INTRODUCTION

Signal transducer and activator of transcription protein 3 (STAT3) (OMIM accession number: *102582) is encoded by one of seven STAT family genes located in chromosomal band 17q21.2 and extends over 75kb (Aggarwal et al., 2009). Its expression is induced by cytokines, hormones, and growth factors. STAT3 protein is activated by phosphorylation of its tyrosine and serine residues via signaling from upstream regulators (Klemm et al., 1998). STAT3 is reported to regulate the expression of many genes such as Bcl-xL, cyclin D1, c-myc, VEGF, IL-10, IL-2, subsequently leading...
to cellular proliferation and slowing-down of apoptosis (Herrmann et al., 2010; Rane & Reddy, 2000). Moreover, experimental models revealed important immune functions for STAT3, including innate and adaptative immunity (Hillmer et al., 2016).

In addition to its physiological role, studies have revealed the role of STAT3 in diseases. Indeed, some mutations in the STAT3 gene are associated with human immune disorders (Gutiérrez et al., 2018; Milner et al., 2015; Velayos et al., 2017). STAT3 is also implicated in tumorigenesis by enhancing tumor growth, survival, invasion, immune suppression, and angiogenesis on the one hand and, decreasing tumor cell apoptosis on the other hand. Moreover, the Jak-STAT3 signaling pathway has been shown to have central roles in inflammation-mediated cancer, including cancer stem cells (CSCs) (Schroeder et al., 2014) and pre-metastatic niche formation (Deng et al., 2012).

Indeed, many proteins whose expression is driven by over-expression of unphosphorylated STAT3 have been implicated in many cancers (Yang et al., 2005). Activated STAT3 has been implicated in multiple human cancers including lung (Du et al., 2012), gastric (Wu et al., 2012), ovarian, breast (Hsieh et al., 2005) (Sansone et al., 2007), colon (Calon et al., 2012) (Liang et al., 2013), prostate (Kroon et al., 2013), hepatocellular carcinoma (Hatziapostolou et al., 2011), and lymphoma (Liu et al., 2012).

A fundamental role for STAT3 in the normal development of the mammary gland and the pathogenesis of human breast cancer (BC) has been established (Clevenger, 2004; Watson, 2001). Several studies have demonstrated increased levels of STAT3 in primary mammary tumors. Immunohistochemical approaches in humans have found increased levels of nuclear-localized STAT3 in malignant BCs when compared with normal tissues (Watson, 2001). A recent study has identified that STAT3 expression was found to have a significantly higher correlation with luminal breast cancer (Eroglu et al., 2020). Hence, according to these data and because of its implication in cancer development and progression, Signal transducer and activator of transcription protein 3 (STAT3), has been recognized as a type of oncogene (Bromberg et al., 1999).

Despite the identification of around fifteen single nucleotide polymorphisms (SNPs) in the STAT3 gene, only a few studies have investigated the association of SNPs in this gene with the susceptibility to cancer. For example, Vaclavicek et al. (2007) reported that the STAT5B rs6503691 and the STAT3 rs7211777 polymorphisms were associated with an increased risk for breast cancer in German patients with familial breast cancer (Vaclavicek et al., 2007). Wang et al. (2011) detected an association between STAT3 polymorphism rs4769793 and cervical cancer. Indeed, women with a G allele appeared to have a higher risk for cervical cancer. Further, the G allele was associated with poor tumor differentiation and positive parametrial invasion (Wang et al., 2011). Moreover, Jiang et al. (2011) have shown that a haplotype in the STAT3 gene may have a protective role in the development of non-small cell lung cancer (NSCLC) (Jiang et al., 2011). Hence, Zhao et al. (2015) proposed that STAT3 polymorphisms might be a candidate pharmacogenomic factor to assess susceptibility and prognosis of cancer (Zhao et al., 2015). Moreover, it has been shown that rs957971 polymorphism in STAT3 gene may predict an unfavorable response to first-line platinum-based therapy for women with advanced serous epithelial ovarian cancer in an American population sample of European ancestry (Permuth-Wey et al., 2016).

