select random neighbor from the second neighbors for the first 100 generations, then select random neighbor first neighbors for more 150 generations, and after 250 generations select random neighbor from the second neighbors again.
4. Random Method – Select a random configuration from the non-simulated configurations

Those methods were tested in 4 different stop conditions:

1. Regular - global 1240 generations.
2. Aggressive - global 1240 generations and local if the concept for 30 generations in a row.
3. Medium - global 1240 generations and local if the concept for 50 generations in a row.
4. Ease - global 1240 generations and local if the concept for 100 generations in a row.

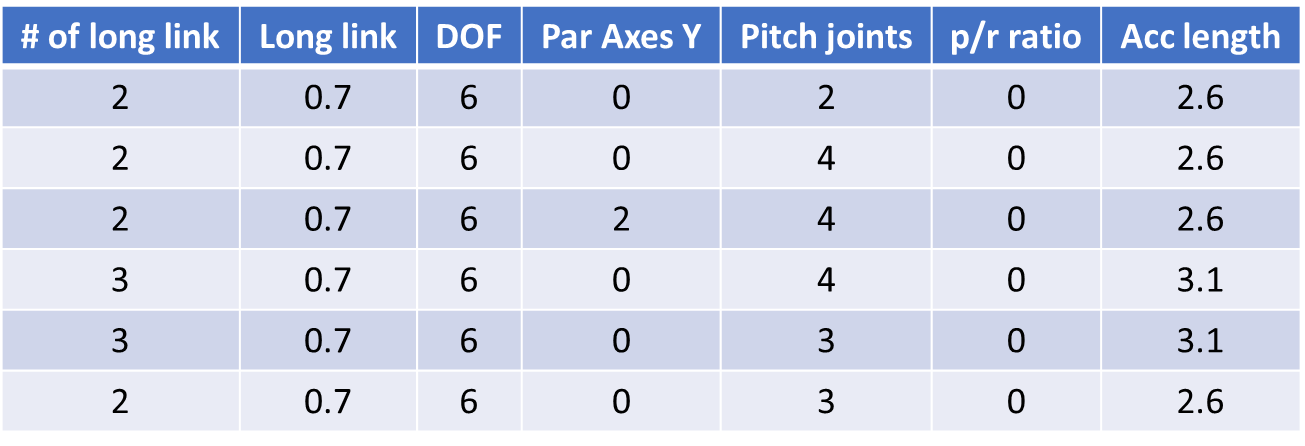
(Results)

**Exploration - Exploitation first method**

**Exploitation - Exploration first method**

To find the best mutation method 6 large concepts were chosen randomly (can be seen in Table 7).

Table 1- Checked Concepts



The selected concepts were fully sorted, all the configurations in those concepts were simulated and their results are known. To compare the methods, each method runs 30 times, for a maximum of 1240 generations. For each method, two indices were calculated: Hyper–Volume(HV) which bigger value means better result and minimum value of the manipulability, this index selected because it can be shown in figure 1 that big part of the results of the mid proximity joint is close to 0 and the manipulability is spreading more. Another metric used to compare the Inverse Generational Distance (IGD). IGD computes the average distance between the true set and the current set. The lower the IGD value, the better is the front.

In figure 2 it can be seen HV at each generation for each method.

