SOX2 expression in the pathogenesis of premalignant lesions of the uterine cervix: its histo-topographical distribution distinguishes between low and high grade CIN

Jobran M. Moshi
Maastricht University Medical Center
Monique Ummelen
Maastricht University Medical Center
Jos L.V. Broers
Maastricht University Medical Center
Frank Smedts
Cork University Hospital
Department Pathology
Maastricht University Medical Center
Jack P.M. Cleutjens
Maastricht University Medical Center
Rogier J.N.T.M. Litjens
Laurentius Hospital Roermond
Frans C.S. Ramaekers
Maastricht University Medical Center
Anton H. N. Hopman (hopman@maastrichtuniversity.nl)
Maastricht University Medical Center

Research Article

Keywords: Cervical preneoplasia, CIN, squamous intraepithelial lesions, SOX2 distribution, genetic aberrations, HPV infection.

Posted Date: April 19th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1560634/v1

Abstract

Aims: SOX2 expression was evaluated in cervical intraepithelial neoplastic lesions (CIN1, 2 and 3). The SOX2 distribution patterns in the epithelial compartments were correlated to their genetic make-up and presence of HPV copies.

Methods and Results: SOX2 expression levels and histological distribution patterns were studied in normal squamous epithelium and in p16 positive CIN lesions. Copy numbers of the *SOX2*, *SOX17* and *TERC* genes and of chromosomes 1, 3q and 7, as well as HPV genotypes and viral physical status were correlated with SOX2 distribution patterns.

In general SOX2 expression increased with severity of CIN, but a specific expression pattern was found in CIN3. This pattern was characterized by absence of SOX2 in the basal compartment of the epithelium and variable levels in the intermediate and superficial compartments. This SOX2 staining pattern is significantly associated with CIN3 (p=0.004), not found in CIN1 and seen only in a small fraction of the CIN2 lesions. The chromosome ploidy and gene copy numbers in the HPV positive basal cell compartments of CIN3 were normal, in contrast to CIN2 where aneusomic cells were detected in these layers.

Conclusions: Based on the SOX2 immunostaining pattern CIN1 and CIN2 could be delineated from CIN3. These data shed new light on the pathogenesis and dynamics of progression in premalignant cervical lesions, as well as on the target cells for HPV infection. Sox2 staining patterns may become a useful marker to distinguish CIN3 from low grade CIN particularly in diagnostically challenging cases.

Introduction

SOX2 is highly expressed in pluripotent cells of the inner cell mass of a developing embryo, but SOX2 expression is also reported to be correlated with carcinogenesis, chemoresistance and maintenance of the stem cell-like phenotype in cancer cells[1]. Increased expression of SOX2 has been demonstrated in a range of epithelial and non-epithelial malignancies, often correlating with adverse prognostic factors, poor patient outcomes and resistance to therapies. SOX2 is involved in the tumorigenesis of amongst others squamous cell carcinoma of skin, vulvar carcinoma, gastric cancer, glioblastoma, colorectal cancer, lung cancer, oral squamous cell carcinomas and breast cancer[2–7]. In addition, SOX2 is highly expressed in premalignant lesions, such as squamous dysplasia and carcinoma in situ of the lung[8].

Recently we reported that SOX2 expression in the normal uterine cervix is restricted to the squamous epithelium and is not found in normal endocervical columnar cells or reserve cells[9]. When HPV infects the basal cells of (metaplastic) squamous epithelium or reserve cells underlying the columnar epithelium in the transformation zone, SOX2 positive squamous cervical intraepithelial neoplastic (CIN) lesions may arise. Although the data in the literature are contradictory, increasing SOX2 expression has been described along the spectrum going from normal squamous epithelium, via CIN to invasive squamous

cell carcinoma (SCC). Furthermore, SOX2 positive cells derived from cervical cancers show cancer stem cell characteristics[10, 11]. Finally, SOX2 can promote the proliferation, clonogenicity and tumorigenicity of cervical cancer cells in vitro[12]. Taken together, these findings suggest that SOX2 plays a role not only in the initial stages of cervical carcinogenesis but also in the critical step from CIN to invasive carcinoma[10, 11].

Several studies on cervical cancer have described instability of the chromosomal 3q26 region, containing both the *SOX2* and *TERC* gene. The amplification or gain of this specific chromosomal region is now considered a marker of carcinogenesis in CIN and invasive SCC.

In this study we focus on the distribution pattern of SOX2 immunostaining and on its intensity of expression, and correlate this with the genetic make-up and viral load in the different stages of CIN[13–15].

Materials And Methods

Tissue material

Tissue samples obtained during colposcopy were selected from the archives of the Pathology Departments of the Maastricht University Medical Center, Maastricht and the Reinier de Graaf Hospital Delft, the Netherlands.