In the present study, we characterized the genetic variation of STAT3 gene in Tunisian and Libyan populations. The association status of STAT3 polymorphism and cancer in Tunisian populations is still unknown. Only two case/control studies on STAT3 polymorphism have been conducted in these populations. The first on rs744166, in Pemphigus patients, with no association was carried out (Ben Jmaa et al., 2018), and the second on rs1053023 and rs1053004 in relation to the Idiopathic Recurrent Miscarriage (IRM) showing that STAT3 rs1053023 was positively associated with IRM in Tunisian women (Messoudi et al., 2013).

In this paper, we focused on two single nucleotide polymorphisms rs7211777 (g.42382057G>A) and rs3869550 (g.42340869T>C), chosen for their association with cancer in German and Chinese populations (Jiang et al., 2011; Vaclavicek et al., 2007). In addition to these two SNPs, we also investigated rs957971 (g.42367907C>G), located in the STAT3 gene between these two SNPs; rs957971 has been associated with the response to chemotherapy (Permuth-Wey et al., 2016). Results will be discussed according to haplotypic diversity in this gene among North African populations.

2 | MATERIALS AND METHODS

2.1 | Ethical compliance

This work is approved by Ethics Committee for Research in Life Sciences and Health of the ISBM (CER-SVS/ISBM).

2.2 | DNA samples and STAT3 SNP typing

A total of 349 North African individuals were collected including 279 Tunisians from 6 populations well distributed throughout Tunisia: Kesra (n = 42) to the north, Sousse (n = 46), Mahdia (n = 45) and Kairouan (n = 40) to the center, Smar (n=62) to the south, and a population of Kerkennah island (n = 44), in addition to 70 Libyans (Figure 1). All
individuals sampled were unrelated and healthy persons and all individuals gave informed consent for the study of DNA sequence variants.

Total human genomic DNA was isolated from peripheral blood samples collected into EDTA tubes using the phenol-chloroform method.

The 3 SNPs (rs3869550, rs957971, rs7211777) of the STAT3 gene have been typed in 3 μl reactions using TaqMan® Assay-on-demand following the manufacturer's protocol. Assays were obtained from Applied Biosystems, Thermo Fisher AB TaqMan Catalog Numbers C___7530575_10 C___1952199_10 C___1952182_10, respectively. 384-well plates were read on an AB7900 thermocycler using SDS software. The SNP frequency results are in Appendix Table A; see also the ALFRED database (Cherni et al., 2016; Rajeevan et al., 2012) at https://alfred.med.yale.edu.

The Reference sequence gene of STAT3 is: RefSeqGene (LRG_112) on chromosome 17 (Accession NG_007370, Region: 5001..80171, Version NG_007370.1).
2.3 | Statistical analysis

The analysis of allelic and genotypic frequencies was performed using Plink 1.09 software (Purcell et al., 2007) [http://pngu.mgh.harvard.edu/purcell/plink/], and the calculation of haplotypes has been done with the HAPLO program (Hawley & Kidd, 1995) based on the EM algorithm (Dempster et al., 1977). The determination of linkage disequilibrium (LD)

| SNP alleles | Kairouan | Kerkennah | Kesra | Mahdia | Smar | Sousse | Libya |
|-------------|----------|-----------|-------|--------|------|--------|-------|
| rs3869550   | 2n = 80  | 2n = 86   | 2n = 78 | 2n = 86 | 2n = 116 | 2n = 86 | 2n = 118 |
| C*         | 47 (0.59)| 47 (0.55) | 22 (0.29) | 45 (0.52) | 54 (0.47) | 40 (0.47) | 58 (0.49) |
| T          | 33 (0.41)| 39 (0.45) | 56 (0.71) | 41 (0.48) | 62 (0.53) | 46 (0.53) | 60 (0.51) |
| rs957971   | 2n = 76  | 2n = 84   | 2n = 84 | 2n = 90 | 2n = 122 | 2n = 92 | 2n = 134 |
| C*         | 44 (0.58)| 44 (0.52) | 29 (0.35) | 46 (0.51) | 59 (0.48) | 43 (0.47) | 66 (0.49) |
| G          | 32 (0.42)| 40 (0.48) | 55 (0.65) | 44 (0.49) | 63 (0.52) | 49 (0.53) | 68 (0.51) |
| rs7211777  | 2n = 70  | 2n = 88   | 2n = 84 | 2n = 90 | 2n = 124 | 2n = 92 | 2n = 140 |
| G*         | 39 (0.56)| 49 (0.56) | 29 (0.35) | 45 (0.5)  | 57 (0.46) | 45 (0.49) | 73 (0.52) |
| A          | 31 (0.44)| 39 (0.44) | 55 (0.65) | 45 (0.5)  | 67 (0.54) | 47 (0.51) | 67 (0.48) |