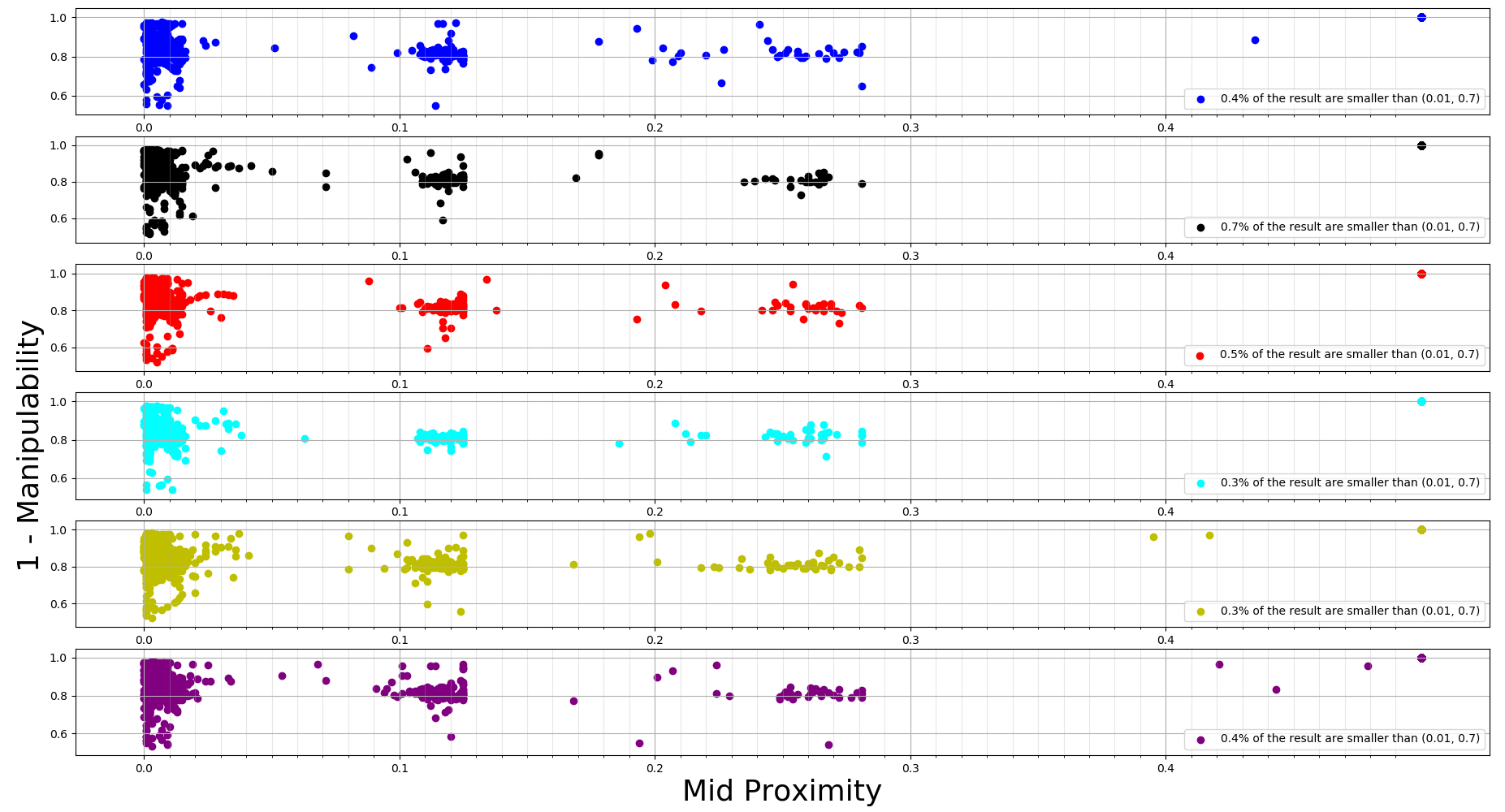


Figure - Concepts Full sorting

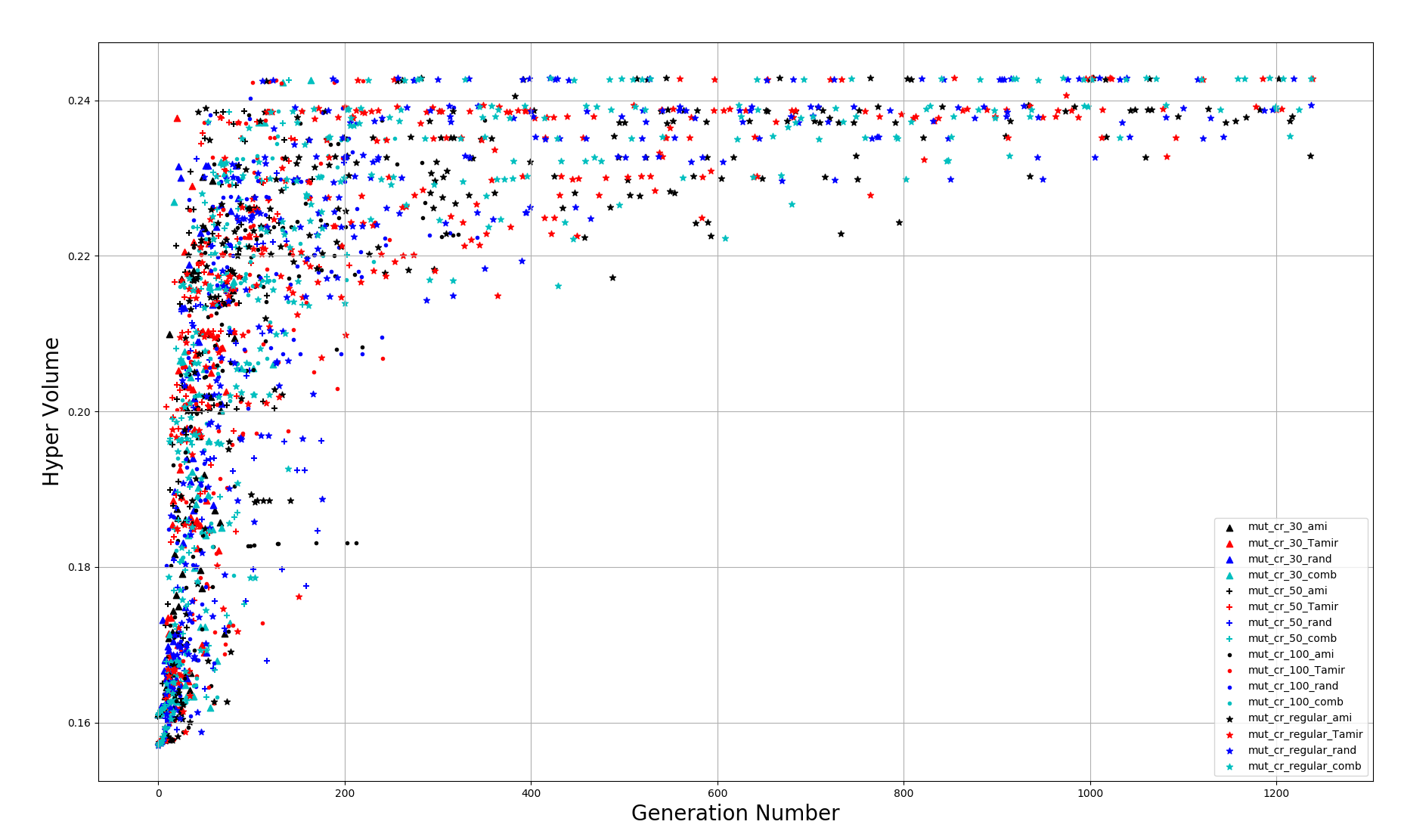


Figure - HyperVolume Vs Generation

In figure 3 it can be seen minimum manipulability at each generation for each method.

From those figures, it can be seen that the most significant changes have occurred in the first 300 generations. Another thing that can be seen that except for the regular stop condition, all the stop conditions get their final position until generation number 200.