Formalin fixed and paraffin embedded (FFPE) tissues from 53 patients were available. In these samples 25 areas were selected with low grade squamous intraepithelial lesions (LSIL; 21 areas of CIN1), and 60 areas of high grade SIL (HSIL; 4 areas of CIN 1–2, 26 areas of CIN2, 16 areas of CIN2-3 and 38 areas of CIN3). The notation CIN1-2 and CIN2-3 refers to areas in which we could not unequivocally distinguish between CIN1 or CIN2, and CIN2 or CIN3, respectively. Corresponding tissue was available for PCR HPV typing.

Research on tissue and cell samples has been performed in accordance with the Code for Proper Secondary Use of Human Tissue in the Netherlands (http://www.federa.org/) and has been approved by the board of the Maastricht Pathology Tissue Collection at the Maastricht University Medical Centre (Registration Number MPTC 2011-05).

Immunohistochemistry, in situ hybridization and quantification

Immunohistochemical staining of 4 µm thick FFPE tissue sections was performed using primary antibodies against SOX2, p16 and Ki-67 (for detailed information on antibody characteristics and detection methods see Supplemental Data, Table S1).

Immunohistochemical results were evaluated by means of bright field microscopy (161 areas in total) or fluorescence microscopy (5 normal areas and 5 areas within 5 samples for CIN1, 2 and 3 each). Whole

slide bright field TIF images were made with a Ventana iScan HT slide scanner (Roche, Ventana Medical Systems, Inc. Tucson, Arizona, USA). The fluorescent images were analysed by confocal laser scanning microscopy (Leica SPE confocal microscope; Amsterdam, the Netherlands) and Image J.

The HPV subtypes were localized in the FFPE tissue sections by means of in situ hybridization (ISH) with DNA probes for HPV 16, 18, 31 or 33 ((PanPath, Budel, The Netherlands). For cytogenetic analysis probes were used for chromosomes 1, 3 and 7, and locus specific probes for the genes *TERC, SOX2* and *SOX17* (for detailed information see Supplemental Data and Table S2) [9, 17–19].

Results

SOX2 distribution patterns in normal squamous epithelia and CIN lesions

In the normal cervix SOX2 is strongly expressed in the basal and parabasal layers of the squamous epithelium (Figs. 1A-E) and is absent in the endocervical columnar cells and in reserve cells (result not shown; see [19]).

In the dysplastic epithelium we distinguished 3 different SOX2 staining patterns (see also Table 1). In CIN1 the expression of SOX2 was found in the basal/parabasal cell layers and reached into the intermediate cell layers (Figs. 1F-J). This pattern was classified as **Pattern 1** SOX2 staining. In CIN2 and CIN3 SOX2 expression reached the superficial layers and intensity of expression increased compared to CIN1. This staining pattern is classified as **Pattern 2** and is seen in 14 out of 26 CIN2 and 14 out of 38 CIN3 lesions (Figs. 1K-N and Figs. 1O-S, respectively). This staining pattern was associated with a strong immunostaining from the basal/parabasal compartment up to the superficial cell layers. In the CIN1-2 and CIN2-3 this pattern was recognized in 1 out of 4 and 9 out of 16 areas, respectively.

Table 1
SOX2 immunostaining and distribution patterns in normal cervical squamous epithelium and premalignant lesions.

Histological Classification		Number of areas	Negative areas	SOX2 distribution patterns		
				Pattern 1	Pattern 2	Pattern 3
	Normal	53	3	50	0	0
LSIL	CIN1	21	2	18	1	0
HSIL	CIN1-2	4	0	3	1	0
	CIN 2*	26	3	4	14	5
	CIN 2–3*	16	0	0	9	7
	CIN 3*	38	0	1	14	23
Different patterns of SOX2 staining observed in normal and cervical interepithelial neoplasia (CIN1-CIN3). SOX2 staining patterns:						
Pattern 1) lower one third: SOX2 seen predominantly in the basal/parabasal layer and weaker in the intermediate layer.						
Pattern 2) Lower two third: SOX2 seen in the basal/parabasal and intermediate layer.						
Pattern 3) Upper one third: SOX2 seen only in the intermediate layer, no or very low expression in basal/ parabasal layer.						
LSIL: low grade squamous intraepithelial lesion, HSIL: high grade squamous intraepithelial lesion.						
* When a premalignant lesion contains an area of Pattern 2 adjacent to Pattern 3, both areas were counted.						

