**Genomic Validation – Run level, samples as statistical units (n=63)**

**Overall and stratified descriptive statistics**

In this step we present descriptive statistics for each section. Results are provided by median and interquartile range both overall and stratified on the guideline.

**SECTION: *DNA CONTAMINATION SCORES AND RELATED***

**Table 1. Descriptive statistics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **n=63 samples** |  | **All samples** | **Guideline =0** | **Guideline =1** |
|  | **Valid values°**  **(DNA samples)** | **n or median**  **(min; max)** | **n or median**  **(min; max)** | **n or median**  **(min; max)** |
| Contamination (0/1) | 40 | 1/39 | 1 | 39 |
| Contamination score (u) | 40 | 1305 (72; 7416) | 5162 | 1270 (72; 7416) |
| P value contamination (u) | 40 | 0.896 (0; 1) | 0.227 | 0.996 (0; 1) |

**Note.** n: absolute frequencies; °non-missing values.

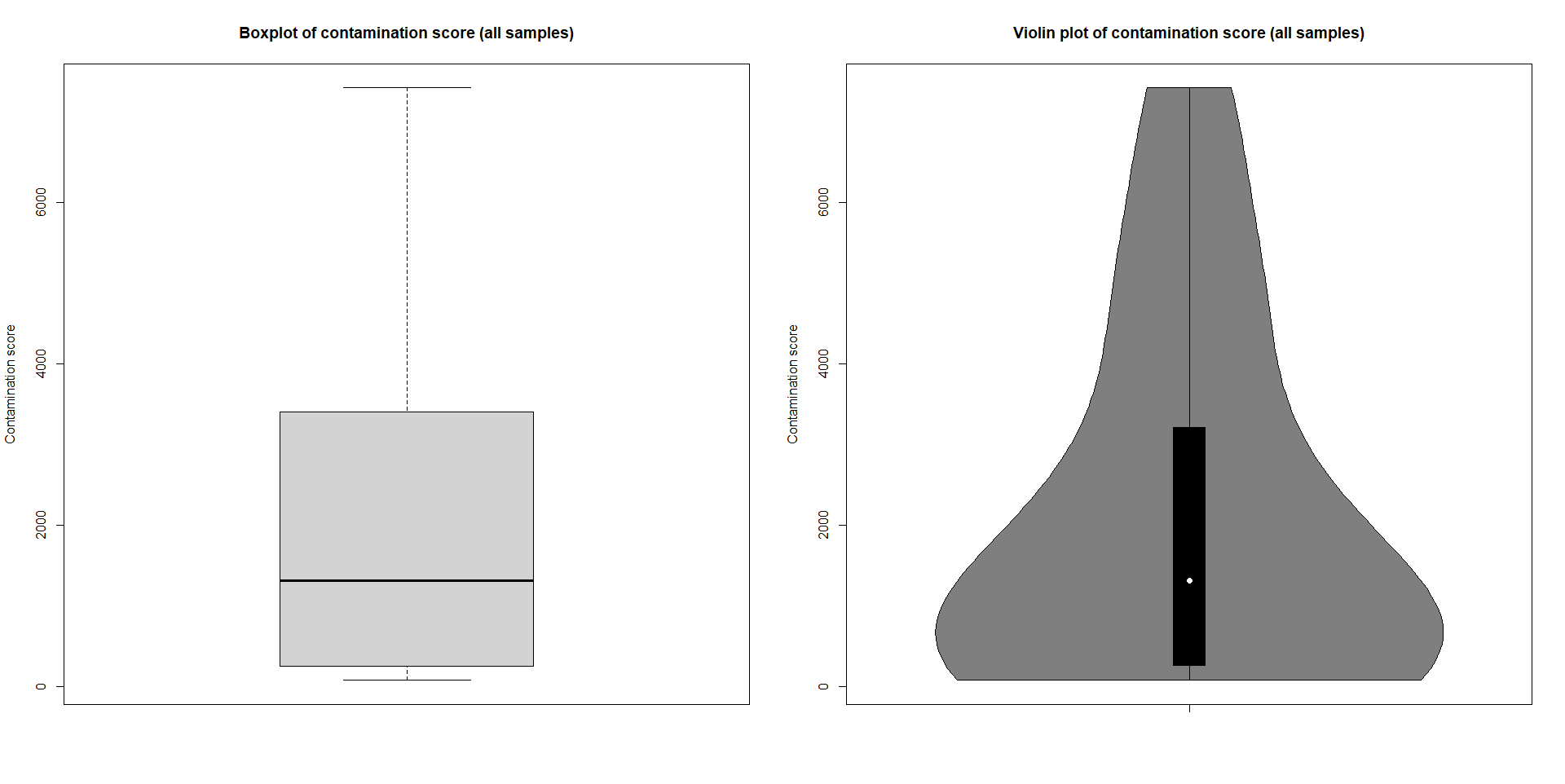
**Table 2. Descriptive statistics of potential CNV**