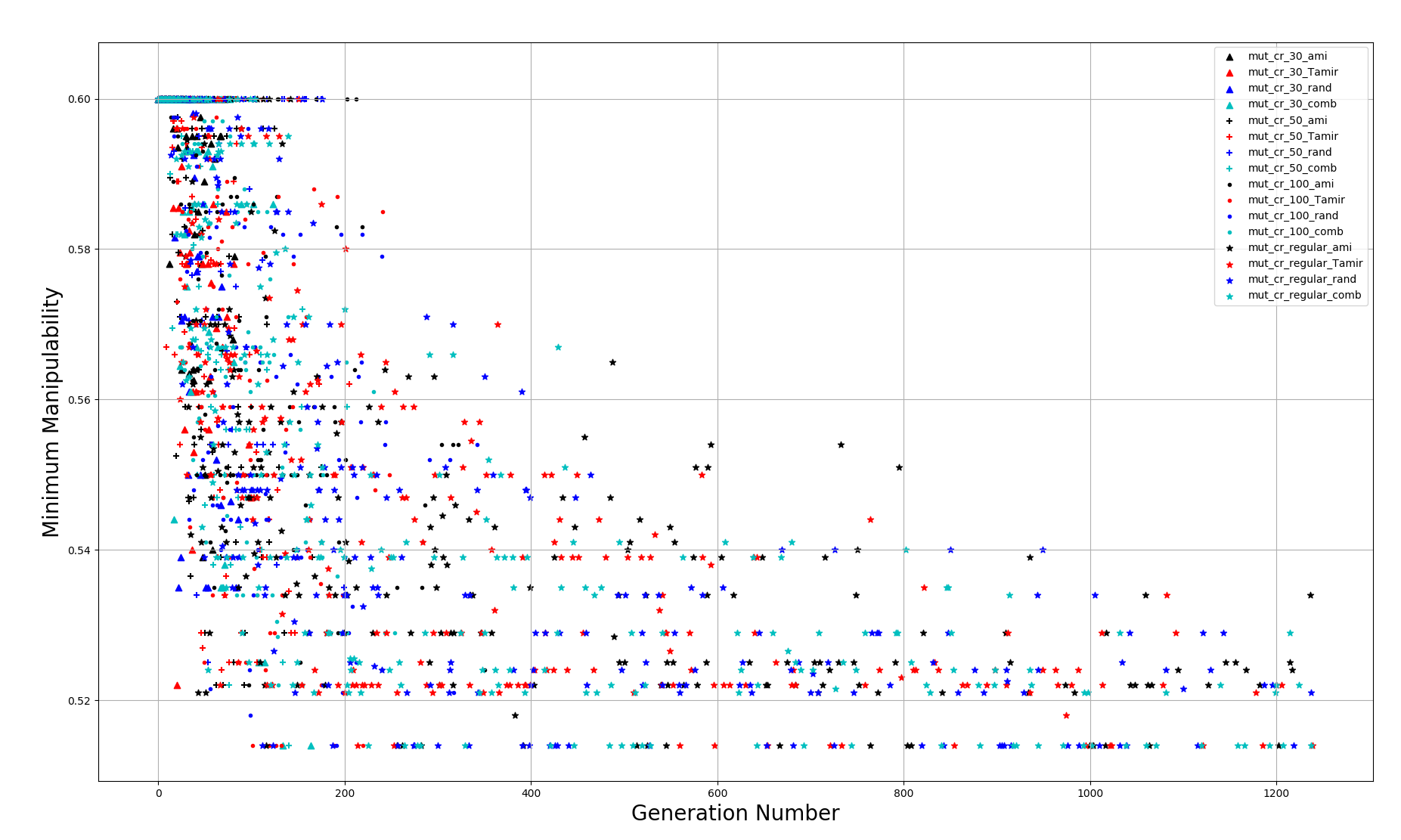


Figure - Minimum Manipulability Vs Generation

In figure 4 it can be seen the IGD at each generation for each method. From this figure, it's hard to get a lot of relevant information about the behavior of the methods.