Remarkably, in 5 out of 26 CIN 2 and 23 out of 38 CIN3 the SOX2 staining pattern was weak or no nuclear expression for SOX2 was observed in the basal and parabasal cell layers. In the non-unequivocally distinguishable HSIL in 7 out of 20 areas, this pattern was recognized, all belonging to CIN2-3. The intermediate and superficial cell layers showed a variable SOX2 expression in these cases (Figs. 1T-X). This **Pattern 3** is clearly different from Patterns 1 and 2. The frequency correlation of Patterns 2 and 3 with CIN2 and CIN3 demonstrated that Pattern 3 is strongly associated with CIN3 ($p = 0.004$; Student T-test). In the non-unequivocally CIN2-3 in 7 out of 16 cases pattern 3 was recognized. The distribution of SOX2 staining throughout the dysplastic epithelium is sometimes complex, with a mixed or a discontinuous pattern. A typical example is depicted in Fig. 2 where CIN tissue exhibiting Pattern 2 adjoins tissue with Pattern 3 (see also Supplemental Figure S1).

Quantification of SOX2 expression in the different staining patterns

The immunofluorescence intensity in the different patterns was quantified (Fig. 3) to obtain a more objective interpretation of the SOX2 expression levels in CIN. Table 2 summarizes the distribution of the average SOX2 immunofluorescence intensities (and range) over the different cell layers quantified in the three different SOX2 distribution patterns.

Table 2
Quantitative analysis of SOX2 expression levels in areas with Patterns 1, 2 or 3 as determined by means of fluorescence intensity measurements (Arbitrary Units).

		Basal/Parabasal cell layers	Intermediate cell layers	Superficial cell layers
Pattern 1	Average	341*	11	14
	Range	75–499	3–26	3–35
Pattern 2	Average	1767	264	27
	Range	547–3202	85–496	4–89
Pattern 3	Average	51	498	54
	Range	2–108	62–878	3–143

Fluorescence intensity is measured after an immunohistochemical staining procedure using FITC-tyramide as peroxidase substrate. Per pattern the SOX2 expression intensity is measured in 5 different sections. Per cell layer in total 50 nuclei were measured, nuclei were not overlapping, nuclear truncation was minimal (large nuclear size) and on average 3.5 (3–5 slices per nucleus were used for reconstruction. In total 450 nuclei were reconstructed and measured (*50 nuclei per SOX2 Pattern and per cell compartment). For quantification, the images are captured with the same fluorescence integration time during confocal microscopical imaging. Average fluorescence intensity and the range of fluorescence intensity are expressed in Arbitrary Units. Image J, (NIH, Bethesda, Maryland, USA) was used for further image analysis, processing and merging/stitching of the fluorescent images for reconstruction of the sections[16].

In Patterns 1 and 2 a high fluorescence intensity was measured in the basal/parabasal compartment, while in Pattern 3 the basal/parabasal compartment showed the lowest intensity. In Pattern 3, however, the nuclei in the intermediate layers showed the highest SOX2 expression levels, with a ten times higher fluorescence intensity as compared to that of the basal/parabasal layers.

Correlation of SOX2 distribution patterns with genetic aberrations in CIN lesions

SOX2 (located at 3q26.32) copy number variations were analyzed in 95 tissue areas. In the superficial layers of most CIN1 and CIN2 cases no genetically aberrant cells were detected (see Fig. 4A-C). Most CIN3 lesions, however, showed genetically aberrant cells in the superficial compartment, with chromosomal copy numbers exceeding 2. Typical examples of FISH results in the basal/parabasal layer are shown in Fig. 4D-F, illustrating a disomy for the *SOX2* gene in Pattern 1 (Fig. 4D), aberrant (aneusomic) cells in Pattern 2 (Fig. 4E) and a disomy again in Pattern 3 (Fig. 4F). Table 3 summarizes the

areas in which genetically aberrant cells were detected in the normal squamous epithelium and (pre)malignant lesions by means of FISH targeting the *SOX2* gene copy number. Several chromosomal probe sets were also tested and correlated with the different SOX2 patterns (see Supplemental Materials and Methods, Tables S2 and S3 and Figure S2).

Table 3

Presence of genetically aberrant (aneusomic) cells in the different cellular compartments of premalignant CIN lesions as detected by FISH targeting of *SOX2* copy numbers.

Histological		SOX2	Number of	Size of	Para)basal	Intermediate	Superficial
Classification		Pattern	areas*	aberrant area	cell layers	cell layers	cell layers
	Normal	1	29	absent	N	N	N
LSIL	CIN1	1	16	absent	N	N	N
		2	1	small	A/N	A/N	N
HSIL	CIN1-2	1	2	small	A/N	N	N
	CIN2	1	8	medium	A	A/N	N
		2	5	large	A	A	N
	CIN 2–3	2	4	large	A	A	N
		3	4	large	N	A/N	N
	CIN3	2	8	large	A	A	N
		3	15	large	N	A	A/N

*Not all cases mentioned in Table 2 were available for FISH analysis. Only the number of analyzed areas is indicated. N: Normal (Disomy; copy number for *SOX2* = 2); A: Aneusomy (copy number for *SOX2* > 2); A/N: Aneusomy/Normal (Aneusomic cells mixed with disomic cells). LSIL: low grade squamous intraepithelial lesion, HSIL: high grade squamous intraepithelial lesion. For a detailed description of the genetic classification see the Supplemental Materials and Methods section. Note that Pattern 3 strongly deviated from Pattern 2. In all 19 cases the SOX2 negative cells in the basal/parabasal cell layers were disomic for the *SOX2* gene while the SOX2 positive intermediate cell layers showed aneuploidy for both targets in nearly all cases.