| Population | L1 | L2 | D’  | R²   | Distance basepairs |
|------------|----|----|-----|------|-------------------|
| Kairouan   | rs3869550 | rs957971 | 1.0 | 1.0  | 27,038            |
|            | rs3869550 | rs7211777 | 1.0 | 1.0  | 41,188            |
|            | rs957971  | rs7211777 | 1.0 | 1.0  | 14,150            |
| Kerkennah  | rs3869550 | rs957971 | 1.0 | 1.0  | 27,038            |
|            | rs3869550 | rs7211777 | 0.892 | 0.757 | 41,188            |
|            | rs957971  | rs7211777 | 0.944 | 0.77  | 14,150            |
| Kesra      | rs3869550 | rs957971 | 0.935 | 0.875 | 27,038            |
|            | rs3869550 | rs7211777 | 0.935 | 0.875 | 41,188            |
|            | rs957971  | rs7211777 | 0.947 | 0.897 | 14,150            |
| Mahdia     | rs3869550 | rs957971 | 1.0  | 0.907 | 27,038            |
|            | rs3869550 | rs7211777 | 1.0  | 0.864 | 41,188            |
|            | rs957971  | rs7211777 | 1.0  | 0.957 | 14,150            |
| Smar       | rs3869550 | rs957971 | 0.963 | 0.896 | 27,038            |
|            | rs3869550 | rs7211777 | 0.964 | 0.897 | 41,188            |
|            | rs957971  | rs7211777 | 1.0  | 0.936 | 14,150            |
| Sousse     | rs3869550 | rs957971 | 1.0  | 0.952 | 27,038            |
|            | rs3869550 | rs7211777 | 0.951 | 0.905 | 41,188            |
|            | rs957971  | rs7211777 | 1.0  | 0.915 | 14,150            |
| Libya      | rs3869550 | rs957971 | 1.0  | 0.932 | 27,038            |
|            | rs3869550 | rs7211777 | 0.931 | 0.866 | 41,188            |
|            | rs957971  | rs7211777 | 1.0  | 0.942 | 14,150            |
| Africa     | rs3869550 | rs957971 | 1.0  | 0.976 | 27,038            |
|            | rs3869550 | rs7211777 | 0.993 | 0.868 | 41,188            |
|            | rs957971  | rs7211777 | 1.0  | 0.859 | 14,150            |
| Europe     | rs3869550 | rs957971 | 1.0  | 0.996 | 27,038            |
|            | rs3869550 | rs7211777 | 1.0  | 0.996 | 41,188            |
|            | rs957971  | rs7211777 | 1.0  | 1.0   | 14,150            |
between the studied SNPs was performed with Haploview Software for all the North African populations.

For comparative analysis, we included data from 59 populations from the Kidd Lab (Brissenden et al., 2015; Cherni et al., 2016) and the 26 worldwide populations from The 1000 Genomes Project (1KG) (Consortium, 2015). The haplotypic data of the three SNPs were downloaded from LD link website of the 1000 Genome Project (Machiela & Chanock, 2015). Data obtained from the 7 North African populations analyzed in this study were merged with data from the 1KG project subset (Table S1: Population list file, supplementary material). Principal Component Analysis (PCA) was performed with PAST software (Hammer et al., 2001).

3.2 | Linkage disequilibrium

We compared the linkage disequilibrium (LD) structure of STAT3 SNP (rs3869550, rs957971, rs7211777) among the studied populations.

Linkage disequilibrium among the three SNPs of the STAT3 gene in North African populations was also compared to European populations and African populations in Table 2 which illustrates $r^2$ and D’ values for each pair of SNPs in these populations. Taken together, results revealed a strong linkage disequilibrium among the 3 SNPs. The lower value of D’ between rs3869550 and rs957971 was observed in the Kesra population which is the only population considered as Berber among the studied populations. We also noticed that the population of Mahdia displayed a total LD similar to European populations while the remaining North African populations studied presented similar LD to African populations.