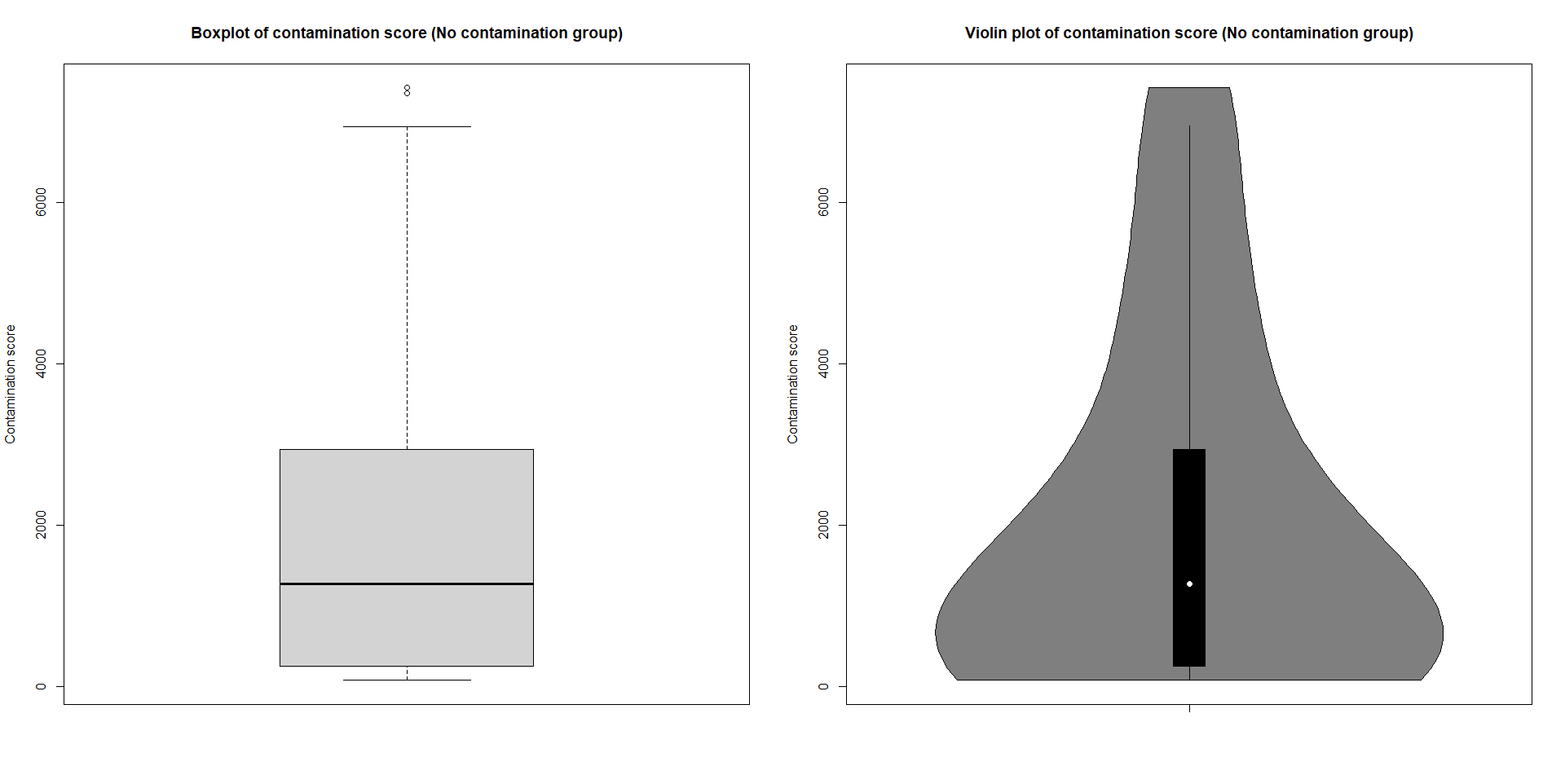
|  |  |
| --- | --- |
| Potential CNV\* | 9 |
| Contamination score (u) | 5040 (3766; 7416) |
| P value contamination (u) | 0.006 (0; 0.044) |
| coverage\_mad\_count |  |
| median\_bin\_count\_cnv\_target\_count |  |