Table 4
Presence of genetically aberrant (aneusomic) cells in the different cellular compartments of (pre)malignant CIN lesions as detected by FISH targeting of chromosome 1 centromere (1C) copy numbers.

Histological		SOX2	Number of	Size of	(Para)basal	Intermediate	Superficial
Classification		Pattern	areas*	aberrant area	cell layers	cell layers	cell layers
LSIL	CIN1	1	10	small	N	N	N
		2	1	small	A/ N	A/ N	N
HSIL	CIN1-2	1	3	small	A/ N	N	N
	CIN2	1	14	small	A	A/ N	A/ N
		2	3	medium	A	A	A/ N
	CIN2-3	2	4	large	A	A	A/ N
		3	4	large	N	A	A/ N
	CIN3	2	9	large	A	A	A/ N
		3	12	large	N	A	A/ N

*Not all cases mentioned in Table 2 were available for FISH analysis. Only the number of analyzed areas is indicated. N: Normal (Disomy; copy number for centromere of chromosome 1 (1C) = 2); A: Aneusomy (copy number for centromere of chromosome 1 (1C) > 2); A/N: Aneusomy/Normal (Aneusomic cells mixed with disomic cells). For a detailed description of the genetic classification see the Supplemental Materials and Methods section. All histological normal areas were genetically normal (disomic) and p16 negative. LSIL: low grade squamous intraepithelial lesion, HSIL: high grade squamous intraepithelial lesion.

In summary, disomic cells were detected in all CIN1 lesions (Pattern 1), while genetically aberrant cells were detected in the (para)basal and intermediate compartments of the CIN2 and CIN3 lesions exhibiting Pattern 2. In Pattern 3, occurring in about 20% of the CIN2 lesions and 60% of the CIN3 lesions, an unexpected distribution pattern of genetically aberrant nuclei was observed. In the basal/parabasal compartment the nuclei showed to be disomic (for the mentioned chromosomal probe sets), while in the majority of these cases the intermediate compartment contained nuclei that were genetically aberrant.

Correlation of SOX2 patterns with presence of HPV copies

The physical presence of HPV in the different cellular compartments could be correlated with the different SOX2 patterns, A typical example with a sharp delineation between SOX2 Patterns is shown in Fig. 2. The panel illustrates the positive p16 staining and a high proliferative activity in the basal/parabasal and intermediate cell layers on both sides of the delineation. In the low grade lesions the HPV ISH revealed extensive viral replication (a productive viral infection) in the superficial layer (Fig. 2D and Fig. 5A, B) with a low viral load in the basal/parabasal compartment. In contrast, areas with SOX2 Pattern 3 showed a p16 staining in all cell compartments of the lesion, with no explicit HPV replication pattern in the

superficial layers (Fig. 2D and Fig. 5C). Most strikingly, the genetically normal basal/parabasal cell layers contained a low HPV viral load.

Discussion

Cervical intraepithelial neoplastic (CIN) lesions are SOX2 positive, with increasing expression of this transcription factor upon progression from CIN1 to CIN3 [8,10,11,19]. In the underlying study we report on a discriminating. SOX2 expression pattern in CIN3 which can be clearly distinguished from the SOX2 expression patterns mainly seen in the lower grade preneoplastic lesions (for a schematic representation see Fig. 6). This particular pattern triggered us to more carefully study the molecular characteristics of the different epithelial layers in these high-grade premalignant lesions. On basis of the finding that the basal and parabasal compartments in a majority of CIN3 lesions show a normal genetic make-up, and the observation of direct morphological transitions between immature metaplasia and CIN3 with concomitant molecular switches, we propose three different models for the route of HPV infection, and the origin and progression of CIN lesions.

In the following paragraphs we will discuss different progression models for cervical carcinogenesis (Fig. 7A-D) that can explain the molecular make up of CIN3, taking into account the recent literature on the progression of CIN lesions [20–28].

Model 1: HPV infection of the (para)basal cell layers in normal squamous epithelium or mature metaplasia results in CIN1, which progresses to CIN2 and CIN3 (Fig. 7A, B).