Linkage disequilibrium is very strong among the 3 SNPs considered pairwise. In Table 2, the $r^2$ measurements are in the range of 0.85 to 1.00. The high degree of similarity of the SNP frequencies can be seen visually in supplementary Figure (SF1) or by inspecting the SNP frequencies in supplementary Table S2. Overall, the LD is very high and the deviations from complete (D’ = 1.0) LD is attributable to rare haplotypes (Figure 3).

3.3 | Analysis of haplotype frequencies

The three STAT3 SNPs can occur in eight possible combinations. Direct gene counting evidence supports the occurrence of all eight haplotypes among the populations sampled from around the world. As shown in Figure 3 and Table S3 (Supplementary Data), there are, in the 92 populations (>5600 individuals) studied, two haplotype alleles CCG, the ancestral allele, and TGA, the fully derived allele, that occur at very common frequencies worldwide. The other six haplotypes usually occur at low to rare frequencies (<5%) but in some populations, they do occur at moderately common frequencies (5%–21%).

The ancestral allele occurs at the highest frequencies (often at 80%–90%) in sub-Saharan Africa and in some populations in the Pacific region. The fully derived allele is found at very common frequencies (>90%) in Native American populations. In other world regions (North Africa, Europe, and Asia) the TGA and CCG haplotypes both occur at very common frequencies with TGA usually being the more frequent allele.

In Figure 4, a Network summarizes this relationship between the haplotypes constructed by these three SNPs of the STAT3 gene.
Considering the functional aspects of the studied SNPs that are located at the intronic level, we hypothesized that sites containing these SNPs could be the target of microRNAs with a possible effect on the splicing or stability of STAT3 mRNA. So, we sifted the databases for possible miRs that could target these regions at the level of human STAT3 gene. No target miR was identified for rs3869550 and rs957971, but the derived allele of rs7211777 was targeted by a miR centric...
hsa-miR-3606-5p when G is replaced by A with a score of 65 which can be significant (Figure 5).

4 | DISCUSSION

STATs are ligand-induced transcriptional factors that are activated in response to a wide range of cytokines, growth factors, and hormones. STAT3 is constitutively activated in various cancers including breast cancer. The association of STAT3 polymorphism and cancer in the Tunisian population is still unknown.

The importance of this transcription factor and its involvement in various biological processes and different types of pathologies justifies its attention from the point of view of its activity, which is often a function of genetic polymorphism. STAT3 SNP data in human populations show frequency differences that need to be clarified especially for North African populations on which little data is available.

Our results on the genetic diversity of STAT3, considering 3 SNPs associated with cancer in populations, show the distinctiveness of Sub-Saharan Africans and Native American populations from each other and from populations in other world regions. Although there is very strong linkage disequilibrium present among these STAT3 SNPs in the many populations examined, the allele and haplotype frequency levels do vary around the world. Two of the eight possible haplotypes (the ancestral CCG and the fully derived TGA) are observed to occur at predominant frequencies in the 92 populations studied. This could have arisen by random genetic drift or it could be related to positive selection but studies differentiating among these possibilities have not yet been carried out. However, to the extent that beneficial or harmful genetic variants for the development of various types of cancers or for the response to therapeutic interventions exist in linkage disequilibrium with these observed haplotypes, it is clear from the genetic variation demonstrated in this report that we should expect that population differences will be observed.

Our results on the genetic diversity of STAT3 showed a high level of diversity in North African populations with seven haplotypes observed. In the PCA plot, North African populations resemble South Asians populations and occupy an intermediate position between Sub-Saharan African and the rest of worldwide populations, with a particular behavior of Berber population from Kesra which was close to Europeans, and the island of Kerkennah population which was isolated (Supplementary Figure SF2). This feature seems to be related to the presence of rare STAT3 haplotypes such CCA and TCA which were specific to North African populations. Two rare haplotypes CCG and TGG could have been generated from the ancestral haplotype CCG by point mutation or recombination, then the four others (CGA, TCG, CCA, TCA) could be mostly obtained by recombination between the two major haplotypes (CCG and TGA), excepted for TCA haplotype that could have been generated by recombination between the two rare haplotypes (Figure 6). One has to ask for the cause of such haplotype diversity in North-African populations. Since the two major haplotypes are present in all human populations, heterozygous genotypes CCG/TGA could lead to the possibility of recombination.