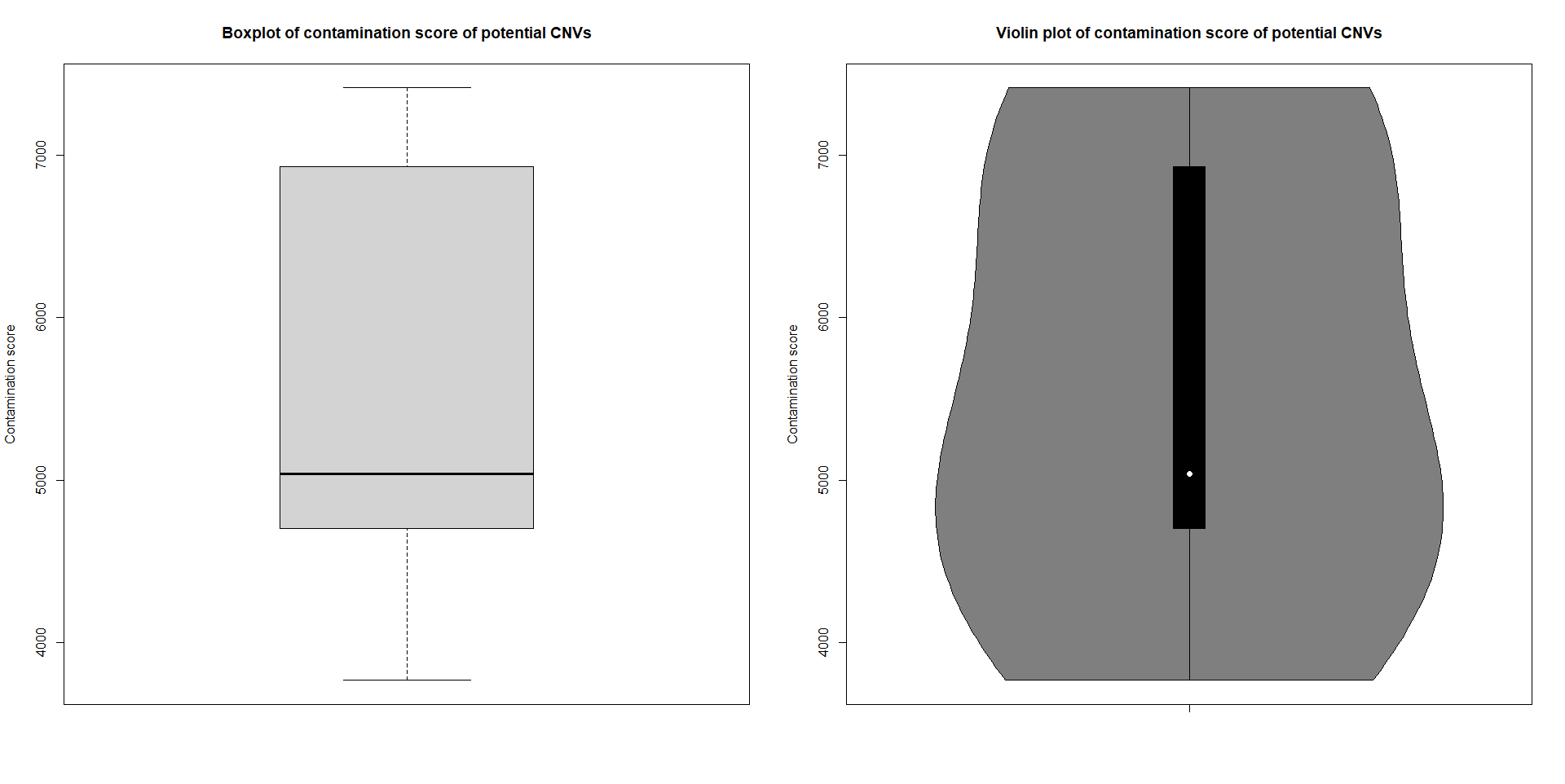
**Note.** CNV: Copy Number Variation; u: unit

\*Potential CNV: contamination= 1 AND contamination score > 3106

**Figure 1. Boxplot and violin plot of the DNA CONTAMINATION SCORES**







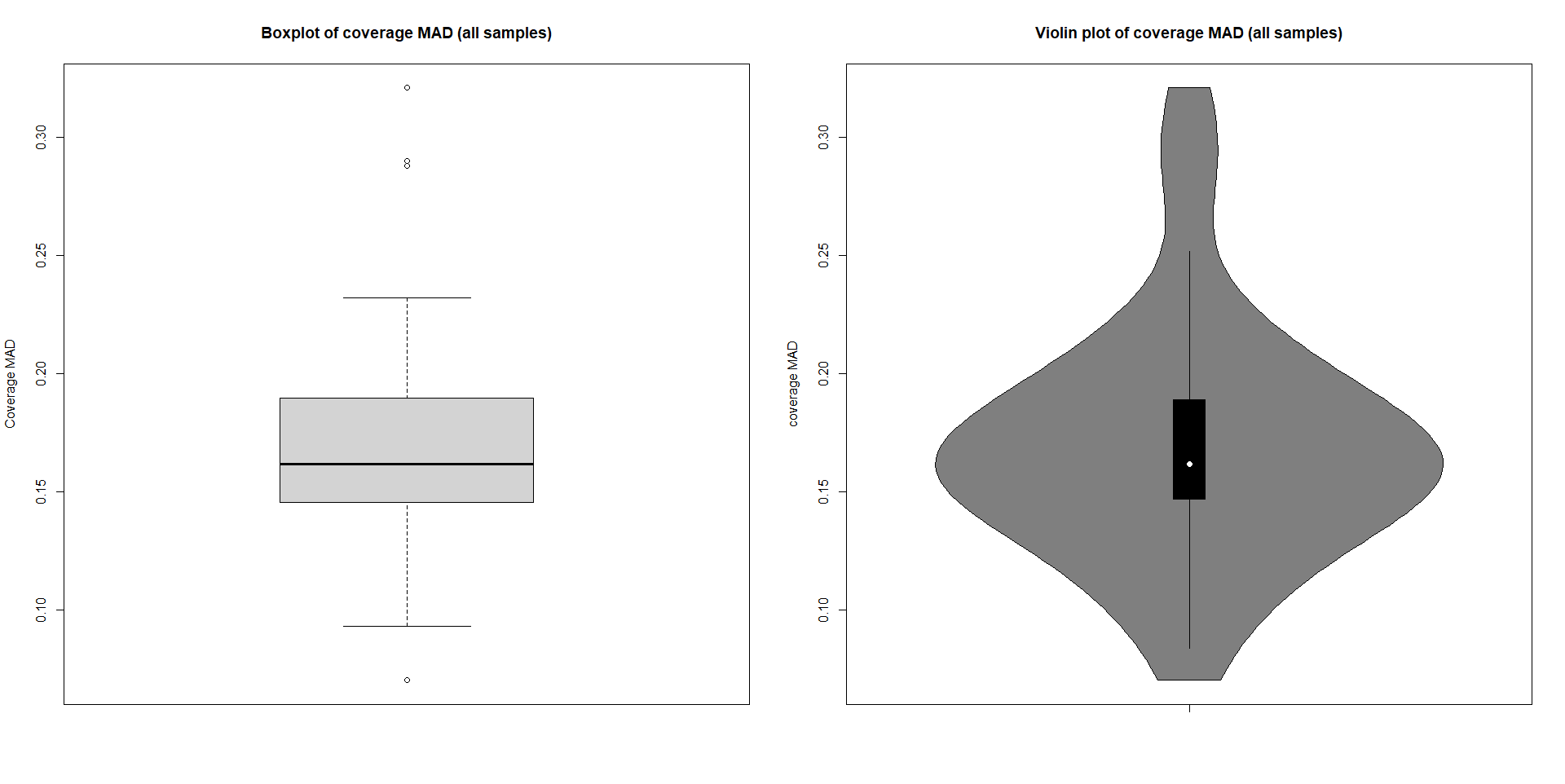
**SECTION: *COPY NUMBER VARIATION (CNV)***