The conventional model for the carcinogenic process in the uterine cervix, as proposed by amongst others Bosch et al[20] and Woodman et al [21], suggests that HPV infection of the basal cells of the squamous epithelium or in mature squamous metaplasia initiates a cascade of events resulting in CIN1 (see also Herfs[22]) which in a minor fraction of patients progresses to invasive squamous cell carcinoma (SCC) via CIN2 and CIN3[23]. Most of the CIN1 lesions have been described to originate from infected (para)basal cells in the ectocervix[24]. Our observation that the SOX2 positive basal cell compartments of normal and mature metaplastic epithelium (Fig. 7A) show a SOX2 expression pattern similar to that of CIN1 (Pattern 1), with subsequent increase of SOX2 positive cell layers with increasing severity of the lesion (Fig. 7B), indeed supports the suggestion that such a sequence of events can take place during progression of cervical preneoplasia. The observation that some of the CIN3 lesions show a SOX2 expression pattern similar to that of most CIN2 lesions suggests that at least this part of the high-grade lesions develops through progression from CIN2.

Model 2: HPV infection of the intermediate cell layers in immature metaplasia results in CIN3 (Fig. 7C).

A significant fraction (60%) of the CIN 3 lesions shows the typical SOX2 Pattern 3, often together with Pattern 2, typical for CIN2. Pattern 3 is characterized by a low or no expression of SOX2 in the (para)basal cell compartments and more extensive staining in the intermediate cell layers upwards. The fact that

these (para)basal cell layers are SOX2 negative or weakly positive indicates that they may be part of immature metaplasia[19]. In the underlying study the basal compartments in a high frequency of CIN3 lesions were shown to exhibit a normal genetic make-up as defined by ploidy for chromosomes 1 and 3, and copy numbers for the SOX2 and SOX17 genes. This suggests that a HPV infection could have occurred in the intermediate cell layers in the normal or metaplastic epithelium in the transformation zone, resulting directly in CIN3, omitting the involvement of lower grade lesions as precursors.

This model is supported by reports in the literature that indicate that CIN3 lesions are rarely preceded by a CIN1 lesion [23, 25]. The majority of metachronous CIN1 and CIN3 lesions, for example, were described to be caused by different HPV genotypes, indicating that a progressive biologic continuum from CIN1 via CIN3, leading finally to cervical cancer may be unlikely in at least part of the cases[26]. The SOX2 immunostaining results for Pattern 3 suggest a relatively rapid development of a high-grade lesion without a well-defined low-grade state. Indeed, Woodman et al[21] have shown that detecting high-grade CIN was maximal 6 to 12 months after first detection of HPV16.

Model 3: HPV infection of the cuboidal squamocolumnar junction cells results in CIN3 (Fig. 7D).

HPV infection of a discrete population of cuboidal squamocolumnar junctional (SCJ) cells with a unique morphology and gene-expression profile, has been suggested by Herfs et al [22] to result in a premalignant squamous cervical lesion. The biomarker expression profile typical for these cells was also detected in a high percentage of high-grade CIN lesions.

Also, this model, suggesting a downward rather than an upward evolution from progenitor cell to the premalignancy, could explain the observation by several authors[23, 25,26] that CIN3 can originate without previous precursor lesions. HPV infection of squamo-columnar junction (SCJ) cells was suggested to result in a trans-differentiation process with an outgrowth of subjacent squamous cells (so-called top-down differentiation) often leading to high grade lesions[27]. In the model as proposed in Fig. 7D the HPV infected SCJ cells overgrow the normal squamous epithelium and initiate a downward proliferation and differentiation to SOX2 negative (immature) squamous metaplasia, that further develops into CIN3, proliferating in an upward direction. This model cannot only explain the normal genetic make-up of the (para)basal cells, but also the aneuploidy and viral load detected in the intermediate cell layers.

Is the absence of SOX2 expression and a normal genetic make-up in the (para)basal cell layers an indication for regression of the CIN3 lesion?

As described by Doorbar et al[29] lesion regression, when it does occur after HPV infection, is not associated with significant apoptosis or cell death. Animal model studies have shown that lesions are cleared by the replacement of actively infected cells with apparently normal basal cells that continue to divide[29]. These histologically normal cells can still contain viral genomes but without obvious viral gene expression. The characteristics of the (para)basal cells in Pattern 3 answer to a certain extent to these

properties in that they are p16 positive, proliferating and show a low HPV load, but exhibit an apparently normal genomic composition.

If Pattern 3 is specific for a regressive process it could be expected that this SOX2 distribution pattern should be detected in the majority of CIN1 lesions. However, the cells targeted by HPV in these low-grade lesions are the basal cells shown to be SOX2 positive, while in the formation of higher grade lesions HPV infection occurs in SOX2 negative immature metaplasia.