Moreover, the presence of recombinant specific haplotypes seems to be characteristic of the North African populations studied. This has been reported for other genes, such as BRCA1, in Tunisian breast cancer patients that displayed several distinct SNP haplotypes, corresponding to different evolution forms, which were less numerous than haplotypes observed in US patients (Troudi et al., 2007). In fact, the American melting pot is recent compared to the very ancient admixture that occurred in North Africa and shown by genetic analysis of actual populations (Ennafaa et al., 2011; Frigi et al., 2010) and also genome analysis of Neolithic fossils (Fregel et al., 2018). Indeed, according to these studies, four components of diverse origins (Sub-Saharan, European, Middle Eastern, and North-African) have been found to be present in North African genomes since at least the Neolithic period. This high level of admixture from prehistoric times should have given possibilities for recombination between distinct haplotypes, leading to new combinations. These conditions may not have been met in other ancient human communities, whose low numbers would have generated a tendency to consanguinity and homozygosity. Considering the pleiotropic role of STAT3 as a transcription factor particularly involved in inflammation and immune response on one hand and the impact of the studied SNPs at the functional level, we can argue that positive selection should have played a role out of Africa, when human migrants settled in a new infectious environment toward which their immune system
**Table 3** Targeted genes by miR-3606-5p in breast cancer (BRCA type) according to ONCOMIR

| Gene     | Gene Description                                                                 | Correlation | Correlation P-value | Correlation FDR  | miRDB Score |
|----------|----------------------------------------------------------------------------------|-------------|---------------------|------------------|-------------|
| MKL2     | MKL/myocardin-like 2                                                             | -0.0824     | 1.65e-02            | 9.46e-01         | 53          |
| RMND1    | required for meiotic nuclear division 1 homolog (S. cerevisiae)                  | -0.0822     | 1.69e-02            | 9.46e-01         | 83          |
| NFAT5    | nuclear factor of activated T-cells 5, tonicity-responsive                       | -0.0763     | 2.65e-02            | 9.46e-01         | 73          |
| MEF2D    | Myocyte Enhancer factor 2D                                                        | -0.0729     | 3.41e-02            | 9.46e-01         | 90          |
| EDNRB    | Endothelin Receptor type B                                                        | -0.0720     | 3.64e-02            | 9.46e-01         | 58          |
| MED12L   | Mediator complex subunit 12-like                                                 | -0.0697     | 4.28e-02            | 9.46e-01         | 60          |

**Table 4** Linkage disequilibrium between pairs of the SNPs (rs3869550, rs957971, rs7211777, rs3736164, rs4796793) of the STAT3 gene worldwide populations

| Rs number | rs3869550 | rs957971 | rs7211777 | rs3736164 | rs4796793 |
|-----------|-----------|----------|-----------|-----------|-----------|
| D'        | 1.0       | 0.999    | 0.983     | 1.0       | 0.996     |
| rs3869550 |           |          |           |           |           |
| rs957971  | 0.999     | 1.0      | 1.0       | 1.0       | 0.996     |
| rs7211777 | 0.983     | 1.0      | 1.0       | 1.0       | 0.996     |
| rs3736164 | 1.0       | 1.0      | 1.0       | 1.0       | 0.996     |
| rs4796793 | 0.996     | 0.996    | 0.996     | 0.996     | 1.0       |

| R²        | 1.0       | 0.98     | 0.96      | 0.623     | 0.557     |
|-----------|-----------|----------|-----------|-----------|-----------|
| rs3869550 |           |          |           |           |           |
| rs957971  | 0.98      | 1.0      | 0.976     | 0.635     | 0.567     |
| rs7211777 | 0.96      | 0.976    | 1.0       | 0.619     | 0.553     |
| rs3736164 | 0.623     | 0.635    | 0.619     | 1.0       | 0.894     |
| rs4796793 | 0.557     | 0.567    | 0.553     | 0.894     | 1.0       |
had to adapt. One has to ask how the location of these SNPs in STAT3 introns might impact the function of the protein. Analysis performed using micro-RNA databases allowed to assess the possibility of targeting these SNPs regions by specific miR. Our results showed the STAT3 region with the derived allele of rs7211777 (G>A) was targeted by miR hsa-miR-3606-5p.