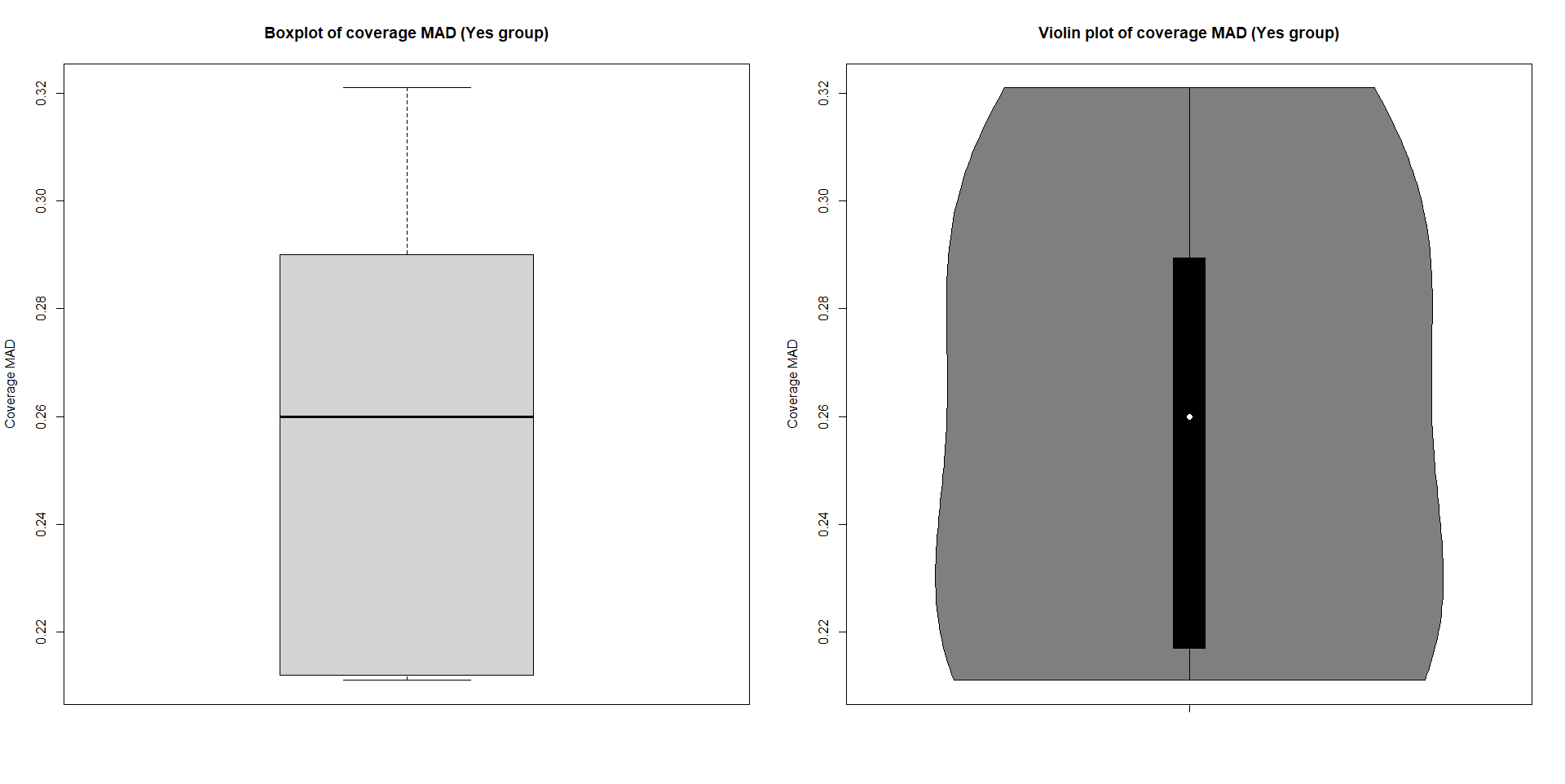
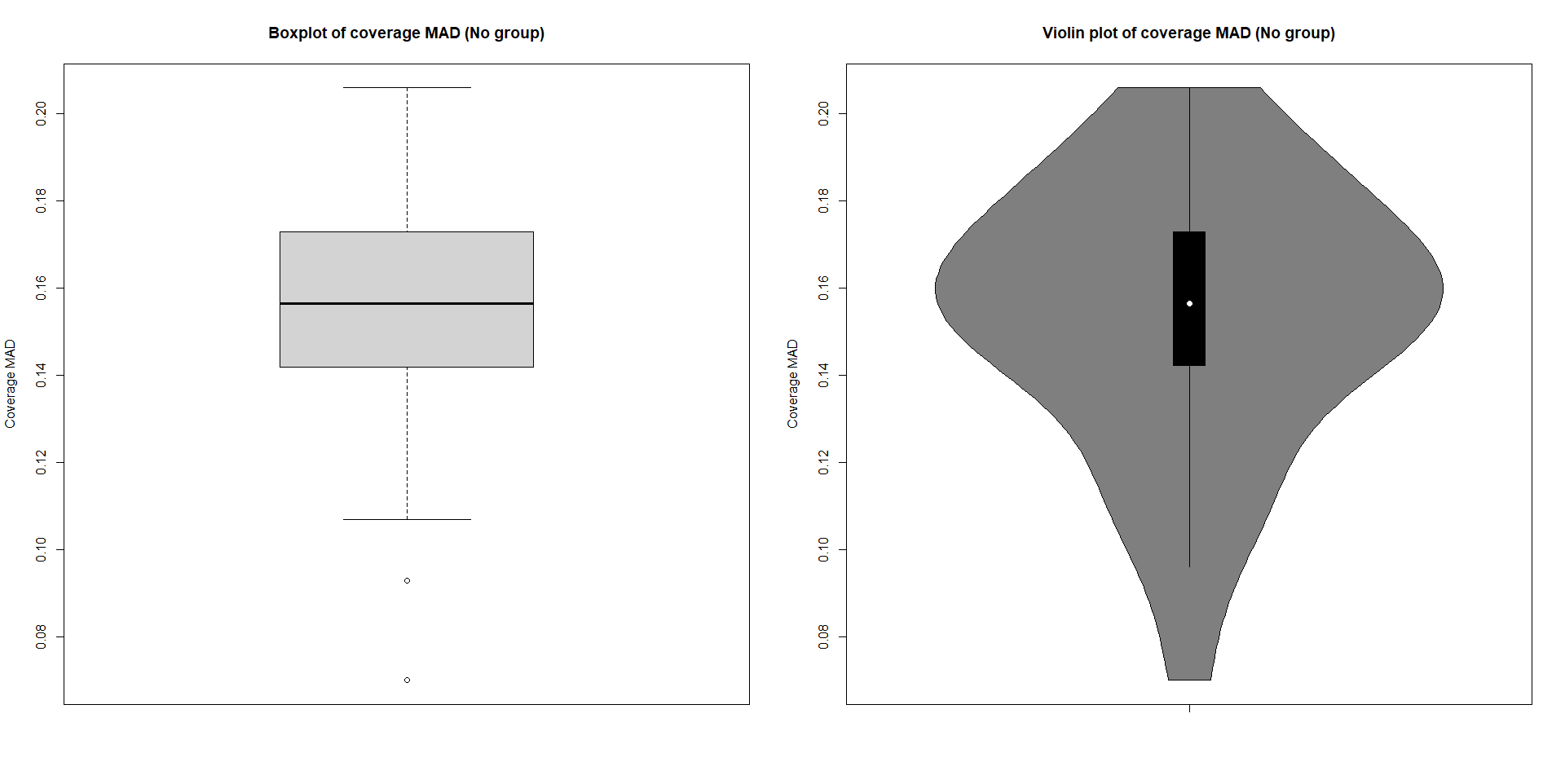
**Table 3. Descriptive statistics**