We conclude that our data shed new light on the biological characteristics and dynamics in the development of premalignant cervical lesions, as well as on the initial viral infection of the target cells for HPV in the area of the squamo-columnar junction.

Declarations

Acknowledgements

The authors would like to thank Mieke Henfling (Dept. of Molecular Cell Biology, Maastricht University Medical Center, The Netherlands) for assistance with immunohistochemical procedures.

Authorship contributions

J.M.M., A.H.N.H. and F.C.S.R. designed and conceptualized the research. J.M.M., J.L.V.B., J.P.M.C. and M.U. performed and recorded the immunohistochemical and ISH analyses. K.K.V.d.V., R.J.N.T.M.L and F.S. collected and diagnosed tissue samples. J.M.M. wrote the first concept of the article, A.H.N.H. and F.C.S.R. reviewed and edited the subsequent versions. A.H.N.H. and F.C.S.R. coordinated all aspects of this study. All authors have read and agreed to the published version of the manuscript.

Funding declaration

This work was supported by a grant from the Jazan University, Jazan, Saudi Arabia (grant number 1067568103).

Data Availability Statement

The study includes original data, A.H.N. Hopman confirms that he had full access to all the data in the study, and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Competing interest statement

The author(s) declare no competing interests.

References

1. Liu K, Lin B, Zhao M *et al.* The multiple roles for Sox2 in stem cell maintenance and tumorigenesis. Cell Signal 2013;**25**;1264–1271.
2. Hütz K, Mejías-Luque R, Farsakova K *et al.* The stem cell factor SOX2 regulates the tumorigenic potential in human gastric cancer cells. Carcinogenesis 2014;**35**;942–950.
3. Weina K, Utikal J. SOX2 and cancer: current research and its implications in the clinic. Clin Transl Med 2014;**3**;19.
4. Laga AC, Lai CY, Zhan Q *et al.* Expression of the embryonic stem cell transcription factor SOX2 in human skin: relevance to melanocyte and Merkel cell biology. Am J Pathol 2010;**176**;903–913.
5. Chen Y, Shi L, Zhang L *et al.* The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. J Biol Chem 2008;**283**;17969–17978.
6. Du L, Yang Y, Xiao X *et al.* Sox2 nuclear expression is closely associated with poor prognosis in patients with histologically node-negative oral tongue squamous cell carcinoma. Oral Oncol 2011;**47**;709–713.
7. Prince ME, Sivanandan R, Kaczorowski A *et al.* Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. Proc Natl Acad Sci U S A 2007;**104**;973–978.
8. Yuan P, Kadara H, Behrens C *et al.* Sex determining region Y-Box 2 (SOX2) is a potential cell-lineage gene highly expressed in the pathogenesis of squamous cell carcinomas of the lung. PLoS One 2010;**5**;e9112.
9. Hopman ANH, Moshi JM, Hoogduin KJ *et al.* SOX17 expression and its down-regulation by promoter methylation in cervical adenocarcinoma in situ and adenocarcinoma. Histopathology 2020;**76**;383–393.
10. Stewart CJ, Crook M. SOX2 expression in cervical intraepithelial neoplasia grade 3 (CIN3) and superficially invasive (stage IA1) squamous carcinoma of the cervix. Int J Gynecol Pathol 2016;**35**;566–573.
11. Herbert M. Sox2 and Cancer Biology Research Literatures. Cancer Biology 2018;8(4):78–133. http://www.cancerbio.net. doi:10.7537/marscbj080418.10.
12. Chang X, Zhang J, Huang C *et al.* Sex-determining region Y-related high mobility group box (SOX)-2 is overexpressed in cervical squamous cell carcinoma and contributes cervical cancer cell migration and invasion in vitro. Tumour Biol 2015;**36**;7725–7733.
13. Hopman AH, Theelen W, Hommelberg PP *et al.* Genomic integration of oncogenic HPV and gain of the human telomerase gene TERC at 3q26 are strongly associated events in the progression of uterine cervical dysplasia to invasive cancer. J Pathol 2006;**210**;412–419.
14. Rao PH, Arias-Pulido H, Lu XY *et al.* Chromosomal amplifications, 3q gain and deletions of 2q33-q37 are the frequent genetic changes in cervical carcinoma. BMC Cancer 2004;**4**;5.
15. Foster N, Carter S, Ng G, Pett M, Roberts I, Coleman N. Molecular cytogenetic analysis of cervical squamous cell carcinoma cells demonstrates discordant levels of numerical and structural

chromosomal instability and identifies ‘selected’chromosome rearrangements. Cytogenetic and genome research 2009;**127**;9–20.