The previous study of STAT3 polymorphism rs3869550 of Jiang et al. conducted on (NSCLC) non-small cell lung cancer showed that the STAT3 protective haplotype GCCGCC contains the ancestral allele (G) instead of the derived allele A (Jiang et al., 2011). Analysis of haplotypes deduced by Vaclavicek et al (Vaclavicek et al., 2007) that the rare haplotype CAGCC which contained the derived allele from each SNP (STAT3 rs721177 and STAT5B rs6503691), was associated with an increased risk of Breast Cancer (OR = 5.83, 95% CI 1.51–26.28, p = .002).

Interestingly, several genes are targeted by miR-3606-5p in breast cancer according to ONCOMIR (Table 3). Expression of hsa-miR-3606-5p has been quantified in normal and breast cancer tissues (TCGA.BRCA.sampleMap/miRNA HiSeq gene database). According to these results, miR-3606-5p might explain the association of rs7211777 at STAT3 gene with Breast cancer. This hypothesis should be confirmed particularly in North African populations where the risk allele (A) is associated with different STAT3 haplotypes and also by assessment of STAT3 mRNA expression and mir-3606-5p in normal and pathological conditions along with STAT3 SNP genotypes.

Moreover, miR-3606-5p is not the only one that regulates STAT3, it is also regulated by other miRs. For example, Mir-520 blocks the progression of EMT by targeting STAT3, in addition, mir-544 inhibits Bcl6 and STAT3, it is also regulated by other miRs. For example, Mir-3606-5p.

Moreover, rs7211777, as located in intron 1, is close to exon 1 and the promoter region. Indeed, the investigation about LD with two other common SNPs in the STAT3 promoter region (rs 3736164, rs4796793) revealed a strong linkage disequilibrium among the 3 SNPs in worldwide populations (LD Supplementary Data). Table 4 illustrates r² and D’ values for each pair of SNPs. Hence this strong LD does not exclude that another mechanism could be associated with the regulation of STAT3 haplotypes expression at the transcriptional level due to functional variants affecting the promoter region of the gene. Indeed, the rs4796793 SNP was previously shown to be associated with cervical cancer, women with a G allele at rs4769793 being submitted to a higher risk for cervical cancer (K. Wang et al., 2011).

In conclusion, previous research has shown that polymorphisms at the STAT3 gene appear to have functional effects on the development and pathological course of various cancers. Assessment of such effects should be investigated in North African populations, considering the presence of specific recombinant STAT3 haplotypes.

ACKNOWLEDGMENTS
Special thanks go to the thousands of individuals around the world who volunteered to give blood and saliva samples to make this study possible and all the authors who contributed to this study.

CONFLICT OF INTEREST
The authors declared no conflict of interest.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID
Wafa Ziadi https://orcid.org/0000-0002-1174-9469
Sarra Elkamel https://orcid.org/0000-0002-1110-5225