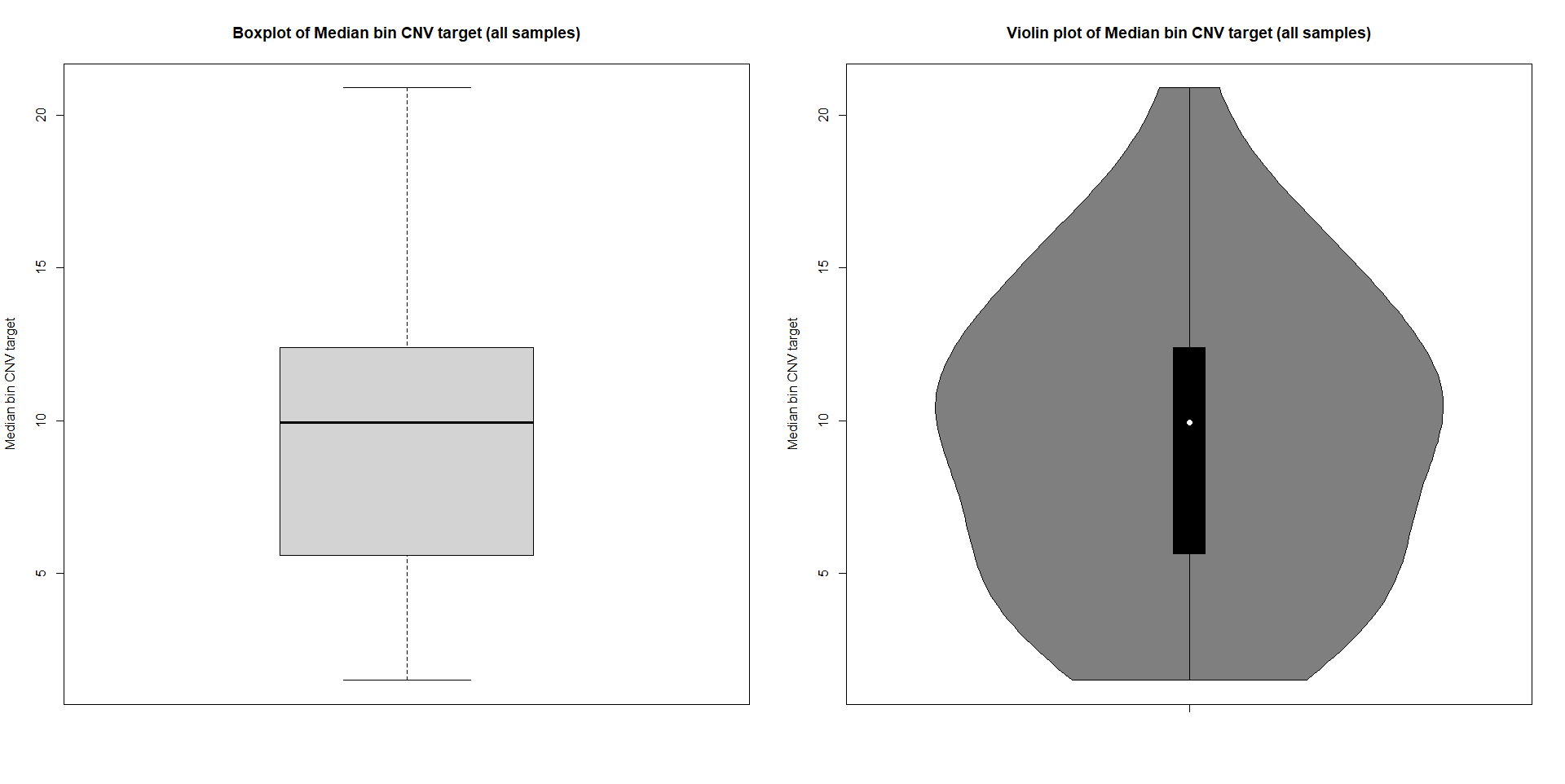
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **n=63 samples** |  | **All samples** | **Guideline =0** | **Guideline =1** | **P-value\*** |
|  | **Valid values°**  **(DNA samples)** | **n or median**  **(min; max)** | **n or median**  **(min; max)** | **n or median**  **(min; max)** |  |
| Coverage MAD (0/1) | 40 | 6/34 | 6 | 34 |  |
| Coverage MAD (u) | 40 | 0.162 (0.070; 0.321) | 0.260 (0.211; 0.321) | 0.156 (0.070; 0.206) | <0.001 |
| Median bin CNV target (u) | 40 | 9.95 (1.5; 20.9) | 9.1 (3.2; 13.6) | 10.05 (1.5; 20.9) | 0.719 |