16. Schindelin J, Arganda-Carreras I, Frise E *et al.* Fiji: an open-source platform for biological-image analysis. Nat Methods 2012;**9**;676–682.
17. Hopman AH, Kamps MA, Smedts F, Speel EJ, Herrington CS, Ramaekers FC. HPV in situ hybridization: impact of different protocols on the detection of integrated HPV. Int J Cancer 2005;**115**;419–428.
18. Hopman AH, Smedts F, Dignef W *et al.* Transition of high-grade cervical intraepithelial neoplasia to micro-invasive carcinoma is characterized by integration of HPV 16/18 and numerical chromosome abnormalities. J Pathol 2004;**202**;23–33.
19. Moshi JM, Hoogduin KJ, Ummelen M *et al.* Switches of SOX17 and SOX2 expression in the development of squamous metaplasia and squamous intraepithelial lesions of the uterine cervix. Cancer Med 2020; 9;6330–6343. doi: 10.1002/cam4.3201.
20. Bosch FX, Broker TR, Forman D *et al.* Comprehensive control of human papillomavirus infections and related diseases. Vaccine 2013;**31**;H1-H31.
21. Woodman CB, Collins SI, Young LS. The natural history of cervical HPV infection: unresolved issues. Nat Rev Cancer 2007;**7**;11–22.
22. Herfs M, Yamamoto Y, Laury A *et al.* A discrete population of squamocolumnar junction cells implicated in the pathogenesis of cervical cancer. Proc Natl Acad Sci U S A 2012;**109**;10516–10521.
23. Mills AM, Paquette C, Terzic T, Castle PE, Stoler MH. CK7 immunohistochemistry as a predictor of CIN1 progression: A retrospective study of patients from the quadrivalent HPV vaccine trials. Am J Surg Pathol 2017;**41**;143–152.
24. Mirkovic J, Howitt BE, Roncarati P *et al.* Carcinogenic HPV infection in the cervical squamo-columnar junction. J Pathol 2015;**236**;265–271.
25. Chen EY, Tran A, Raho CJ, Birch CM, Crum CP, Hirsch MS. Histological 'progression' from low (LSIL) to high (HSIL) squamous intraepithelial lesion is an uncommon event and an indication for quality assurance review. Mod Pathol 2010;**23**;1045–1051.
26. Litjens RJ, Van de Vijver KK, Hopman AH *et al.* The majority of metachronous CIN1 and CIN3 lesions are caused by different human papillomavirus genotypes, indicating that the presence of CIN1 seems not to determine the risk for subsequent detection of CIN3. Hum Pathol 2014;**45**;221–226.
27. Herfs M, Vargas SO, Yamamoto Y *et al.* A novel blueprint for 'top down' differentiation defines the cervical squamocolumnar junction during development, reproductive life, and neoplasia. J Pathol 2013;**229**;460–468.
28. Doorbar J, Egawa N, Griffin H, Kranjec C, Murakami I. Human papillomavirus molecular biology and disease association. Rev Med Virol 2015;25 Suppl 1;2–23. doi: 10.1002/rmv.1822.

Figures

Figure 1

Comparison of SOX2 expression patterns in normal squamous epithelium and CIN1, CIN2 and CIN3 lesions as detected in bright field and fluorescence microscopy.

A-E) Expression of SOX2 in normal squamous epithelium in bright field microscopy (A, B; higher magnification of A in B) and fluorescence microscopy (C-E) with higher magnifications of superficial and basal layer showing no or very weak SOX2 expression in superficial cells (D) and strong expression in basal and parabasal cells (E).

F-J) Expression of SOX2 in a CIN1 lesion with positive staining in basal and parabasal cells and moving gradually into the intermediate cell layer in bright field (F, G; higher magnification of F in G) and fluorescence microscopy (H-J). Higher magnifications clearly show the absence of SOX2 staining in the superficial area (I) and strong positivity in the basal and parabasal cell layer (J). Positivity in H and I on top of superficial cells is caused by a high autofluorescence of erythrocytes.

K-N) Expression of SOX2 in a CIN2 lesion with a strong positive staining throughout the squamous epithelium except for the superficial cells that are weakly positive in bright field (K, L; higher magnification of K in L) and fluorescence microscopy (M, N). The higher magnification in (N) clearly shows the positive staining with a speckled pattern in the basal, parabasal and intermediate cell layers.

O-S) Expression of SOX2 in a CIN3 lesion in bright field microscopy (O, P; higher magnification of O in P) and fluorescence microscopy (Q-S). Higher magnifications show a strong fluorescence in the basal/parabasal layer up to the intermediate layers (R) and a weaker/negative reaction in the more superficial cell layers (S).

T-X) Expression of SOX2 in a CIN3 lesion in bright field microscopy (T, U; higher magnification of T in U) and fluorescence microscopy (V-X), showing a negative reaction in the basal/parabasal and superficial cell layers, while showing SOX2 expression only in intermediate cell layers. Also, higher magnifications in W and X show that most epithelial cells are negative except for cells in the intermediate layers, which are positive with a speckled pattern.