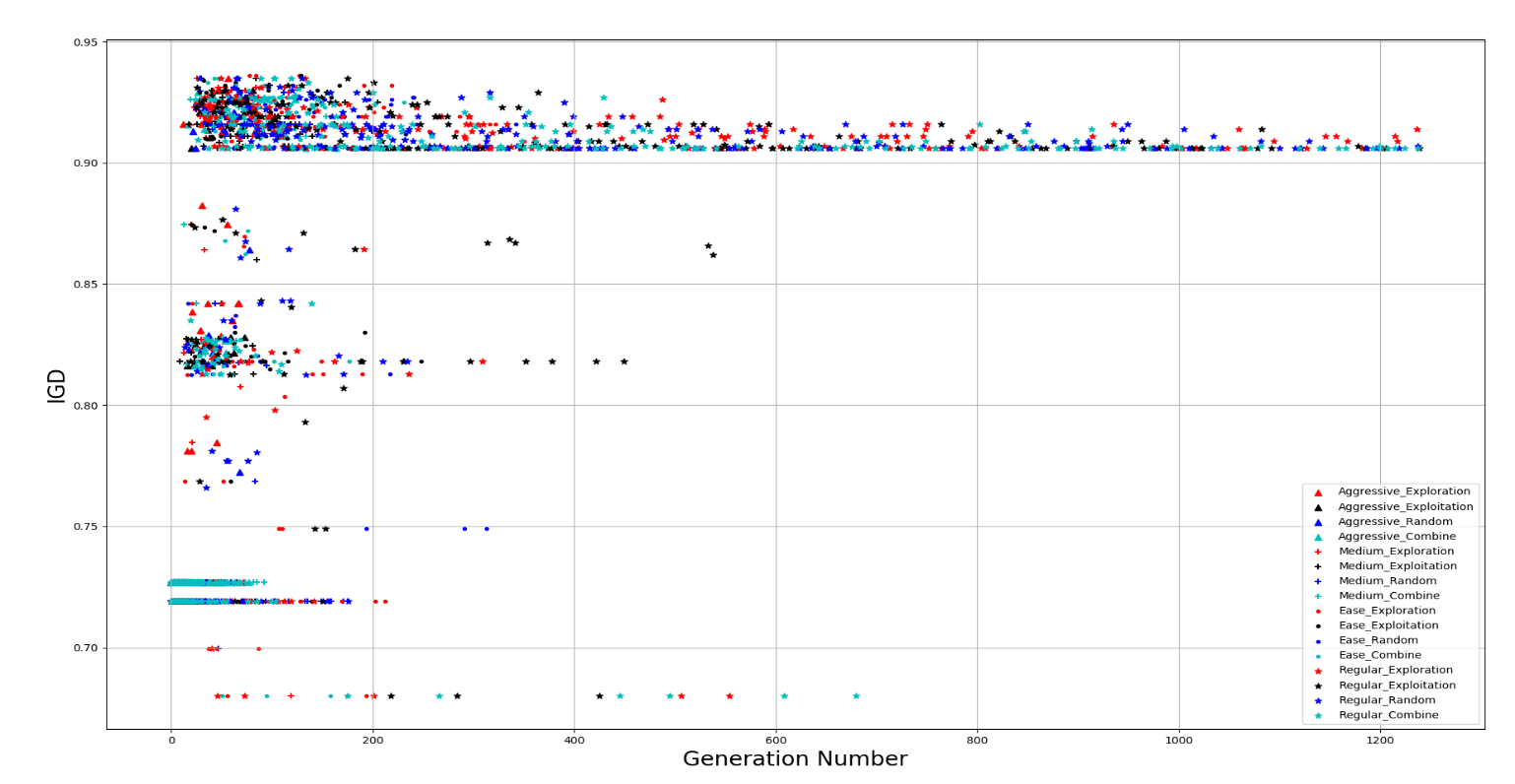


Figure - IGD Vs Generation

To compare the algorithms Wilcoxon test was applied over the results of each method of the 30 runs. Wilcoxon test return P-Value, which P-Value lower than 0.05 means there is a statistical significance that one method its better than another. from figures 5-7 it can be seen that methods with the same stop conditions are in the same population. The selected methods are with the highest median and the lowest Min-Manipulability median and IGD (Table 2).

In Table 2 there are the medians and variances of HV, IGD, Min Manipulabiliry, and the generation in which the WOI was last change over 30 runs on each method. The last four columns, under the Wilcoxon title, show if the method its less than (-) the reference method or there is uncertainty(~) according to the Wilcoxon test.

It can be seen that the regular stop condition (which simulate the Fair resource allocation method), gives the best results, but between the mutation methods, there is no statistical significance for any method. But because of this test compare only six fully sort concepts, the disadvantages of the fair method don’t take into account, such as the simulation time, and the number of concepts to optimize which will make the algorithm to stop early, after about 150-200 generations.

From Table 3, it can be seen that the 3 best methods that not in the regular stop condition are:

Ease-Random, Aggressive-Exploitation, and Medium-Exploitation.

From figures 5 -7, it can be seen that Aggressive-Exploitation is statistical significance as not good as Ease-Random in HV and Min-Manipulability metrics.

In Table 4 it can be seen a comparison of the HV at different generations. The selected numbers were taken from the median values in Table 2.

From Table 4 the method that was selected is **Medium-Exploitation** because it's one of the three methods that were selected as reference methods, but any of the other methods are statistical significance better than it.

Medium-Exploitation is a greedy method, and therefore **Regular-Exploitation** also will be tested to perform a fair comparison between the data allocation methods. Another method that will be tested is **Regular-Random** to compare the methods with random selection.