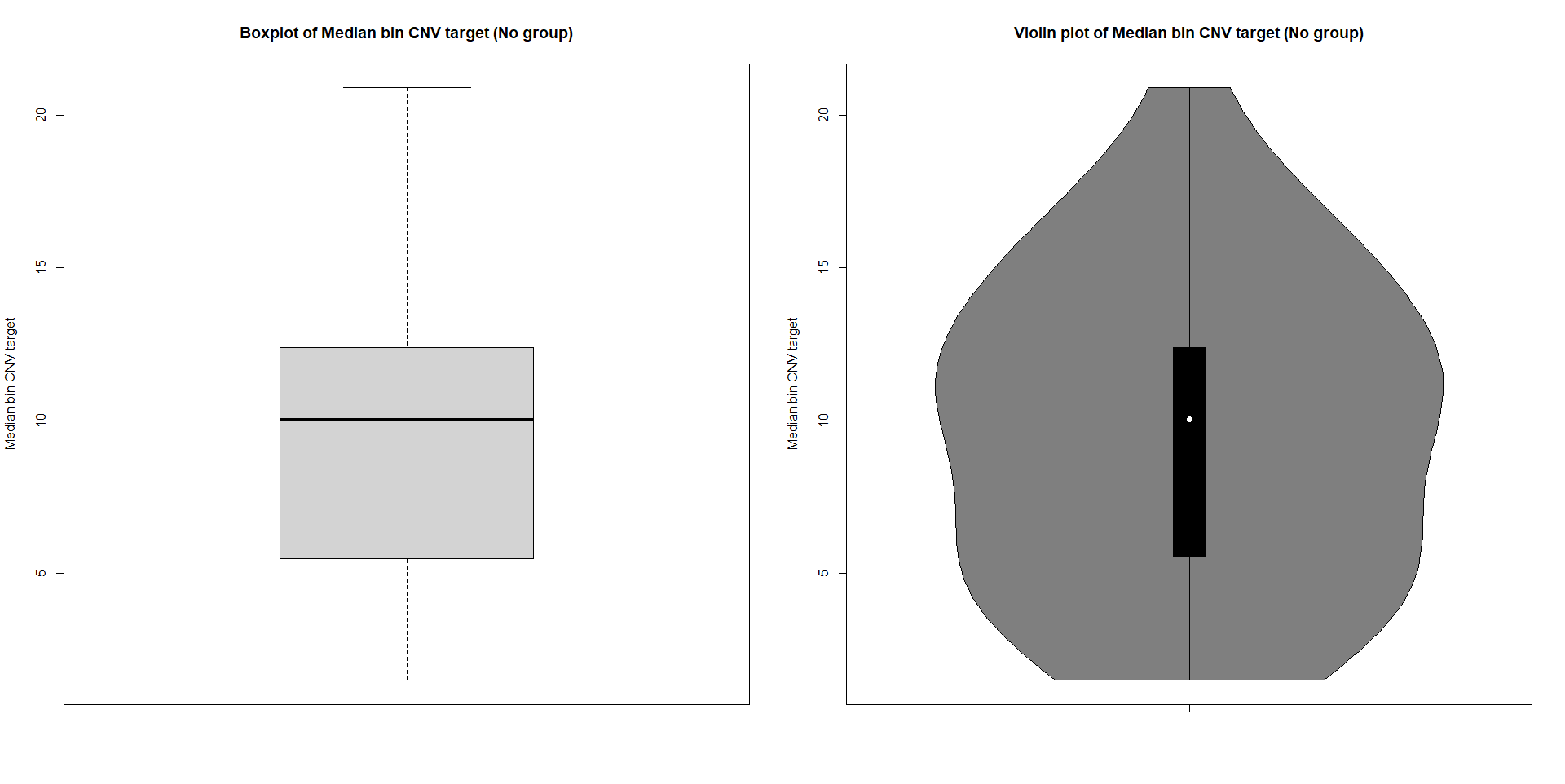
**Note.** n: absolute frequencies; MAD: Median Absolute Deviation; u: unit; °non-missing values; \*Mann-Whitney test