REFERENCES
Aggarwal, B. B., Kunnumakkara, A. B., Harikumar, K. B., Gupta, S. R., Tharakan, S. T., Koca, C., Dey, S., & Sung, B. (2009). Signal transducer and activator of transcription-3, inflammation, and cancer: how intimate is the relationship? Annals of the New York Academy of Sciences, 1171, 59–76. https://doi.org/10.1111/j.1749-6632.2009.04911.x
Ben Jmaa, M., Abida, O., Fakhfakh, R., Bahliou, E., Sellami, K. H., Gaddour, L., Elloumi, N., Ben Ayed, M., Masmoudi, A., Dhouib, M., Abdelmoula, M., Mahfoudh, N., Makni, H., Turki, H., & Masmoudi, H. (2018). Involvement of the IL23/Th17 pathway in the pathogenesis of Tunisian pemphigus foliaceus. Mediators of Inflammation, 2018, 1–12. https://doi.org/10.1155/2018/8206983
Brisenden, J. E., Kidd, J. R., Esvaanaa, B., Togtokh, A. J., Pakstis, A. J., Friedlaender, F., Kidd, K. K., & Roscoe, J. M. (2015). Mongolians in the genetic landscape of Central Asia: Exploring the genetic relations among Mongolians and other world populations. Human Biology, 87(2), 73–91. https://doi.org/10.13110/humanbiology.87.2.0005
Bromberg, J. F., Wrzeszynska, M. H., Devgan, G., Zhao, Y., Pestell, R. G., Albanese, C., & Darnell, J. E. Jr. (1999). Stat3 as an oncogene. Cell, 87(3), 295–303. https://doi.org/10.1016/s0092-8674(00)81959-5
Calon, A., Espinet, E., Palomo-Ponce, S., Taurello, D. V. F., Iglesias, M., Céspedes, M. V., Sevillano, M., Nadal, C., Jung, P., Zhang, X.-F., Byrom, D., Riera, A., Rossell, D., Mangues, R., Massagüé, J., Sancho, E., & Batlle, E. (2012). Dependency of colorectal cancer on a TGF-β-driven program in stromal cells for metastasis initiation. Cancer Cell, 22(5), 571–584. https://doi.org/10.1016/j.ccr.2012.08.013
Cherni, L., Pakstis, A. J., Boussetta, S., Elkamel, S., Frigi, S., Khodjet-Al-Khil, H., Barton, A., Haigh, E., Speed, W. C., Ben Ammar Elgaied, A., Kidd, J. R., & Kidd, K. K. (2016). Genetic variation in Tunisia in the context of human diversity worldwide. American Journal of Physical Anthropology, 161(1), 62–71. https://doi.org/10.1002/ajpa.23008
Clevenger, C. V. (2004). Roles and regulation of stat family transcription factors in human breast cancer. American Journal of Pathology, 165(5), 1449–1460. https://doi.org/10.1016/s0002-9440(10)63403-7
Priceman, S. J., Kujawski, M., Shen, S., Cherryholmes, G. A., Lee, H., Zhang, C., Kruper, L., Mortimer, J., Jove, R., Riggs, A. D., & Yu, H. (2013). Regulation of adipose tissue T cell subsets by Stat3 is crucial for diet-induced obesity and insulin resistance. *Proceedings of the National Academy of Sciences of the United States of America*, 110(32), 13079–13084. https://doi.org/10.1073/pnas.1311557110

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., De Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81(3), 559–575. https://doi.org/10.1086/519795

Rajeevan, H., Soudararajan, U., Kidd, J. R., Pakstis, A. J., & Kidd, K. K. (2012). ALFRED: an allele frequency resource for research and teaching. *Nucleic Acids Research*, 40(Database issue), D1010–1015. https://doi.org/10.1093/nar/gkr924

Rane, S. G., & Reddy, E. P. (2000). Janus kinases: Components of multi-Priceman, S. J., Kujawski, M., Shen, S., Cherryholmes, G. A., Lee, H., Zhang, C., Kruper, L., Mortimer, J., Jove, R., Riggs, A. D., & Yu, H. (2013). Regulation of adipose tissue T cell subsets by Stat3 is crucial for diet-induced obesity and insulin resistance. *Proceedings of the National Academy of Sciences of the United States of America*, 110(32), 13079–13084. https://doi.org/10.1073/pnas.1311557110

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., De Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81(3), 559–575. https://doi.org/10.1086/519795

Rajeevan, H., Soudararajan, U., Kidd, J. R., Pakstis, A. J., & Kidd, K. K. (2012). ALFRED: an allele frequency resource for research and teaching. *Nucleic Acids Research*, 40(Database issue), D1010–1015. https://doi.org/10.1093/nar/gkr924

Rane, S. G., & Reddy, E. P. (2000). Janus kinases: Components of multi-signaling pathways. *Oncogene*, 19(49), 5662–5679. https://doi.org/10.1038/sj.onc.1203925

Sansone, P., Storci, G., Tavolari, S., Guarnieri, T., Giovanni, C., Taffurelli, M., Ceccarelli, C., Santini, D., Paterini, P., Marcu, K. B., & Chiec, P. (2007). IL-6 triggers malignant features in mammospheres from human ductal breast carcinoma and normal mammary gland. *J Clin Invest*, 117(12), 3988–4002. https://doi.org/10.1172/jci32533

Schroeder, A., Herrmann, A., Cherryholmes, G., Kowolik, C., Buettner, R., Pal, S., Yu, H., Müller-Newen, G., & Jove, R. (2014). Loss of androgen receptor expression promotes a stem-like cell phenotype in prostate cancer through STAT3 signaling. *Cancer Research*, 74(4), 1227–1237. https://doi.org/10.1158/0008-5472.can-13-0594