Table 2 - Summary of the results



Table 3 - Summary of the results without Regular

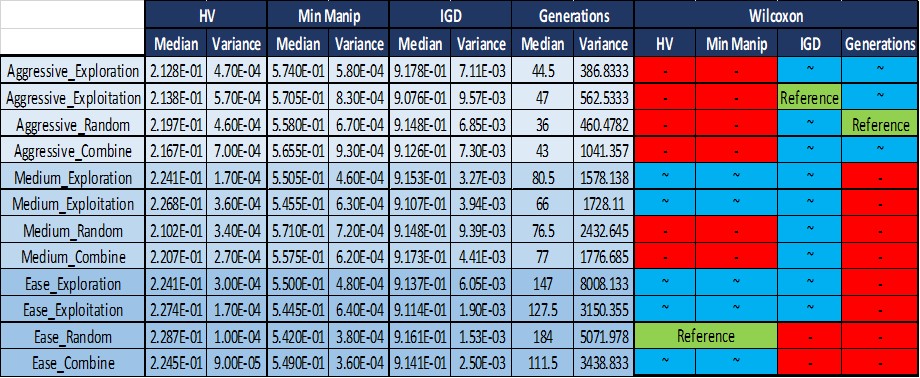
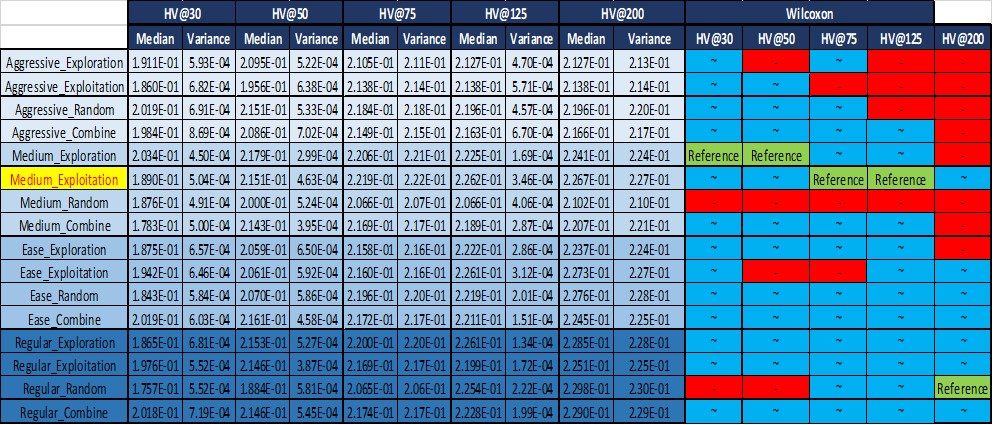