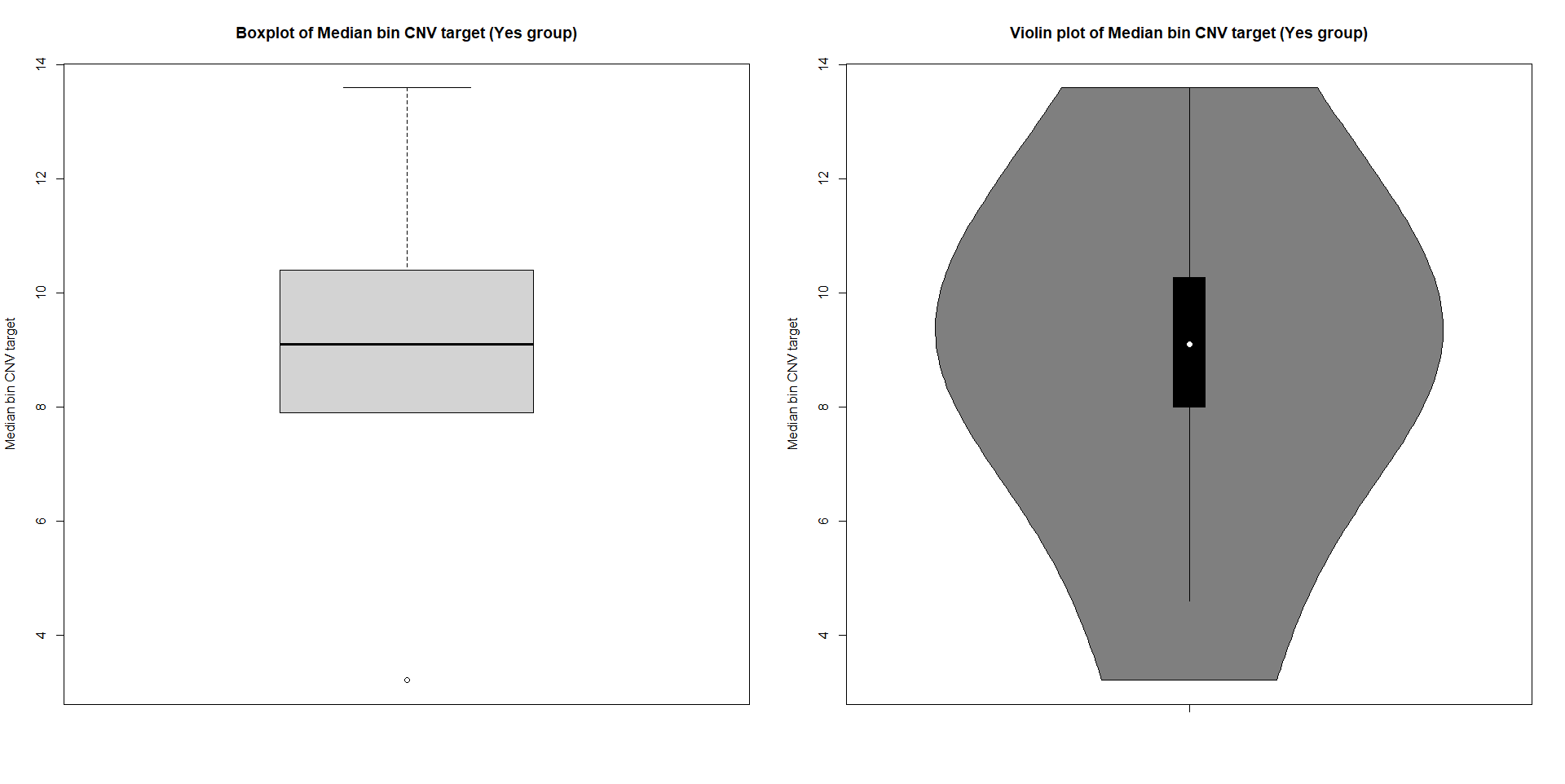
**Figure 2. Boxplot and violin plot of the COPY NUMBER VARIATION**