Figure 2

Transition area in a CIN3 lesion exhibiting adjoining Pattern 2 and Pattern 3 SOX2 distributions, with corresponding areas assessed for HPV and the proliferation marker Ki-67.

(A, C, E, F) SOX2 expression level and distribution differences seen in Pattern2 (left side) and Pattern 3 (right side) as visualized in bright field microscopy (A) and fluorescence microscopy (C). Higher

magnifications of the transition zone exhibit a speckled SOX2 pattern in individual cells (E and F).

(B) P16 staining is strong in both Pattern 2 and Pattern 3 with no expression in the superficial layer in Pattern 2.

(D) HPV load and physical status as detected by chromogenic in situ hybridization, showing viral replication in the superficial layer (productive lesion), low viral load in the basal/parabasal nuclei in the Pattern 3 area (see insert).

(G) Ki-67 expression in areas with SOX2 Pattern 2 and Pattern 3. Note the difference in nuclear size and the frequency of Ki-67 positive cells in the two different areas and that both areas harbor HPV.

Figure 3

Quantitative analysis of SOX2 expression in areas with Patterns 1, 2 and 3.

Expression levels of SOX2 were assessed by means of fluorescence intensity measurements in a confocal laser scanning microscope using a line scan (white line) in (A) normal squamous epithelium with a SOX2 Pattern 1, (B) a CIN1 lesion with Pattern 1, (C) a CIN2 lesion with Pattern 2, and (D) a CIN3 lesion with Pattern 3.

In Pattern 1 a high SOX2 fluorescence intensity is measured in the basal/parabasal cell layer only, while in Pattern 2 the fluorescence intensity moves gradually upwards to reach the intermediate cell layer. In Pattern 3 the basal/parabasal layer shows no SOX2 staining.

Figure 4

Fluorescence in situ hybridization showing copy number variations for the *SOX2* (visualized in green) and *SOX17* genes (visualized in red) in areas with SOX2 expression Patterns 1, 2 and 3.

The intermediate and superficial cell layers are shown in A, B and C, while the basal/parabasal cell compartments are depicted in D, E and F. The Pattern 1 area was selected from a CIN1 lesion (A, D), while the Pattern 2 and Pattern 3 areas were from a CIN3 lesion (B, C, E, F). Genetically aberrant cells were mainly detected in areas with SOX2 Pattern 2, while in Patterns 1 and 3 the basal/parabasal cells were disomic. The intermediate layer in Pattern 3 showed genetically aberrant cells in most cases.

Figure 5

Fluorescence in situ hybridization targeting copy number variations for the centromere region of chromosome 1 (C1; visualized in green) and load for HPV 16 (visualized in red).

The intermediate and superficial cell layers are shown in A, B and C, while the basal/parabasal cell compartments are depicted in D, E and F. The Pattern 1 area was selected from a CIN1 lesion (A, D), while the Pattern 2 and Pattern 3 areas were from a CIN3 lesion (B, C, E, F). Genetically aberrant cells were seen in areas with SOX2 Pattern 2, while in Patterns 1 and 3 the basal/parabasal cells were disomic. The intermediate layer in Pattern 3 showed genetically aberrant cells in most cases, with the highest viral load is also seen in the intermediate/superficial cells showing replication of the viral sequences. In Pattern 2 a relatively high copy number for the virus is also seen in the basal/parabasal cells, while in Patterns 1 and 3 a low HPV copy number is seen in these basal/parabasal layers (see insert in F).

Figure 6

Schematic representation of the different SOX2 distribution patterns, with concomitant molecular characteristics, for the normal squamous epithelium and the different stages of cervical premalignant lesions.

Figure 7

Schematic representation of the different models for the origin of preneoplastic CIN lesions after HPV infection (indicated by the arrow) based on the different SOX2 expression patterns and the genetic characteristics.

A) Origin of metaplastic epithelial changes and increasing SOX2 expression with higher degree of maturation.

B) HPV infection of the basal cell layer(s) in the normal or metaplastic epithelium of the transformation zone results in CIN1, that progresses to invasive squamous cell carcinoma via CIN2 and CIN3.

C) HPV infection of the intermediate cell layers in the normal or metaplastic epithelium of the transformation zone resulting directly in CIN3, subsequently progressing to invasive squamous cell carcinoma.

D) HPV infection of the columnar cell layer overgrowing the normal or metaplastic epithelium of the transformation zone (as suggested by Herfs et al[22]) results directly in CIN3, subsequently progressing to invasive squamous cell carcinoma.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- J.Moshietal.SOX2HistochemistryandCellBiologySupplementalData14042022.docx
- supplementalfigure1J.Moshietal.SOX2.png
- supplementalfigure2J.Moshietal.SOX2.png