Troudi, W., Uhrhammer, N., Sibille, C., Dahan, C., Mahfoudh, W., Bouchlaka Souissi, C., Jalabert, T., Chouchane, L., Bignon, Y. J., Ben Ayed, F., & Ben Ammar Elgaaied, A. (2007). Contribution of androgen receptor expression promotes a stem-like cell phenotype in prostate cancer through STAT3 signaling. *Cancer Research*, 74(4), 1227–1237. https://doi.org/10.1158/0008-5472.can-13-0594

Vaclavicek, A., Bermejo, J. L., Schmutzler, R. K., Sutter, C., Wappenschmidt, B., Meindl, A., Kiechle, M., Arnold, N., Weber, B. H. F., Niederacher, D., Burwinkel, B., Bartram, C. R., Hemminki, K., & Försti, A. (2007). Polymorphisms in the Janus kinase 2 (JAK2)/signal transducer and activator of transcription (STAT) genes: putative association of the STAT gene region with familial breast cancer. *Endocrine-Related Cancer*, 14(2), 267–277. https://doi.org/10.1677/erc-06-0077

Velayos, T., Martínez, R., Alonso, M., Garcia-Etxebarria, K., Aguayo, A., Camarero, C., Urrutia, I., de LaPiscina, I. M., Barrio, R., Santin, I., & Cañete, L. (2017). An activating mutation in STAT3 results in neonatal diabetes through reduced insulin synthesis. *Diabetes*, 66(4), 1022–1029.

Wang, K., Zhou, B., Zhang, J., Xin, Y., Lai, T., Wang, Y., Hou, Q., Song, Y., Chen, Y., Quan, Y. I., Xi, M., & Zhang, L. (2011). Association of signal transducer and activator of transcription 3 gene polymorphisms with cervical cancer in Chinese women. *DNA and Cell Biology*, 30(11), 931–936. https://doi.org/10.1089/dna.2010.1179

Wang, N., Wei, L., Huang, Y., Wu, Y., Su, M., Pang, X., Ji, F., Zhong, C., Chen, T., & Li, B. (2017). miR520c blocks EMT progression of human breast cancer cells by repressing STAT3. *Oncology Reports*, 37(3), 1537–1544. https://doi.org/10.3892/or.2017.5393

Watson, C. J. (2001). Stat transcription factors in mammary gland development and tumorigenesis. *Journal of Mammary Gland Biology and Development*, 6(1), 115–127. https://doi.org/10.1023/a:1009524817155

Wu, H., Huang, M., Cao, P., Wang, T., Shu, Y., & Liu, P. (2012). MiR-135a targets JAK2 and inhibits gastric cancer cell proliferation. *Cancer Biology & Therapy*, 13(5), 281–288. https://doi.org/10.4161/cbt.18943

Yang, J., Chatterjee-Kishore, M., Staugaitis, S. M., Nguyen, H., Schlessinger, K., Levy, D. E., & Stark, G. R. (2005). Novel roles of unphosphorylated STAT3 in oncogenesis and transcriptional regulation. *Cancer Research*, 65(3), 939–947.

Zhao, H., Wang, Z., Wu, H., Xiao, Q., Yao, W., Wang, E., Liu, Y., & Wei, M. (2015). STAT3 genetic variant, alone and in combination with STAT5b polymorphism, contributes to breast cancer risk and clinical outcomes. *Medical Oncology*, 32(1), 375. https://doi.org/10.1007/s12032-014-0375-z

Zhu, Z., Wang, S., Zhu, J., Yang, Q., Dong, H., & Huang, J. (2016). MicroRNA-544 down-regulates both Bcl6 and Stat3 to inhibit tumor growth of human triple negative breast cancer. *Biological Chemistry*, 397(10), 1087–1095. https://doi.org/10.1515/hsz-2016-0104

**SUPPORTING INFORMATION** Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Ziadi, W., Boussetta, S., Elkamel, S., Pakstis, A. J., Kidd, K. K., Medimegh, I., Ben Ammar Elgaaied, A., & Cherni, L. (2021). *STAT3* polymorphisms in North Africa and its implication in breast cancer. *Molecular Genetics & Genomic Medicine*, 9, e1744. https://doi.org/10.1002/mgg3.1744