Table 4 - HV in specific Generation



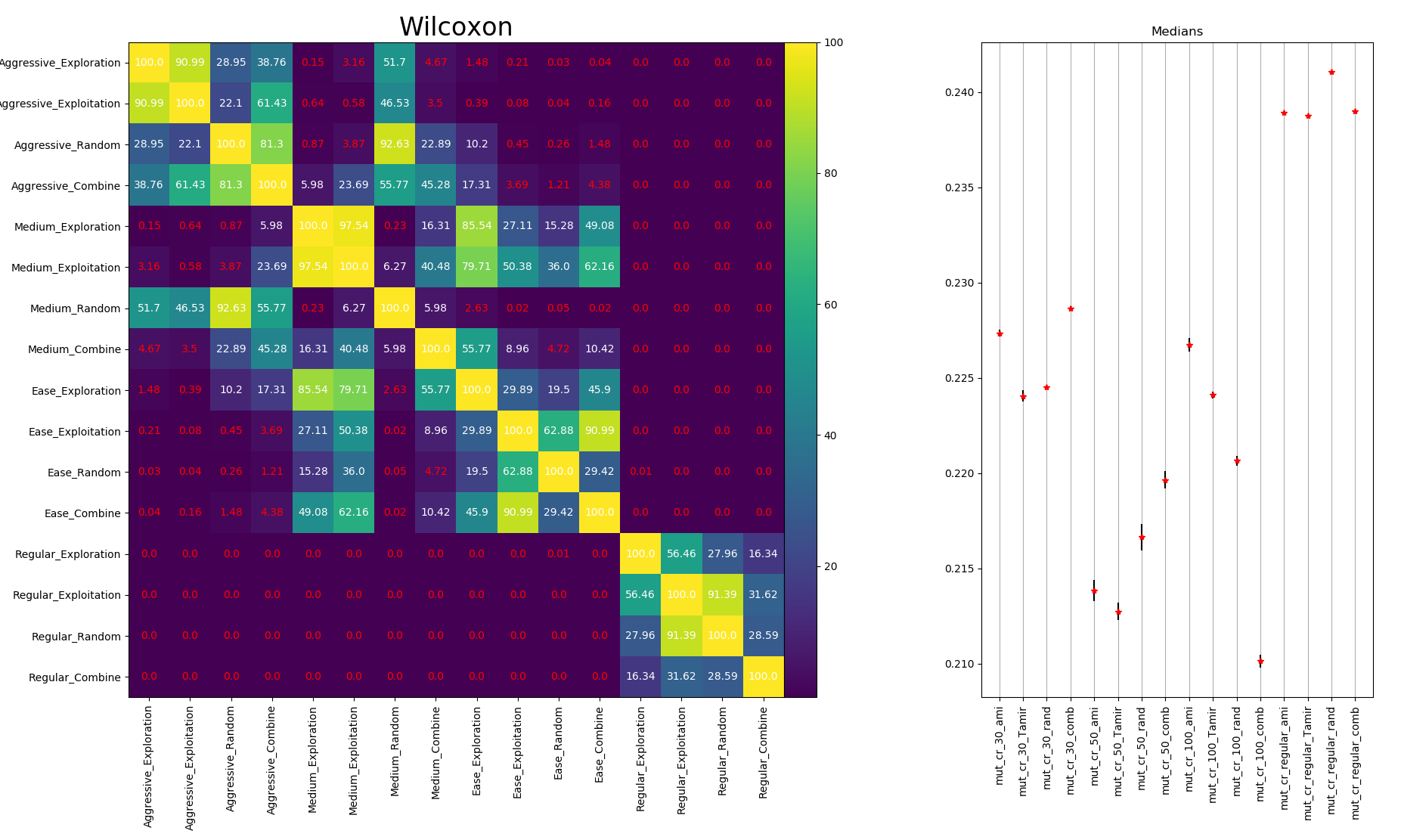


Figure 5- Wilcoxon: HV

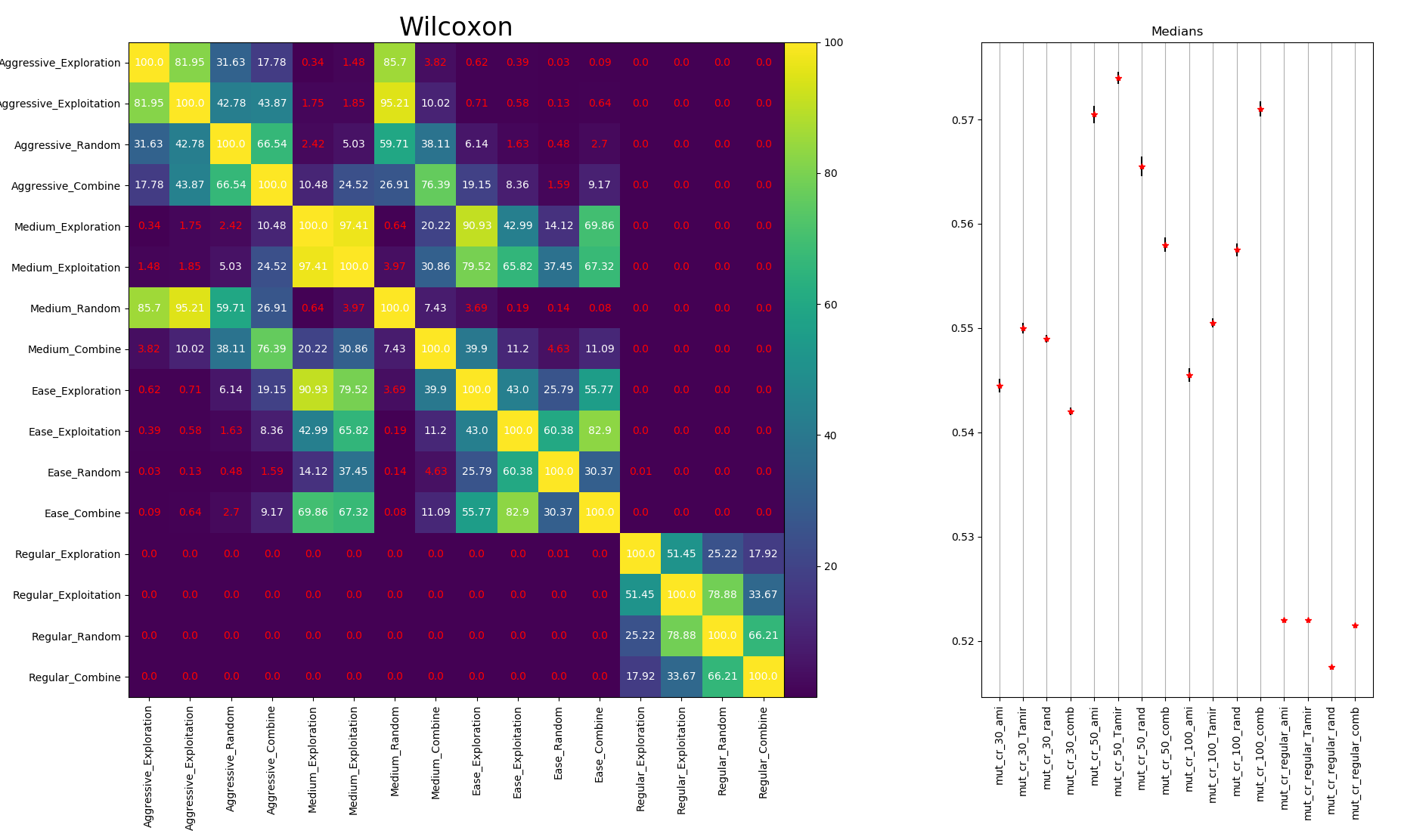