**SECTION: *DNA METRICS***

**Table 4. Descriptive statistics**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **n=63 samples** |  | **All samples** | **Guideline=0** | **Guideline =1** | **P-value\*** |
|  | **Valid values°**  **(DNA samples)** | **n or median**  **(min; max)** | **n or median**  **(min; max)** | **n or median**  **(min; max)** |  |
| *Tumour Mutational Burden* |  |  |  |  |  |
| Median insert size (0/1) | 40 | 0/40 | NA | 40 | NA |
| Median insert size (u) | 40 | 104.5 (73; 149) | NA | 104.5 (73; 149) | NA |
|  |  |  |  |  |  |
| Median exon coverage (0/1) | 40 | 3/37 | 3 | 37 |  |
| Median exon coverage (u) | 40 | 709 (99; 1646) | 112 (99; 119) | 743 (202; 1646) |  |
|  |  |  |  |  |  |
| Exon 50x (0/1) | 40 | 1/39 | 1 | 39 |  |
| Exon 50x (%) | 40 | 99.3 (87; 99.7) | 87 | 99.3 (90.9; 99.7) |  |
|  |  |  |  |  |  |
| *Microsatellite Instability (MSI)* |  |  |  |  |  |
| Unstable (MSI) sites (0/1) | 40 | 7/33 | 7 | 33 |  |
| Unstable (MSI) sites (unit) | 40 | 117.5 (2; 125) |  |  |  |
| mean\_target\_coverage (u) |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

**Note.** n: absolute frequencies; u: unit; °non-missing values; NA: not available; \*Mann-Whitney test.

**SECTION: DNA Expanded Metrics**

|  |  |  |
| --- | --- | --- |
| **n=63 samples** |  | **All samples** |
|  | **Valid values°**  **(DNA samples)** | **median**  **(min; max)** |
|  |  |  |
| Total pf reads (u) | 40 | 181M (66M; 253M) |
| Median target coverage (u) | 40 | 705 (96; 1595) |
| Chimeric reads (%) | 40 | 0.825 (0.050; 4.87) |
| Exon 100x (%) | 40 | 98.7 (49.6; 99.5) |
| Read enrichment (%) | 40 | 71.9 (60.3; 82.8) |
| Usable\_umi\_reads (%) | 40 | 99.9 (99.8; 99.9) |
| Mean target coverage (u) | 40 | 721.3 (110.4; 1619.0) |
| Aligned reads (%) | 40 | 96.55 (55.4; 99.1) |
| Contamination est (%) | 40 | 2 (0.2; 25.7) |
| Pf\_uq\_reads (%) | 40 | 100 (100; 100) |
| Target\_04x\_mean (%) | 40 | 94.95 (64.3; 96.8) |
| Target\_100x (%) | 40 | 98.1 (47.1; 99.4) |
| Target\_250x (%) | 40 | 94.8 (0.4; 98.4) |

**Note.** n: absolute frequencies; u: unit; °non-missing values

**SECTION: Association analysis**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **n=63 samples** |  | **MSI Guideline=0** | **MSI Guideline =1** | **P-value\*** |
|  | **Valid values°**  **(DNA samples)** | **n or median**  **(min; max)** | **n or median**  **(min; max)** |  |
| *Microsatellite Instability (MSI)* |  |  |  |  |
| Unstable (MSI) sites (0/1) | 40 | 7 | 33 |  |
| Total pf reads (u) |  |  |  |  |
| Median target coverage (u) |  |  |  |  |
| Chimeric reads (%) |  |  |  |  |
| Exon 100x (%) |  |  |  |  |
| Read enrichment (%) |  |  |  |  |
| Usable umi reads (%) |  |  |  |  |
| Mean target coverage (u) |  |  |  |  |
| Aligned reads (%) |  |  |  |  |
| Contamination est (%) |  |  |  |  |
| Pf uq reads est (%) |  |  |  |  |
| Target 04x mean (%) |  |  |  |  |
| Target 100x (%) |  |  |  |  |
| Target\_250x (%) |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**Note.** n: absolute frequencies; MSI: Microsatellite Instability; u: unit; °non-missing values; \*Mann-Whitney test.