Figure 6 - Wilcoxon: Min Manipulability

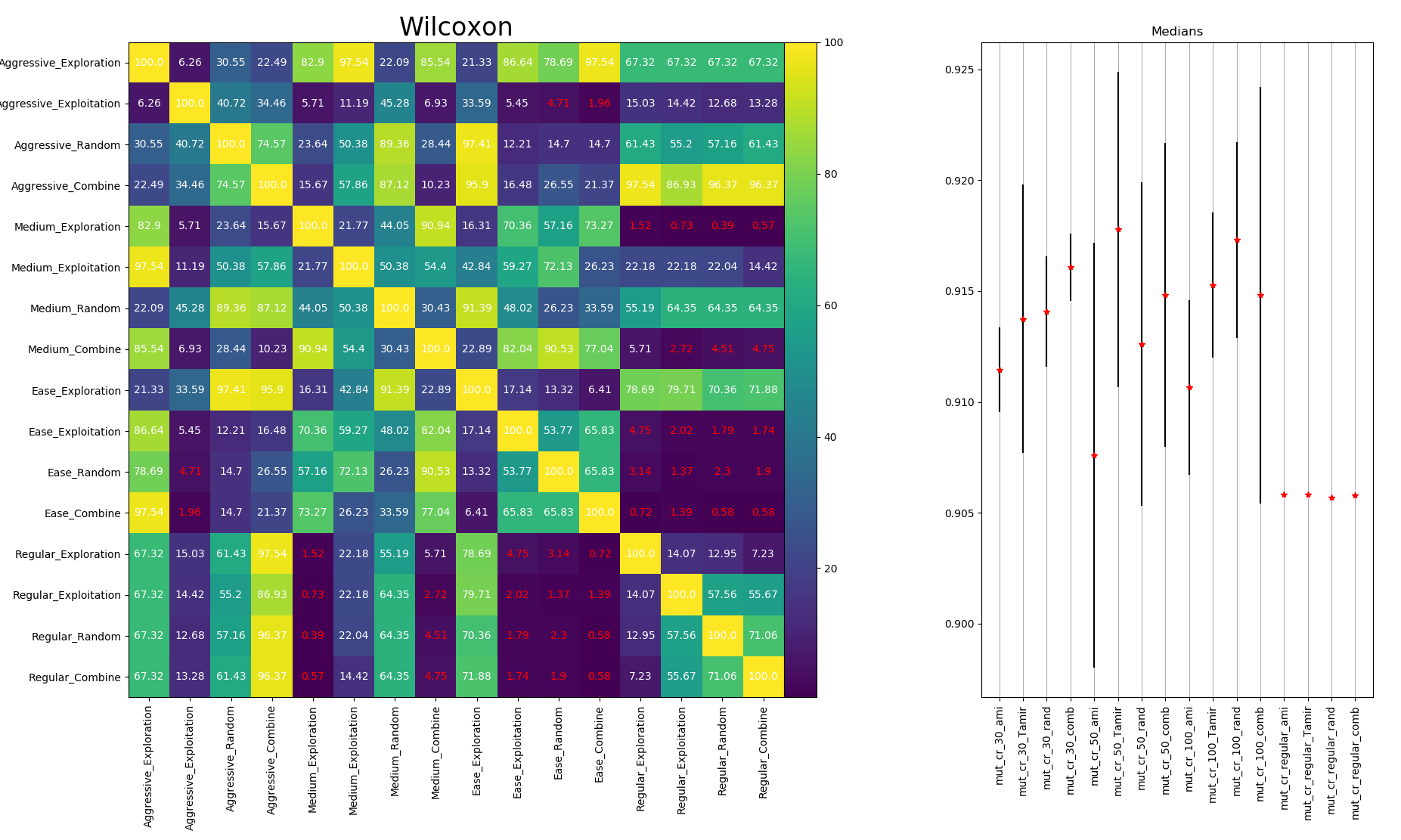


Figure 7 - Wilcoxon IGD